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

M1 MMA 2023-2024
Optimisation

Feuille de TD n°3

Problème d'optimisation et convexité.

Exercice 1

Soit D la droite de $\mathbb{R}^2$ caractérisée par $(x_1, x_2) \in D$ si et seulement si $x_2 = 0$. On considère le problème d'optimisation

$$(\mathcal{P}) \qquad \inf_D f,$$

où pour tout $(x_1, x_2) \in \mathbb{R}^2$, $f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1 - 2x_2 + 5$.

1. Justifier que f est une fonctionnelle quadratique. Déduire l'unique minimum global de f sur $\mathbb{R}^2$.
2. a) Décrire les lignes de niveau de f.

   b) Conjecturer, à partir d'observations géométriques, l'ensemble des solutions du problème d'optimisation.
3. a) Déterminer la nature du problème d'optimisation $(\mathcal{P})$.

   b) Montrer que $(\mathcal{P})$ admet au moins une solution.

   c) Donner une CNS pour que $x^\star = (x_1^\star, x_2^\star) \in D$ soit une solution de $(\mathcal{P})$. Identifier alors l'ensemble des solutions de $(\mathcal{P})$.

<u>Correction</u>.

1. On peut écrire f sous la forme $f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c$, pour tout $x = (x_1, x_2) \in \mathbb{R}^2$, avec
$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad b = (4, 2), \quad c = 5.$$
Donc f est une fonctionnelle quadratique. A est symétrique définie positive donc f est strictement convexe et f admet un unique point critique : $(2, 1) = A^{-1}b$. Comme f est convexe, l'ensemble de ses minimums globaux est l'ensemble de ses points critiques, donc $(2, 1)$ est l'unique minimum global de f.
2. a) On remarque que pour tout $x \in \mathbb{R}^2$, $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$. Donc les lignes de niveau de f sont les cercles de centre $(2, 1)$.

   b) Soit $\mathcal{C}_{\sqrt{\lambda}}$ le cercle de centre $(2, 1)$ et de rayon $\sqrt{\lambda}$ correspondant à la ligne de niveau λ de f. $\mathcal{C}_{\sqrt{\lambda}}$ intersecte D en deux points distincts si et seulement si $\lambda > 1$, car la distance de $(2, 1)$ à D est 1, et dans ce cas la valeur prise par f en ces points est par définition λ. Pour $\lambda = 1$, on a $\mathcal{C}_1 \cap D = \{(2, 0)\}$, donc comme tout point de D différent de $(2, 0)$ est dans l'intersection d'un $\mathbb{C}_{\sqrt{\lambda}}$ avec D, pour $\lambda > 1$, on déduit que $(2, 0)$ est l'unique minimum global de f sur D. Ainsi
$$\operatorname{argmin}_D f = \{(2, 0)\}.$$
   Voir la Figure 1.
3. a) La fonction f est (strictement) convexe et D est un ensemble convexe. Donc $(\mathcal{P})$ est un problème d'optimisation convexe sous contrainte.

   b) La fonction f est continue, coercive et D est un ensemble fermé, donc on sait que $(\mathcal{P})$ admet au moins une solution (f a au moins un minimum global sur D).

   c) Comme $(\mathcal{P})$ est convexe et sous contrainte, on sait que $x^\star \in D$ est solution de $(\mathcal{P})$ si et seulement si
$$\forall y \in D, \quad \langle \nabla f(x^\star), y - x^\star\rangle \geqslant 0,$$
   Et même ici comme D est un sous espace affine (en fait un sev) de $\mathbb{R}^2$, la condition précédente se réécrit de manière équivalente : $\nabla f(x^\star) \in D^\perp$. Voir la Figure 2 pour une résolution géométrique utilisant cette condition d'orthogonalité. Comme pour tout $x \in \mathbb{R}^2$, on a $\nabla f(x) = 2x - (4, 2)$ et

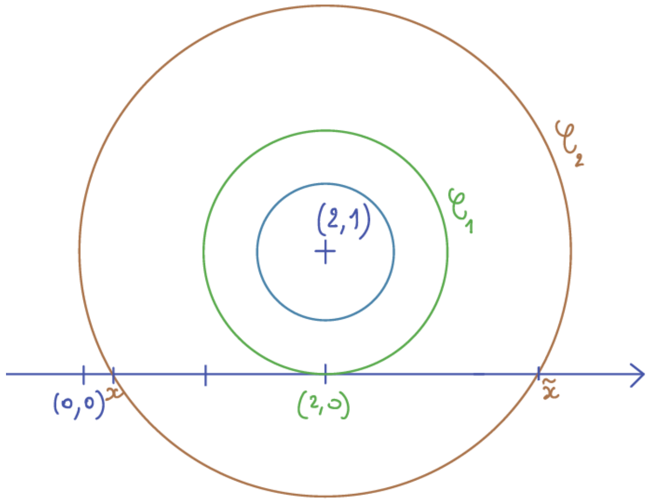

FIGURE 1. On a $f(x) = f(\tilde{x}) > f(2,0)$, donc $(2,0)$ est l'unique minimum global de f sur $D = \{(t,0) : t \in \mathbb{R}\}$.

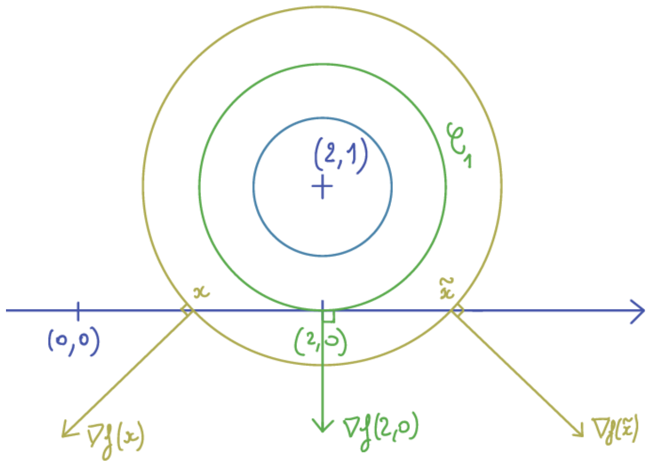

FIGURE 2. Comme pour tout $x^\star$ minimum global de f sur D on doit avoir $\nabla f(x^\star) \in D^\perp$, et que le gradient de f est orthogonal à ses lignes de niveaux, donc aux $\mathcal{C}_{\sqrt{\lambda}}$ pour $\lambda \geqslant 1$, on déduit que la seule possibilité est $x^\star = (2,0)$.

$D^\perp = \{(0,t) : t \in \mathbb{R}\}$, on déduit que $\nabla f(x) \in D^\perp$ si et seulement si $x_1 = 2$. Donc f admet un unique minimum global sur D qui est $(2,0)$.

Exercice 2

Soit $p^{(1)} = (p_1^{(1)}, p_2^{(1)}), p^{(2)} = (p_1^{(2)}, p_2^{(2)}), \dots, p^{(n)} = (p_1^{(n)}, p_2^{(n)})$, n points dans $\mathbb{R}^2$. Le but de cet exercice est de trouver le point qui minimise la somme des distances au carré à tous les points $p^{(i)}$. Ainsi, pour tout $x = (x_1, x_2) \in \mathbb{R}^2$ nous définissons la fonctionnelle suivante

$$J(x) = \sum_{i=1}^{n} \left\| x - p^{(i)} \right\|^2.$$

1. Montrer que J admet un unique point critique $x^\star$ sur $\mathbb{R}^2$. Interpréter géométriquement ce point.
2. Montrer que $x^\star$ est l'unique minimum global de J.

Correction.

1. J est $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ comme somme et composée de fonctions $\mathcal{C}^\infty$ (l'application $x \mapsto \|x\|^2$ est $\mathcal{C}^\infty$). Donc en particulier J est différentiable sur $\mathbb{R}^2$.

   Soit $x \in \mathbb{R}^2$. On a $\nabla J(x) = \sum_{i=1}^n 2(x - p^{(i)})$, donc $\nabla J(x) = 0$ si et seulement si $x = \frac{1}{n} \sum_{i=1}^n p^{(i)}$. Donc J admet un unique point critique qui est le barycentre des points $p^{(i)}$, $i \in \{1, \dots, n\}$.
2. Comme J est $\mathcal{C}^\infty$, J est aussi deux fois différentiable sur $\mathbb{R}^2$ et sa matrice hessienne en tout $x \in \mathbb{R}^2$ est $\nabla^2 J(x) = 2nI_2$ où I_2 est la matrice identité de $\mathcal{M}_2(\mathbb{R})$. En particulier J est strictement convexe

sur $\mathbb{R}^2$. L'ensemble de ses points critique est donc l'ensemble de ses minimums globaux. Donc J admet comme unique minimum global sur $\mathbb{R}^2$ le barycentre des points $p^{(i)}$.

> On vient de montrer que l'on peut définir alternativement le barycentre des points $p^{(i)}$ comme l'unique minimum global de la fonction objectif J. Trouver le barycentre des points $p^{(i)}$ c'est trouver le point x qui minimise la somme des distance euclidienne de x aux $p^{(i)}$. Cette définition équivalente permet de fournir une définition du barycentre plus robuste : il suffit d'avoir un espace métrique pour pouvoir la déployer. Par exemple on peut se demander comment définir un barycentre de n points appartenant à une sphère de sorte que ce barycentre respecte la géométrie de la sphère, c'est à dire appartienne à celle-ci. On remarque que le barycentre classique définie comme la moyenne arithmétique de ces n points n'appartient nécessairement pas à la sphère. Mais la sphère peut être munie d'une structure d'espace métrique en considérant que la distance séparant deux points est la longueur de l'arc du grand cercle les rejoignant [a]. En notant d cette distance, et S la sphère, on peut alors définir le barycentre des points $p^{(i)}$ comme la solution de $\inf_{x\in S} d(x, p^{(i)})$.
>
> a. `https://fr.wikipedia.org/wiki/Distance_du_grand_cercle`

Exercice 3 (Régression linéaire simple)
Soit $((x_i, y_i))_{1\leqslant i\leqslant n}$ un n-uplet de points dans $\mathbb{R}^2$, avec $n \geqslant 2$. On suppose qu'au moins deux points x_i sont distincts.

La régression linéaire simple consiste à chercher la meilleure relation affine du type $y = \alpha x + \beta$ qui relie l'*entrée* x_i à la *sortie* y_i, pour tout $i \in \{1, \ldots, n\}$. Ceci s'obtient en s'intéressant au problème d'optimisation

$$\inf_{(\alpha,\beta)\in\mathbb{R}^2} J(\alpha, \beta),$$

où pour tout $(\alpha, \beta) \in \mathbb{R}^2$, $J(\alpha, \beta) = \sum_{i=1}^{n} (\alpha x_i + \beta - y_i)^2$.

Montrer que J admet un minimiseur global unique et le calculer.

<u>Correction</u>.
La fonction J est $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ car polynomiale en ses arguments.

> On peut résoudre le problème tel qu'il est formulé ici en calculant les dérivées partielles $\partial_\alpha J$ et $\partial_\beta J$ pour obtenir le gradient, puis trouver l'unique point critique en résolvant $\partial_\alpha J(\alpha, \beta) = \partial_\beta J(\alpha, \beta) = 0$, et enfin en montrant que J est convexe en calculant la matrice hessienne de J à partir de son gradient. Je vous conseille de vous entraîner en faisant l'exercice de cette manière.
> Nous allons plutôt choisir ici de réécrire la fonction objectif J de manière vectorielle car cela mène souvent à des calculs plus simples, plus facilement généralisables aux dimensions supérieurs et directement implémentable efficacement dans divers langages informatiques (dont Python) qui sont plus efficaces lorsqu'ils effectuent des opérations vectorielles plutôt qu'à travers des boucles for. Vous pourrez comparer que les deux méthodes sont cohérentes (c'est un bon exercice !).

On peut réécrire J de la forme suivante

$$\forall (\alpha, \beta) \in \mathbb{R}^2, \quad J(\alpha, \beta) = \|\alpha x + \beta \mathbb{1} - y\|^2,$$

où $\mathbb{1} \overset{\text{def.}}{=} (1, \ldots, 1) \in \mathbb{R}^n$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ et $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. En posant A la matrice à 2 colonnes et n lignes définie par $A = (x\ \mathbb{1})$, on a finalement

$$\forall (\alpha, \beta) \in \mathbb{R}^2, \quad J(\alpha, \beta) = \left\| A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - y \right\|^2.$$

On a alors directement pour tout $(\alpha, \beta) \in \mathbb{R}^2$, $\nabla J(\alpha, \beta) = A^T (A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - y)$ (calcul déjà fait plusieurs fois : développer la quantité $J((\alpha, \beta) + (h_1, h_2))$).

> On remarque par une analyse dimensionnelle que $A^T (A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - y)$ est bien une matrice colonne à deux lignes, donc un vecteur de $\mathbb{R}^2$ comme attendu.

Soit $(\alpha, \beta) \in \mathbb{R}^2$, alors $\nabla J(\alpha, \beta) = 0$ si et seulement si $A^T A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = A^T y$. Or

$$A^T A = \begin{pmatrix} \langle x,\, x\rangle & \langle x,\, \mathbb{1}\rangle \\ \langle \mathbb{1},\, x\rangle & \langle \mathbb{1},\, \mathbb{1}\rangle \end{pmatrix} = \begin{pmatrix} \|x\|^2 & \langle x,\, \mathbb{1}\rangle \\ \langle x,\, \mathbb{1}\rangle & n \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

On a donc $\det(A^T A) = n \|x\|^2 - \langle x,\, \mathbb{1}\rangle^2$. Or par l'inégalité de Cauchy-Schwarz : $|\langle x,\, \mathbb{1}\rangle| \leqslant \|\mathbb{1}\| \, \|x\| = \sqrt{n} \, \|x\|$. D'où $\det(A^T A) \geqslant 0$. Et on a égalité dans l'inégalité de Cauchy-Schwarz, i.e. $\det(A^T A) = 0$, si et seulement si x et $\mathbb{1}$ sont positivement liés, i.e. il existe $\lambda \in \mathbb{R}_+$ tel que $x = \lambda \mathbb{1}$. Si c'était le cas alors on aurait $x_1 = x_2 = \ldots = x_n$. C'est défendu par hypothèse de l'énoncé. Par conséquent $\det(A^T A) > 0$, donc $A^T A$ est inversible et J admet un unique point critique : $(A^T A)^{-1} A^T y$.

En différentiant ∇J (on calcule $\nabla J((\alpha, \beta) + (h_1, h_2))$), on obtient $\nabla^2 J(\alpha, \beta) = A^T A$. C'est une matrice symétrique définie positive car c'est une matrice 2×2, son déterminant est strictement positif ainsi que sa trace (nécessairement ses deux valeurs propres sont strictement positives). Donc J est strictement convexe. Par conséquent son unique point critique est un minimum global et c'est le seul, c'est $(A^T A)^{-1} A^T y$.

> Attention si $B \in \mathcal{M}_n(\mathbb{R})$ est une matrice symétrique dont le déterminant et la trace sont strictement positifs, alors on ne peut déduire que B est définie positive lorsque $n > 2$.

Exercice 4 (Modèle linéaire)
Soit $((x_i, y_i))_{1 \leqslant i \leqslant n}$ un n-uplet de points de $\mathbb{R}^2$. On va chercher une manière optimal de représenter la sortie y_i en fonction de l'entrée x_i, pour tout $i \in \{1, \ldots, n\}$, via le modèle défini par

$$\forall x \in \mathbb{R}, \quad f_\beta(x) = \sum_{j=0}^{k} \beta_j w_j(x),$$

où $k \in \mathbb{N}$, et les $w_j : \mathbb{R} \to \mathbb{R}$ sont des fonctions données, et les $\beta_j \in \mathbb{R}$ sont les paramètres cherchés. On s'intéresse au problème d'optimisation

$$(\mathcal{P}) \qquad \inf_{\beta \in \mathbb{R}^{k+1}} J(\beta),$$

où pour tout $\beta \in \mathbb{R}^{k+1}$, $J(\beta) = \sum_{i=1}^{n} \left(f_\beta(x_i) - y_i\right)^2$. On supposera que $n \geqslant k + 1$.

1. Montrer que l'on peut écrire

$$\forall \beta \in \mathbb{R}^{k+1}, \quad J(\beta) = \|M\beta - y\|^2 .$$

où $y = (y_1, \cdots, y_n)$, et M est une matrice à définir.

2. Montrer que $\beta \in \mathbb{R}^{k+1}$ est un point critique de J si et seulement si $M^T M \beta = M^T y$.

3. Montrer que si M est de rang maximal, alors la solution du problème $(\mathcal{P})$ est unique.

<u>Correction</u>.
Cet exercice généralise l'Exercice 3. Les calculs ressemblent donc beaucoup.

1. En définissant la matrice $M \in \mathcal{M}_{n,k}(\mathbb{R})$ dont les lignes sont les $(w_1(x_i)\, w_2(x_i)\, \ldots\, w_k(x_i))$ pour $i \in \{1, \ldots, n\}$, on obtient que pour $\beta \in \mathbb{R}^{k+1}$, $M\beta$ est le vecteur $(f_\beta(x_1), \ldots, f_\beta(x_n)) \in \mathbb{R}^n$, et donc on a bien $J(\beta) = \|M\beta - y\|^2$.

2. J est $\mathcal{C}^\infty$ donc en particulier différentiable et pour tout $\beta \in \mathbb{R}^{k+1}$, $\nabla J(\beta) = M^T M \beta - M^T y$. D'où la réponse.

3. Si M est de rang maximal (i.e. $k+1$ car $n \geqslant k+1$) alors par le théorème du rang on déduit $\mathrm{Ker}(M) = \{0\}$. Donc $\mathrm{Ker}(M^T M) = \{0\}$ puisque si $M^T M \beta = 0$, alors $\beta^T M^T M \beta = 0$, donc $\|M\beta\|^2 = 0$, d'où $M\beta = 0$ i.e. $\beta = 0$. Par conséquent $M^T M$ est inversible car c'est une matrice carré. J admet donc un unique point fixe : $(M^T M)^{-1} M^T y$.

   Il nous reste à montrer que J est convexe pour conclure (puisqu'alors son unique point critique sera son unique minimum global et donc l'unique solution de $(\mathcal{P})$). On a pour tout $\beta \in \mathbb{R}^{k+1}$, $\nabla^2 J(\beta) = M^T M$. Or pour tout $\beta \in \mathbb{R}^{k+1}$, tout $h \in \mathbb{R}^{k+1}$, $\left\langle \nabla^2 J(\beta) h,\, h \right\rangle = \|Mh\|^2 \geqslant 0$. D'où J est convexe sur $\mathbb{R}^{k+1}$ puisque sa matrice hessienne est positive en tout point de $\mathbb{R}^{k+1}$.

> On aurait pu raisonner d'une manière similaire à cette question dans l'Exercice 3 plutôt que d'invoquer l'inégalité de Cauchy-Schwarz. En effet pour montrer que $A^T A$ est inversible, on pouvait remarquer que A est de rang plein (2) car ses deux colonnes ne peuvent être liées (sinon contradiction avec l'hypothèse qu'il existe $i \neq j$ tels que $x_i \neq x_j$), d'où par le raisonnement de cette question $A^T A$ est inversible. Puis directement par le calcul précédent, déduire que $\nabla^2 J(\alpha, \beta)$ est positive pour tout $(\alpha, \beta) \in \mathbb{R}^2$.