DM1 - Optimisation Convexe

Coërchon Colin

14 Octobre 2024

Exercice 0.1 :

Lequel des ensembles suivants est convexe ?

1. Un rectangle, c'est-à-dire un ensemble de la forme :
$$\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, \ldots, n\}.$$
2. L'ensemble hyperbolique :
$$\{x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1\}.$$
3. L'ensemble des points plus proches d'un point donné que d'un ensemble donné, c'est-à-dire :
$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ pour tout } y \in S\},$$
où $S \subseteq \mathbb{R}^n$.
4. L'ensemble des points plus proches d'un ensemble que d'un autre, c'est-à-dire :
$$\{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\},$$
où $S, T \subseteq \mathbb{R}^n$, et $\operatorname{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$.
5. L'ensemble
$$\{x \mid x + S_2 \subseteq S_1\},$$
où $S_1, S_2 \subseteq \mathbb{R}^n$ avec S_1 convexe.

Démonstration.

1. **Oui, l'ensemble est convexe**. Prouvons le.

   Soit $A = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, \ldots, n\}$. On se donne $x, y \in A$ et $\lambda \in [0, 1]$, et on pose $z = \lambda x + (1 - \lambda) y$. On a alors :

$$\begin{aligned}
& \forall i \in [\![1, n]\!], \quad \begin{cases} \alpha_i \leq x_i \leq \beta_i, \\ \alpha_i \leq y_i \leq \beta_i \end{cases} \\
\implies \quad & \forall i \in [\![1, n]\!], \quad \begin{cases} \lambda\alpha_i \leq \lambda x_i \leq \lambda\beta_i, \\ (1-\lambda)\alpha_i \leq (1-\lambda) y_i \leq (1-\lambda)\beta_i \end{cases} \\
\implies \quad & \forall i \in [\![1, n]\!], \quad \lambda\alpha_i + (1-\lambda)\alpha_i \leq \lambda x_i + (1-\lambda) y_i \leq \lambda\beta_i + (1-\lambda)\beta_i \\
\implies \quad & \forall i \in [\![1, n]\!], \quad \alpha_i \leq z \leq \beta_i \\
\implies \quad & z \in A
\end{aligned}$$

On a donc prouvé que A est bien convexe.

2. **Oui, l'ensemble est convexe**. Prouvons le.

   Soit $B = \{x \in \mathbb{R}_+^2 \mid x_1x_2 \geq 1\}$. On se donne $x = (x_1, x_2) \in B, y = (y_1, y_2) \in B$ et $\lambda \in [0,1]$, et on pose $z = \lambda x + (1-\lambda)y$. On a alors :

$$\begin{aligned} z_1z_2 &= (\lambda x_1 + (1-\lambda)y_1)(\lambda x_2 + (1-\lambda)y_2) \\ &= \lambda^2 x_1x_2 + \lambda(1-\lambda)x_1y_2 + \lambda(1-\lambda)x_2y_1 + (1-\lambda)^2 y_1y_2 \\ &\geq \lambda^2 x_1x_2 + (1-\lambda)^2 y_1y_2 && (\text{car } x, y \in \mathbb{R}_+^2 \text{ et } \lambda \in [0,1]) \\ &\geq \lambda^2 + (1-\lambda)^2 && (x, y \in B \text{ donc } x_1x_2 \geq 1 \text{ et } y_1y_2 \geq 1) \\ &\geq \lambda^2 + 1 - 2\lambda + \lambda^2 \\ &\geq 1 + 2\lambda(1-\lambda) \\ &\geq 1 && (\text{car } \lambda(1-\lambda) \geq 0) \end{aligned}$$

   Ainsi, on en déduit que $z \in B$. Et donc, B est convexe.

3. **Oui, l'ensemble est convexe**. Prouvons le.

   Premièrement, on définie les demi-plans par :

$$\forall p, q \in \mathbb{R}^n, \quad H(p,q) = \{x \in \mathbb{R}^n \mid \|x-p\|_2 \leq \|x-q\|_2\}$$

   $\forall p, q \in \mathbb{R}^n, H(p,q)$ est **convexe**. En effet, on peut remarquer que :

$$\begin{aligned} \|x-p\|_2 \leq \|x-q\|_2 &\iff \|x-p\|_2^2 \leq \|x-q\|_2^2 \\ &\iff (x-p)^T(x-p) \leq (x-q)^T(x-q) \\ &\iff x^Tx - 2p^Tx + p^Tp \leq x^Tx - 2q^Tx + q^Tq \\ &\iff 2(q-p)^Tx \leq q^Tq - p^Tp \end{aligned}$$

   On a donc, en posant $a = 2(q-p)^T$ et $b = q^Tq - p^Tp$: $H(p,q) = \{x \in \mathbb{R}^n \mid ax \leq b\}$. Cet ensemble est forcément convexe parce qu'il est défini par une inégalité linéaire.

   Ainsi, en posant $C = \{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \leq \|x-y\|_2 \text{ pour tout } y \in S\}$, on peut écrire :

$$C = \bigcap_{y \in S} H(x_0, y)$$

   C s'écrit comme une intersection d'ensembles convexes, il est donc convexe.

4. **Non, l'ensemble n'est pas forcément convexe**. Prouvons le.

   On peut en effet prendre le contre exemple suivant.
   On se place dans $\mathbb{R}^2$, et on pose : $S = \{x \in \mathbb{R}^2 \mid \|x\|_2 = 1\}$ et $T = \{0\}$.

   On a alors : $D \triangleq \{x \in \mathbb{R}^2 \mid \text{dist}(x,S) \leq \text{dist}(x,T)\} = \left\{x \in \mathbb{R}^2 \,\middle|\, \|x\|_2 \geq \frac{1}{2}\right\}$.

   L'ensemble D est clairement non-convexe. En effet, il suffit de prendre $\lambda = \frac{1}{2}$ avec $x = (-1, 0)$ et $y = (1, 0)$ pour s'en convaincre. On remarque que $x, y \in D$ (car $\in S$) et alors : $\lambda x + (1-\lambda)y = 0$. Et $0 \notin D$.

5. **Oui, l'ensemble est convexe**. Prouvons le.

   On a :
   $$E \triangleq \{x \mid x + S_2 \subseteq S_1\} \underset{\text{def.}}{=} \bigcap_{y \in S_2} \{x \mid x + y \in S_1\} = \bigcap_{y \in S_2} (S_1 - y)$$

   Comme S_1 est convexe par hypothèse, les ensembles $(S_1 - y)$ avec $y \in S_2$ sont donc convexes. Vu comme une intersection d'ensembles convexes, on en conclut que E est convexe.

$\square$

Exercice 0.2 :

Pour chacune des fonctions suivantes, déterminez si elle est convexe, concave ou aucun des deux.

1. $f(x_1, x_2) = x_1 x_2$ sur $\mathbb{R}^2_{++}$.
2. $f(x_1, x_2) = \frac{1}{x_1 x_2}$ sur $\mathbb{R}^2_{++}$.
3. $f(x_1, x_2) = \frac{x_1}{x_2}$ sur $\mathbb{R}^2_{++}$.
4. $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, où $0 \leq \alpha \leq 1$, sur $\mathbb{R}^2_{++}$.

Démonstration.

1. La Hessienne de f est :
$$\nabla^2 f(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
Ces valeurs propres sont 1 et -1. Par conséquent, elle n'est ni semi-définie positive, ni semi-définie négative.

   Ainsi, f **n'est ni convexe ni concave**.

2. La Hessienne de f est :
$$\nabla^2 f(x) = \frac{1}{x_1 x_2} \begin{pmatrix} \frac{2}{x_1^2} & \frac{1}{x_1 x_2} \\ \frac{1}{x_1 x_2} & \frac{2}{x_2^2} \end{pmatrix}$$
Le calcul des mineurs principaux donne $\Delta_1 = \frac{2}{x_1^3 x_2} \geq 0$ et $\Delta_2 = \frac{3}{x_1^4 x_2^4} \geq 0$. La matrice est donc semi-définie positive. Donc f **est convexe**.

3. La Hessienne de f est :
$$\nabla^2 f(x) = \begin{pmatrix} 0 & \frac{-1}{x_2^2} \\ \frac{-1}{x_2^2} & \frac{2x_1}{x_2^3} \end{pmatrix}$$
Le calcul des mineurs principaux donne $\Delta_1 = 0$ et $\Delta_2 = \frac{-1}{x_2^4}$. Par conséquent, elle n'est ni semi-définie positive, ni semi-définie négative.

   Ainsi, f **n'est ni convexe ni concave**.

4. La Hessienne de f est :
$$\begin{aligned} \nabla^2 f(x) &= \begin{pmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^{\alpha}x_2^{\alpha-1} \end{pmatrix} \\ &= \alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{pmatrix} \frac{-1}{x_1^2} & \frac{1}{x_1 x_2} \\ \frac{1}{x_1 x_2} & \frac{-1}{x_2^2} \end{pmatrix} \\ &= -\alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{pmatrix} \frac{1}{x_1} \\ \frac{-1}{x_2} \end{pmatrix} \begin{pmatrix} \frac{1}{x_1} \\ \frac{-1}{x_2} \end{pmatrix}^T \end{aligned}$$
Or, pour tout $x \in \mathbb{R}^2_{++}$, une matrice de la forme $A = xx^T$ est toujours semi-définie positive. Donc :
$$\begin{pmatrix} \frac{1}{x_1} \\ \frac{-1}{x_2} \end{pmatrix} \begin{pmatrix} \frac{1}{x_1} \\ \frac{-1}{x_2} \end{pmatrix}^T \succeq 0$$

Et donc, comme $\left[\alpha(1-\alpha)x_1^\alpha x_2^{1-\alpha} \geq 0\right]$ (puisque $\alpha \in [0,1]$), on a :

$$\nabla^2 f(x) \preceq 0$$

Ainsi, f **est concave**.

$\square$

Exercice 0.3 :

Montrer que les fonctions suivantes sont convexes :

1. $f(X) = \mathrm{Tr}(X^{-1})$ sur $\mathbf{dom}\, f = S^n_{++}$.
2. $f(X, y) = y^T X^{-1} y$ sur $\mathbf{dom}\, f = S^n_{++} \times \mathbb{R}^n$.
   Indice : exprimez-la comme un supremum.
3. $f(X) = \sum_{i=1}^n \sigma_i(X)$ sur $\mathbf{dom}\, f = S^n$, où $\sigma_1(X), \dots, \sigma_n(X)$ sont les valeurs singulières d'une matrice $X \in \mathbb{R}^{n \times n}$.
   Indice : exprimez-la comme un supremum.

Démonstration.

1. Pour prouver que la fonction f est convexe, j'ai utilisé la propriété de la page 7/67 du cours 2 ("*Restriction of a convex function to a line*").

   On définit donc $g : \mathbb{R} \to \mathbb{R}$ telle que : $\forall t \in \mathbb{R}$, $g(t) = f(X + tV)$ où $X \in S^n_{++}$ et $V \in S^n$. On sait que $X \in S^n_{++}$, donc elle admet une racine carré $Y \in S^n_{++}$ telle que $X = Y^2$.

$$\begin{aligned} g(t) &= \mathrm{Tr}[(X + tV)^{-1}] \\ &= \mathrm{Tr}[(Y^2 + tV)^{-1}] && \text{(car } X = Y^2\text{)} \\ &= \mathrm{Tr}\left[(Y(I_n + tY^{-1}VY^{-1})Y)^{-1}\right] \\ &= \mathrm{Tr}\left[Y^{-2}(I_n + tY^{-1}VY^{-1})^{-1}\right] && \text{(par propriété de la trace)} \\ &= \mathrm{Tr}\left[X^{-1}(I_n + tY^{-1}VY^{-1})^{-1}\right] \end{aligned}$$

   Si A et B sont symétriques, alors ABA est symétrique. Ici, Y^{-1} et V sont symétriques, donc le produit $Y^{-1}VY^{-1}$ est symétrique. On peut alors l'orthodiagonaliser :

$$Y^{-1}VY^{-1} = QDQ^T \qquad \text{avec } D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

   On a donc :

$$\begin{aligned} g(t) &= \mathrm{Tr}\left[X^{-1}(I_n + tY^{-1}VY^{-1})^{-1}\right] \\ &= \mathrm{Tr}\left[X^{-1}(I_n + tQDQ^T)^{-1}\right] \\ &= \mathrm{Tr}\left[X^{-1}Q(I_n + tD)^{-1}Q^T\right] \\ &= \mathrm{Tr}\left[Q^T X^{-1} Q(I_n + tD)^{-1}\right] && \text{(par propriété de la trace)} \\ &= \sum_{i=1}^n (Q^T X^{-1} Q)_{ii}(1 + t\lambda_i)^{-1} \end{aligned}$$

   Les fonctions $t \longmapsto (1 + t\lambda_i)^{-1}$ sont convexes. Comme la fonction g s'écrit comme une somme pondérée de ces fonctions là, on en déduit que **la fonction g est convexe en t.**

   Ainsi, par propriété, f **est convexe**.

□