Devoir maison n^{o} 1 convex optimization

Maria Cherifa

18 octobre 2020

Exercice 2.12 :

Dans cet exercice on veut voir si les ensembles donnés sont convexes.

1- L'ensemble de la forme $A = \{x \in \mathbf{R}^n \mid \alpha \leq a^T x \leq \beta\}$: la première idée est de prendre deux points $x_1 \in A$ et $x_2 \in A$ et de voir si pour $\theta \in [0,1]$ on a bien $\theta x_1 + (1-\theta) x_2 \in A$, cela revient à voir si $\alpha \leq a^T (\theta x_1 + (1-\theta) x_2) \leq \beta$. En effet comme x_1 et x_2 sont dans A on a :

$$\alpha \leq a^T x_1 \leq \beta$$

$$\alpha \leq a^T x_2 \leq \beta$$

En multipliant la première inégalité par θ et la deuxièmes par $1-\theta$ et en faisant la somme des deux inégalités on trouve bien :

$$\alpha \leq a^T (\theta x_1 + (1-\theta) x_2) \leq \beta$$

Donc A est un ensemble convexe.

Méthode 2 : une méthode plus simple consiste à voir que A est l'intersection de deux demi-espaces, donc comme vu dans le cours que les demi-espaces sont convexes et que l'intersection d'ensembles convexes est un ensemble convexe, ceci implique que A est convexe.

2- L'ensemble de la forme $A = \{x \in \mathbf{R}^n \mid \alpha_i \leq x_i \leq \beta_i \, i = 1, ..., n\}$: on peut voir facilement que A est une intersection finie de demi-espaces, comme vu dans le cours A est donc convexe.

3- L'ensemble de la forme $A = \{x \in \mathbf{R}^n | a_1^T x \leq \beta_1 \, a_2^T x \leq \beta_2\}$: de même A est l'intersection de deux demi-espaces, donc c'est un ensemble convexe.

4- L'ensemble de la forme $A = \{x \mid ||x - x_0||_2 \leq ||x-y||_2 \, , y \in S\}$ avec $S \subset \mathbf{R}^n$: on peut voir A comme :

$$A = \bigcap_{y \in S} \{x \mid ||x - x_0||_2 \leq ||x-y||_2 \, y \in S\}$$

une intersection finie de demi-espaces, car nous savons que si y est fixé alors $\{x \mid ||x - x_0||_2 \leq ||x-y||_2 \, y \in S\}$ cet ensemble est un demi-espace, donc A est convexe.

5- L'ensemble de la forme $A = \{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\}$ avec $S, T \subset \mathbf{R}^n$, et $\mathbf{dist}(x, S) = inf\{||x - z||_2 | \, z \in S\}$: si on prend par exemple $S = \{a\}$ et $T = \{0\}$ alors $A = \{x | \, |x-a| \leq |x|\} = \{x | \, x \leq \frac{-a}{2} \, x \geq \frac{a}{2}\}$, qui n'est pas convexe.

6- L'ensemble de la forme $A = \{x \mid \; x + S_2 \subset S_1\}$ avec $S_1, S_2 \subset \mathbf{R}^n$ et S_1 est convexe : tout d'abord nous pouvons réecrire A de cette manière $A = \{x \mid \; x + y \subset S_1, \text{pour tout } y \in S_2\}$ et donc :

$$A = \bigcap_{y \in S_2} \{x | \, x \in S_1 - y\}$$

Donc c'est une intersection finie d'ensembles convexes de la formule $S_1 - y$, ce qui implique que A est convexe.

7- L'ensemble de la forme $A = \{x| \, ||x-a||_2 \leq \theta \, ||x-b||_2\}$ avec $a \neq b$ et $\theta \in [0,1]$: cet ensemble est un demi-espace car si on passe au carré, A est équivalent à :

$$\{x| \, ||x-a||_2^2 \leq \theta^2 \, ||x-b||_2^2\}$$

$$\{x| \, x^T x - 2 < x, a > + a^T a \leq \theta^2 \, (x^T x - 2 < x, b > + b^T b)\}$$
$$\{x| \, (1-\theta^2) x^T x - 2(a - \theta^2 \, b)^T \, x + (a^T a - \theta^2 b^T b) \leq 0\}$$

et donc pour le cas particulier $\theta = 1$, l'ensemble précédent est un demi-espace, donc A est convexe.

Exercice 3.21 :

Dans cet exercice nous allons prouver que les fonction $f : \mathbf{R}^n \rightarrow \mathbf{R}$ sont convexes.

1- $f(x) = \mathbf{max}_{i=1,\ldots k} ||A^{(i)} \, x - b^{(i)}||$ avec $A^{(i)} \in \mathbf{R}^{m \times n}$ et $b^{(i)} \in R^m$ et $||.||$ est une norme de $\mathbf{R}^n$: cette fonction est convexe car c'est le maximum de fonction convexe, car comme vu dans le cours les fonction composées d'une fonction affine et une norme sont convexes donc $||A^{(i)} \, x - b^{(i)}||$ est convexe et donc le maximum de k fonctions convexes est convexe aussi, donc f est convexe.

2- $f(x) = \sum_{i=1}^{r} |x|_{[i]}$ sur $\mathbf{R}^n$, $|x|_{[i]}$ est le plus grand élément de $|x|$: cette fonction est convexe car cette fonction peut s'écrire sous la forme $f(x) = \mathbf{max}_{1 \leq i_1 \leq \ldots \leq i_r \leq n} |x_{i_1}| + \ldots + |x_{i_r}|$ et donc f est le maximum de nombre finie de fonctions convexes donc par le cours f est convexe.

Exercice 3.32 :

Dans cet exercice nous allons essayer de faire des preuve concernant le produit et les fractions de fonctions convexes.

1- Prouvons que si f et g sont convexes et sont croissantes (ou decroissantes) et positives sur un intervalle alors $f \, g$ est convexe : utilisons la définitions de la convexité, soit $\theta \in [0,1]$:

$$f(\theta \, x + (1-\theta) y) \leq \theta \, f(x) + (1-\theta) f(y)$$

$$g(\theta \, x + (1-\theta) y) \leq \theta \, g(x) + (1-\theta) g(y)$$

Le produit des deux inégalités donne :

$$\begin{aligned} f(\theta \, x + (1-\theta) y) g(\theta \, x + (1-\theta) y) &\leq (\theta \, f(x) + (1-\theta) f(y)) \, (\theta \, g(x) + (1-\theta) g(y)) \\ &\leq \theta f(x) g(x) + (1-\theta) \, f(y) g(y) + \theta(1-\theta)(f(y) - f(x))(g(x) - g(y)) \end{aligned}$$

Comme f et g sont soit croissantes ou décroissantes donc dans le deux cas on $(f(y) - f(x)) \leq 0$ et $(g(x) - g(y)) \geq 0$ ce qui implique que $\theta(1-\theta)(f(y) - f(x))(g(x) - g(y)) \leq 0$ et donc on a bien

$$f(\theta \, x + (1-\theta) y) f(\theta \, x + (1-\theta) y) \leq \theta f(x) g(x) + (1-\theta) \, f(y) g(y)$$

Nous avons donc prouver que $f \, g$ est convexe.

2- Prouvons que si f et g sont concaves et sont croissantes (ou decroissantes) et positives sur un intervalle alors $f\,g$ est concave : utilisons la définitions de la concavité, soit $\theta \in [0,1]$

$$f(\theta\, x + (1-\theta)y) \geq \theta\, f(x) + (1-\theta) f(y)$$

$$f(\theta\, x + (1-\theta)y) \geq \theta\, g(x) + (1-\theta) g(y)$$

Le produit des deux inégalités donne :

$$\begin{aligned} f(\theta\, x + (1-\theta)y) f(\theta\, x + (1-\theta)y) &\geq (\theta\, f(x) + (1-\theta) f(y))\, (\theta\, g(x) + (1-\theta) g(y)) \\ &\geq \theta f(x) g(x) + (1-\theta)\, f(y) g(y) + \theta(1-\theta)(f(y)-f(x))(g(x)-g(y)) \end{aligned}$$

Comme f et g sont soit croissantes ou décroissantes donc dans le deux cas on $(f(y) - f(x)) \leq 0$ et $(g(x) - g(y)) \geq 0$ ce qui implique que $\theta(1-\theta)(f(y)-f(x))(g(x)-g(y)) \leq 0$ et donc on a bien

$$f(\theta\, x + (1-\theta)y) f(\theta\, x + (1-\theta)y) \geq \theta f(x) g(x) + (1-\theta)\, f(y) g(y)$$

Nous avons donc prouver que $f\,g$ est concave .

3- Prouvons que f est convexe, croissante et positive et g concave, décroissante et positive alors f/g est convexe : comme g est concave alors $1/g$ est convexe et croissante, positive alors on applique le résultat trouvé dans (1) à f et $1/g$.

Exercice 3.36 :

Dans cet exercice nous allons calculer les fonctions conjuguées des fonctions données :

1- $f(x) = \mathbf{max}_{i=1,..n} x_i$ sur $\mathbf{R}^n$, nous savons que $f^*(y) = \mathbf{sup}_{x\in \mathbf{dom} f}(y^T\, x - f(x))$, alors la fonction conjuguée de f est donnée par :

$$f^*(y) = \begin{cases} 0 \quad \text{si } y \geq 0 \ \text{ et } \mathbf{1}^T\, y = 0 \\ \qquad\qquad\qquad \infty \quad \text{sinon} \end{cases}$$

Nous savons que le domaine de la fonction conjuguée est constitué des $y \in \mathbf{R}^n$ pour lequel le supremum est fini. Prouvons que ce domaine est le bon, supossons que y a une composante negative $(y_j < 0)$, et que par exemple on prend un x tel que $x_j = \alpha$ avec $\alpha < 0$ et $x_i = 0$ avec $i \neq j$, nous aurons alors :

$$y^T\, x - \mathbf{max}_i x_i = y_j\, \alpha \to \infty$$

Nous déduisons que y n'appartient pas au domaine de f^* puisque f^* tend vers l'infini. Prenons maintenant $y \geq 0$ et $\mathbf{1}^T\, y > 1$ et on choisit $x = \alpha\, \mathbf{1}$, nous allons avoir alors :

$$y^T\, x - \mathbf{max}_i x_i = y^T\, \alpha\, \mathbf{1} - \alpha$$

Quand t tend vers l'infini l'expression précédente tend vers l'infini donc y n'appartient pas au domaine de f^*. Et si on prend $y \geq 0$ et $\mathbf{1}^T\, y < 1$ et on choisit $x = -\alpha\, \mathbf{1}$, nous allons avoir alors :

$$y^T\, x - \mathbf{max}_i x_i = -y^T\, \alpha\, \mathbf{1} + \alpha$$

Quand t tend vers l'infini l'expression précédente tend vers l'infini, de même y n'appartient pas au domaine de f^*.

Donc pour le cas $y \geq 0$ et $\mathbf{1}^T\, y = 1$, nous avons alors :

$$y^T x \leq \mathbf{max}_i x_i$$

Ceci implique que $y^T\, x - \mathbf{max}_i x_i \leq 0$ et si $x = 0$ on a l'égalité et donc on retrouve bien $f^*(y) = 0$, nous avons prouvé alors que f^* que nous avons trouvé est la bonne fonction.

2- $f(x) = \sum_{i=1}^{r} x_{[i]}$ sur $\mathbf{R}^n$, en utilisant la définition précédente de la fonction conjuguée, on trouve alors :

$$f^*(y) = \begin{cases} 0 \quad \text{si } 0 \leq y \leq 1 \;\text{ et } \mathbf{1}^T y = r \\ \qquad\qquad\qquad\qquad\quad \infty \quad \text{sinon} \end{cases}$$

Nous savons que le domaine de la fonction conjuguée est constitué des $y \in \mathbf{R}^n$ pour lequel le supremum est fini. Supossons que que y a une composante negatives ($y_k < 0$), et que par exemple on prend un x tel que $x_j = \alpha$ avec $\alpha < 0$ et $x_i = 0$ avec $i \neq j$, nous aurons alors :

$$y^T x - f(x) = y_k \alpha \rightarrow \infty$$

Nous déduisons que y n'appartient pas au domaine de f^* puisque f^* tend vers l'infini. Supposons maintenant que l'un des éléments de y est supérieur à 1 par exemple $y_j \geq 1$, et par exemple x est défini par $x_i = \beta$ avec $\beta > 0$ et $x_j = 0$ avec $i \neq j$, nous aurons alors pour β tendant à l'infini :

$$y^T x - f(x) = y_j \beta - \beta \rightarrow \infty$$

Donc y n'appartient pas au domaine de f^* pour ce cas. Supposons à présent que $\mathbf{1}^T y \neq r$ et que $x = \mathbf{1}\,\alpha$, nous avons alors :

$$y^T x - f(x) = y^T \mathbf{1}\,\alpha - \alpha\, r$$

la quantité précédente tend vers l'infini pour α qui tend vers l'infini donc y n'appartient pas au domaine de f^*.

Enfin si y satisfait toutes les conditions précédentes, et si on prend un x tel que $x_1 \geq ... \geq x_n$, alors :

$$\begin{aligned} y^T x - f(x) &= \sum_{i=1}^{n} x_i\, y_i - \sum_{i=1}^{r} x_{[i]} \\ &= \sum_{i=r+1}^{n} x_i\, y_i + \sum_{i=1}^{r} y_i - \sum_{i=1}^{r} x_{[i]} \\ &\leq \sum_{i=r+1}^{n} x_i\, y_i + \sum_{i=1}^{r} (y_i - 1) x_i \\ &\leq x_r \left(\sum_{i=r+1}^{n} y_i + \sum_{i=1}^{r} (y_i - 1) \right) \\ &\leq 0 \end{aligned}$$

et c'est égale à 0 pour $x = 0$. Donc on a bien $f^*(y) = 0$, et donc f^* est la bonne fonction.

3- $f(x) = \mathbf{max}_{i=1,...m} a_i\, x + b_i$, en utilisant la définition précédente de la fonction conjuguée, on peut calculer la fonction conjuguée avec cette expression :

$$f^*(y) = \mathbf{sup}_x (x\, y - \mathbf{max}_{i=1,...m} a_i\, x + b_i)$$

Comme on est sur $\mathbf{R}$ le produit scalaire est juste un produit. On peut voit que le sup de la fonction précédente est fini pour y dans l'intervalle $[a_1, a_m]$ et donc $\mathbf{dom} f = [a_1, a_m]$. Pour $a_i \leq y \leq a_{i+1}$, le sup défini dans la fonction f^* est attient en $x = \frac{b_i - b_{i+1}}{a_{i+1} - a_i}$, et donc on aura :

$$f^*(y) = y(\frac{b_i - b_{i+1}}{a_{i+1} - a_i}) - a_i\, (\frac{b_i - b_{i+1}}{a_{i+1} - a_i}) - b_i$$

avec $a_i \leq y \leq a_{i+1}$.

4- $f(x) = x^p$ sur $\mathbf{R}_{++}$ avec $p > 1$. Pour $p > 1$, prenons $y \leq 0$ alors $f^*(y) = xy - f(x) = xy - x^p$ c'est une fonction positive et décroissante son max est atteint pour $x \rightarrow 0$ alors pour ce cas $f^*(y) = 0$. Si on prend

$y > 0$ la fonction $f^*(y) = xy - f(x) = xy - x^p$ atteint son max en $x = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$, et la valeur de la fonction en ce point est égale à (avec $q = \frac{p}{p-1}$) :

$$f^*(y) = y\left(\frac{y}{p}\right)^{\frac{1}{p-1}} - \left(\frac{y}{p}\right)^{\frac{p}{p-1}} = (p-1)\left(\frac{y}{p}\right)^q$$

On déduit alors que :

$$f^*(y) = \begin{cases} 0 & \text{si } y \leq 0 \\ (p-1)\left(\frac{y}{p}\right)^q & \text{si } y > 0 \end{cases}$$

Pour $p < 0$, on reprend le même raisonnement en réecrivant $x^p = \frac{1}{x^{-p}}$ avec $-p > 0$ et on retrouve alors que $f^*(y) = (p-1)\left(\frac{y}{p}\right)^q$ et le domaine de f^* est pour ce cas $-\mathbf{R}_+$.

5- $f(x) = -(\prod x_i)^{\frac{1}{n}}$ sur $\mathbf{R}^n_{++}$, en utilisant la définition précédente de la fonction conjuguée, on trouve alors :

$$f^*(y) = \begin{cases} 0 \quad \text{si } 0 \leq y \leq 0 \text{ et } (\prod_i(-y_i))^{\frac{1}{n}} \geq \frac{1}{n} \\ \infty \quad \text{sinon} \end{cases}$$

Nous savons que le domaine de la fonction conjuguée est constitué des $y \in \mathbf{R}^n$ pour lequel le supremum est fini. Supossons que y à un élement positif ($_j > 0$), et que par exemple on prend un x tel que $x_j = \alpha$ avec $\alpha > 0$ et $x_i = 1$ avec $i \neq j$, nous aurons alors :

$$y^T x - f(x) = y_j\,\alpha + \alpha^{\frac{1}{n}} + \sum_{i \neq j} y_i \to \infty$$

pour α tendant vers l'infini, donc ici y n'appartient pas au domaine de f^*. Supposons à présent que $y < 0$ et que $(\prod_i(-y_i))^{\frac{1}{n}} \leq \frac{1}{n}$, prenons par exemple $x_i = \frac{-\alpha}{y_i}$, alors :

$$y^T x - f(x) = -\alpha\, n + \alpha\left((\prod_i(-y_i^{-1}))\right)^{\frac{1}{n}}$$

la quantité précédente tend vers l'infini pour α tendant vers l'infini, donc aussi pour ce cas y n'appartient pas au domaine de f^*.

Enfin supposons que $y < 0$ et que $(\prod_i(-y_i))^{\frac{1}{n}} \geq \frac{1}{n}$, prenons $x > 0$, nous savons alors que l'inégalité arithmético-géométrique donne :

$$\frac{y^T x}{n} \geq \left(\prod_i(-x_i\, y_i)\right)^{\frac{1}{n}} \geq \frac{1}{n}\left(\prod_i(x_i)\right)^{\frac{1}{n}}$$

Donc $y^T x \leq f(x)$ et si $x \to 0$ on a $y^T x = f(x)$ ce qui implique que $f^*(y) = 0$ et donc notre fonction est la bonne.

6- $f(x,t) = \log(t^2 - x^T x)$ définie sur $\{(x,t) \in \mathbf{R}^n \times \mathbf{R} \mid ||x||_2 < t\}$, en utilisant la définition de la fonction conjuguée, on trouve :

$$f^*(y,v) = 2 + \log(4) - log(v^2 - y^T y)$$

et le domaine de définition de f^* est très similaire au domaine de f et il est donné par $\mathbf{dom}(f^*) = \{(y,v) \in \mathbf{R}^n \times \mathbf{R} \mid ||y||_2 < -v\}$, en effet pour $||y||_2 > -v$ avec $v < 0$ et en prennant $x = \alpha\, y$, $t = ||x||_2 + 1 \geq \alpha\, ||y||_2 \geq -\alpha\, v$, avec $\alpha \geq 0$. alors :

$$y^T x + t\, v \geq \alpha\, y^T y - \alpha\, v^2 = \alpha(y^T y - v^2) \geq 0$$

$$\log(t^2 - x^T x) = \log(2\,\alpha\, ||y||_2 + 1)$$

Alors :

$$y^T\,x + t\,v + \log(t^2 - x^T\,x)$$

tend vers l'infini et donc est non bornée, donc y n'appartient pas au domaine de f^* dans ce cas.

Supposons maintenant $||y||_2 \leq -v$, en dérivant

$$y^T\,x + t\,v + \log(t^2 - x^T\,x)$$

et en respectant le fait que x et t sont égaux à 0, on trouve que les maximiseurs qui sont définis par :

$$x = \frac{2\,y}{v^2 - y^T\,y}$$

et

$$t = \frac{-2\,v}{v^2 - y^T\,y}$$

alors en remplaçant on trouve :

$$f^*(y,v) = y^T\,x + t\,v + \log(t^2 - x^T\,x)$$

$$f^*(y,v) = 2 + \log(4) - log(v^2 - y^T\,y)$$

donc on a trouvé la bonne fonction.