CALLARD Baptiste

Homework 1 : convex optimization

MVA.

Homework 1

exercice 1

1). Let's show that $C := \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, \forall i = 1, \dots, m\}$

with $\alpha_i, \beta_i \in \mathbb{R}$ $\forall i = 1, \dots, n$ is convex

First $C = \bigcap_{i=1}^{m} \underbrace{\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i\}}_{:= C_i}$,

let $i \in \{1, \dots, m\}$

$C_i = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i\} \cap \{x \in \mathbb{R}^n \mid x_i \leq \beta_i\}$

$= \{x \in \mathbb{R}^n \mid \alpha_i \leq e_i^T x\} \cap \{x \in \mathbb{R}^n \mid e_i^T x \leq \beta_i\}$ with $e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \swarrow i^{th}$

So $\forall i = 1, \dots, n$, C_i is an intersection of 2 half-spaces, which are convex sets so C_i is convex

Hence $\underline{C \text{ is convex}}$ as it is an intersection of m convex sets.

2) Let's show that $C := \{x \in \mathbb{R}_+^2 \mid x_1 x_2 \geq 1\}$ is convex,

We will use the definition, let $\theta \in [0, 1]$ and $x, y \in C$

We want to show that $z := \theta x + (1-\theta) y \in C$

$z_1 z_2 = (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2)$

$= \theta^2 \underbrace{x_1 x_2}_{\geq 1} + (1-\theta)^2 \underbrace{y_1 y_2}_{\geq 1} + \underbrace{\theta(1-\theta)[x_1 y_2 + x_2 y_1]}_{\geq 0}$

$\geq \theta^2 + (1-\theta)^2 = 2\theta^2 + 1 - 2\theta = \underbrace{2\theta(\theta - 1)}_{\geq 0} + 1 \geq 1$

As $z_1 z_2 \geq 1$, $z = \theta x + (1-\theta) y \in C$, therefore $\underline{C \text{ is convex}}$

3) Let's show that $C := \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$ is convex.

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

we note $C_y := \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$

$x \in C_y \iff \|x - x_0\|_2^2 \leq \|x - y\|_2^2$ by growth of square function

$$\iff \|x\|_2^2 - 2\langle x, x_0\rangle + \|x_0\|_2^2 \leq \|x\|_2^2 - 2\langle x, y\rangle + \|y\|_2^2$$

$$\iff \langle x, y - x_0\rangle \leq \frac{\|y\|_2^2 - \|x_0\|_2^2}{2}$$

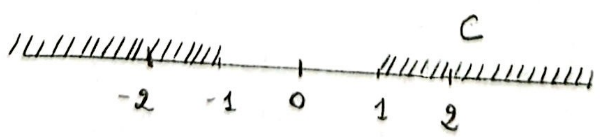
$$\iff (y - x_0)^T x \leq \frac{\|y\|_2^2 - \|x_0\|_2^2}{2}$$

$C_y = \left\{x \mid (y - x_0)^T x \leq \dfrac{\|y\|_2^2 - \|x_0\|_2^2}{2}\right\}$ is an half space hence convex.

Therefore $\underline{C \text{ is convex}}$ because it is an intersection of convex sets.

4) Let's show that $C := \{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\}$ where $S, T \subseteq \mathbb{R}^n$ and $\operatorname{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$ is not convex. For instance, this is a counter example for $n = 1$:

We define $S =]-\infty, 2] \cup [2, +\infty[$ and $T = \{0\}$



$C =]-\infty, -1] \cup [1, +\infty[$ which is clearly non convex.

5) Let's show that $C := \{x \mid x + S_2 \subseteq S_1\}$ where $S_1, S_2 \subseteq \mathbb{R}^2$ with S_1 convex is convex. We use the definition.

Let $\theta \in [0,1]$ and $x, y \in C$,

Let $z \in S_2$

$$\theta x + (1-\theta)y + z = \theta(x+z) + (1-\theta)(y+z)$$

$x \in C$ so $x + z \in S_1$

$y \in C$ so $y + z \in S_1$

so, $\theta x + (1-\theta)y + z = \underbrace{\theta(x+z)}_{\in S_1} + (1-\theta)\underbrace{(y+z)}_{\in S_1} \in S_1$ because S_1 is convex.

It means that $\theta x + (1-\theta)y \in C$.

Hence we prove that <u>C is convex</u>

exercice 2 :

Before starting the question 1 let's derive two usefull results

i) Let $A \in M_2(\mathbb{R})$ symmetric, there exists a matrix P orthogonal such that $A = PDP^{-1}$ with D a diagonal matrix with all the egein values of A on the diagonal. We note λ_1, λ_2 the egein value

Then $\det(A) = \det(PDP^{-1}) = \det(P)\det(D)\dfrac{1}{\det(P)} = \det(D) = \lambda_1\lambda_2$

$\mathrm{Tr}(A) = \mathrm{Tr}(PDP^{-1}) = \mathrm{Tr}(PP^{-1}D) = \mathrm{Tr}(D) = \lambda_1 + \lambda_2$

Hence :

$* \; A \in S_2^+(\mathbb{R}) \iff \begin{cases} \lambda_1 \geq 0 \\ \lambda_2 \geq 0 \end{cases} \iff \begin{cases} \det(A) = \lambda_1\lambda_2 \geq 0 \\ \mathrm{Tr}(A) = \lambda_1 + \lambda_2 \geq 0 \end{cases}$

$* \; A \in S_2^-(\mathbb{R}) \iff \begin{cases} \lambda_1 \leq 0 \\ \lambda_2 \leq 0 \end{cases} \iff \begin{cases} \det(A) = \lambda_1\lambda_2 \geq 0 \\ \mathrm{Tr}(A) = \lambda_1 + \lambda_2 \leq 0 \end{cases}$

ii) $f : A \to B$ convex $\Rightarrow f : A \to B$ quasi-convex

Let $\theta \in [0,1]$, $\alpha \in B$, $x, y \in \{x \in A \mid f(x) \leq \alpha\} = C$

$$f(\theta x + (1-\theta)y) \underset{\text{convex}}{\leq} \theta f(x) + (1-\theta) f(y) \underset{x,y \in C}{\leq} \theta\alpha + (1-\theta)\alpha = \alpha$$

1) Let $x \in \mathbb{R}^2_{++}$ and $f(x_1, x_2) = x_1 x_2$, f is twice differentiable on $\mathbb{R}^2_{++}$.

$$\nabla^2 f(x_1,x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \det(\nabla^2 f(x_1,x_2) - tI_2) = t^2 - 1 = (t+1)(t-1),\ t \in \mathbb{R}$$

But the eigenvalues are the roots of the characteristic polynoma

let $\lambda_1 = -1$, $\lambda_2 = 1$ the eigenvalues. $\nabla^2 f(x_1, x_2) \notin S_+^2$ so <u>f is not convex</u> and $\nabla^2(-f)(x_1, x_2) \notin S_+^2$ so <u>f is not concave</u>.

Let us show that $-f$ is quasi-convex ($\iff f$ is quasi-concave)

$C := \{x \in \mathbb{R}^2_{++} \mid f(x_1,x_2) \geq \alpha\}$, $\alpha \in \mathbb{R}$.

. if $\alpha < 0$ then $C = \mathbb{R}^2_{++}$ hence convex

. if $\alpha \geq 0$, let $\theta \in [0,1]$ and $x, y \in C$ (same as 2.1)

$$(\theta x_1 + (1-\theta)y_1)(\theta x_2 + (1-\theta)y_2) = \theta^2 \underbrace{x_1x_2}_{\geq \alpha} + (1-\theta)^2 \underbrace{y_1y_2}_{\geq \alpha} + \underbrace{\theta(1-\theta)[x_1y_2 + x_2y_1]}_{\geq 0}$$

$$\geq \alpha[\theta^2 + (1-\theta)^2] \geq \underbrace{2\alpha\theta(\theta-1)}_{\geq 0} + \alpha \geq \alpha$$

So $\theta x + (1-\theta)y \in C$. Therefore C is convex.

Hence, <u>f is quasi-concave</u> (but not quasi-convex).

2) Let $x \in \mathbb{R}_{++}^2$, $f(x_1, x_2) = \frac{1}{x_1 x_2}$, f is twice differentiable on $\mathbb{R}_{++}^2$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix}$$

$$\det(\nabla^2 f(x_1, x_2)) = \frac{4}{x_1^4 x_2^4} - \frac{1}{x_1^4 x_2^4} = \frac{3}{x_1^4 x_2^4} > 0.$$

$$\mathrm{Tr}(\nabla^2 f(x_1, x_2)) = \frac{2}{x_1^3 x_2} + \frac{2}{x_1 x_2^3} > 0$$

Hence, with the result derives at the begging we can conclude that <u>f is convex</u>. In particular, f is <u>quasi-convex</u>. (f is not concave and quasi-concave)

3) Let $x \in \mathbb{R}_{++}^2$, $f(x_1, x_2) = \frac{x_1}{x_2}$, f is twice differentiable on $\mathbb{R}_{++}^2$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$$

$\det(\nabla^2 f(x_1, x_2)) = -\frac{1}{x_2^4}$ hence if we note λ_1, λ_2 the eigenvalues of $\nabla^2 f(x_1, x_2)$ we get that $\lambda_1 \lambda_2 < 0$.

It means $[\lambda_1 > 0$ and $\lambda_2 < 0]$ or $[\lambda_1 < 0$ and $\lambda_2 > 0]$.

So $\nabla^2 f(x_1, x_2) \notin S_+^2$ and $\nabla^2 f(x_1, x_2) \notin S_+^2$.

<u>f is not convex</u> and <u>f not concave</u>

Let's prove that f is quasi-linear.

Let $C_1 := \{x \in \mathbb{R}^2_{++} \mid f(x_1, x_2) \leq \alpha\}$

$= \{x \in \mathbb{R}^2_{++} \mid x_1 - \alpha x_2 \leq 0\}$

$= \{x \in \mathbb{R}^2_{++} \mid \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}^T x \leq 0\}$

C_1 is an half-space so convex. f is quasi-convex

$C_2 := \{x \in \mathbb{R}^2_{++} \mid -f(x_1, x_2) \leq \alpha\}$

$= \{x \in \mathbb{R}^2_{++} \mid -\alpha x_2 - x_1 \leq 0\}$

$= \{x \in \mathbb{R}^2_{++} \mid \begin{bmatrix} -1 \\ -\alpha \end{bmatrix}^T x \leq 0\}$

C_2 is an half-space so convex. $-f$ is quasi-convex

By definition, f quasi-concave.

f quasi concave + quasi convex $\Rightarrow$ <u>f is quasi linear</u>

4) Let $x \in \mathbb{R}^2_{++}$, $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$

$f(x_1, x_2) = x_2 \left[\frac{x_1}{x_2}\right]^{\alpha} = x_2\, g\left(\frac{x_1}{x_2}\right)$ with $g : \begin{array}{l} \mathbb{R}_{++} \to \mathbb{R} \\ t \to t^{\alpha} \end{array}$

g is twice differentiable

$(-g)''(t) = -\left[\underbrace{\alpha}_{\geq 0} \underbrace{(\alpha - 1)}_{\leq 0} \underbrace{t^{\alpha-2}}_{\geq 0}\right] \geq 0$, $-g$ is convex

$-f$ is the perspective function of the function $-g$ convex

From the course, $-f$ is convex. It means that <u>f is concave</u>. In particular <u>f is quasi-concave</u>

(f not convex or quasi convex)

exercice 3 :

1) Let's show that $f(X) = Tr(X^{-1})$ is convex.

We use the proposition of retriction of a convex function to a line

$\left[f : S_{++}^m \to \mathbb{R} \text{ convex } \text{(iif)} \quad g : \begin{array}{l} \mathbb{R} \to \mathbb{R} \\ t \to f(X+tV) \end{array} \text{ is convex } \forall X \in S_{++}^m, \forall V \in S^m, t \in \{t \in \mathbb{R} \mid X+tV \in S_{++}^m\} \right]$

Let $V \in S^m$, $X \in S_{++}^m$, $t \in \{t \in \mathbb{R} \mid X+tV \in S_{++}^m\}$

$g(t) := f(X+tV)$

$g(t) = Tr((X+tV)^{-1})$

$X \in S_{++}^m$, from the Cholesky decomposition, there exists $L \in \mathcal{M}_n$ invertible and lower triangular such that $X = LL^T$

$$
\begin{aligned}
g(t) &= Tr((X+tV)^{-1}) \\
&= Tr((LL^T+tV)^{-1}) \\
&= Tr((L^T)^{-1}[I + tL^{-1}V(L^T)^{-1}]^{-1}L^{-1}) \quad , Tr(AB) = Tr(BA) \\
&= Tr(X^{-1}[I+tL^{-1}V(L^T)^{-1}]^{-1})
\end{aligned}
$$

$L^{-1}V(L^T)^{-1}$ is symmetric, from the spectral theorem, there exists $D \in \mathcal{M}_n$ a diagonal matrix and O an orthogonal matrix such that: $L^{-1}V(L^T)^{-1} = ODO^T$, with D a diagonal matrix whose coefficients are precisely the egeinvalues of $L^{-1}V(L^T)^{-1}$

$$
\begin{aligned}
g(t) &= Tr(X^{-1}[I + tODO^T]^{-1}), \\
&= Tr(OX^{-1}O^T(I+tD)^{-1}) \quad , Tr(AB) = Tr(BA) \\
&= \sum_{i=1}^m (OX^{-1}O^T)_{ii}(1+tD_{ii})^{-1} \quad \text{with } D_{ii} = \lambda_i \text{ the } i^{th} \text{ egeinvalues.} \\
&= \sum_{i=1}^m (OX^{-1}O^T)_{ii} \frac{1}{1+t\lambda_i} \quad , h_i(t) := \frac{1}{1+\lambda_i t} \text{ and } h_i''(t) = \frac{2\lambda_i^2}{(1+\lambda_i t)^3}
\end{aligned}
$$

$h_i''(t) > 0$ because $t \in \{t \in \mathbb{R} \mid X+tV > 0, X \in S_{++}^n, V \in S^n\}$ in particular $I_n \in S_{++}^n$, $D \in S^n$, $I+tD > 0$

Let u the i^{th} raw of O and u the i^{th} column of O^T $\Rightarrow$ $(OX^{-1}O^T)_{ii} = u^T X^{-1} u > 0$ because $X^{-1} \in S_{++}^n$

Hence, $g(t) = \sum_{i=1}^{m} \underbrace{(O X^{-1} O^T)_{ii}}_{> 0} h_i(t)$

So $t \to (O^T X^{-1} O)_{ii} h_i(t)$ is convex because h_i convex and $(O X^{-1} O^T)_{ii} > 0$

g is a sum of convex functions therefore g is convex, for all $X \in S_{++}^m$, $V \in S^m$, $t \in \{t \in \mathbb{R} \mid X + tV \succ 0\}$.

So $\underline{f \text{ is convex}}$

2) Let $X \in S_{++}^m$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^m$

From the course we have that $1/2\, y^T X^{-1} y = \sup_{x \in \mathbb{R}^n} \left(y^T x - \frac{1}{2} x^T X x\right)$

Let's prove this result:

It is equivalent to show that $\sup_{x \in \mathbb{R}^n} \left(y^T x - \frac{1}{2} x^T X x - 1/2\, y^T X^{-1} y\right) = 0$

$-y^T x + \frac{1}{2} x^T X x + \frac{1}{2} y^T X^{-1} y = \frac{1}{2}(X^{-1}y - x)^T X (X^{-1} y - x) \geq 0$ as $X \in S_{++}^m$, $(S^{-1}y - x) \in \mathbb{R}^m$

So $y^T x - \frac{1}{2} x^T X x \leq \frac{1}{2} y^T X^{-1} y$, it is true for all $x \in \mathbb{R}^n$ so

$$\sup_{x \in \mathbb{R}^n} \left(y^T x - \frac{1}{2} x^T X x\right) \leq \frac{1}{2} y^T X^{-1} y$$

For $x = X^{-1} y$ we get $y^T x - \frac{1}{2} x^T X x = y^T X^{-1} y - \frac{1}{2}(X^{-1}y)^T X X^{-1} y$

$= \frac{1}{2} y^T X^{-1} y$, the bound is reached

Hence $2 \sup_{x \in \mathbb{R}^n} \left(y^T x - \frac{1}{2} x^T X x\right) = y^T X^{-1} y = f(X, y)$

Let define $g(x, y, X) = 2y^T x - x^T X x$. This function is convex in y and in X for all $x \in \mathbb{R}^n$. $f(X, y) = \sup_{x \in \mathbb{R}^n} g(x, y, X)$

From the course f is convex as a pointwise supremum

3) $f(X) = \sum_{i=1}^{m} \sigma_i(X)$ on dom $f = S^2$ and $\sigma_1(X), \dots, \sigma_n(X)$ are the singular values of $X \in \mathbb{R}^{n\times n}$.

We can sort $\sigma_1(X) \geq \sigma_2(X) \geq \dots \geq \sigma_n(X)$.

By definition, $\sigma_i(X) = \sqrt{\lambda_i(X^TX)}$ with $\lambda_1(X^TX) \geq \dots \geq \lambda_n(X^TX)$ the eigenvalues of X^TX.

Let define the optimization problem:

$$\sup_{\substack{v\in\mathbb{R}^n \\ \|v\|_2=1}} \|Xv\|_2 = \sup_{\substack{v\in\mathbb{R} \\ \|v\|_2=1}} \sqrt{v^TX^TXv}$$

Let's introduce the unit basis $u_1, \dots, u_n$ of egeinvetors associated with the eigenvalues $\lambda_1, \dots, \lambda_n$.

We can write v in this basis: $v = \sum_{i=1}^{m} \alpha_i u_i$; $\alpha_i \in \mathbb{R}$.

So $v^TX^TXv = v^T \sum_{i=1}^{m} \alpha_i \underbrace{X^TXu_i}_{\lambda_i u_i} = v^T \sum_{i=1}^{m} \alpha_i \lambda_i u_i = \sum_{i=1}^{m} \alpha_i^2 \lambda_i$

But we have $\|v\|_2^2 = \langle \sum_{i=1}^{m} \alpha_i u_i, \sum_{i=1}^{m} \alpha_i u_i \rangle = \sum_{i=1}^{m}\sum_{j=1}^{m} \langle \alpha_i u_i, \alpha_j u_j \rangle$.

$= \sum_{i=1}^{m} \alpha_i^2$ because $\langle u_i, u_j \rangle = \delta_{ij}$ $\forall i,j \in 1, \dots, m$

$= 1$ because $\|v\|_2^2 = 1$

So $v^TX^TXv = \sum_{i=1}^{m} \alpha_i^2 \lambda_i \leq \lambda_1 \sum_{i=1}^{m} \alpha_i^2 = \lambda_1$ $\forall v \in \mathbb{R}^n$

$$\Rightarrow \sup_{\substack{v\in\mathbb{R}^n \\ \|v\|_2=1}} v^TX^TXv \leq \lambda_1$$

But $u_1^TX^TXu_1 = u_1^T\lambda_1 u_1 = \lambda_1$, hence $\sup_{\substack{v\in\mathbb{R}^n \\ \|v\|_2=1}} v^TX^TXv = \lambda_1 \underset{X^TX \in S_+^m}{\geq} 0$

We deduce that $\sigma_1 = \sup_{\substack{v\in\mathbb{R}^n \\ \|v\|=1}} \sqrt{v^TX^TXv}$

let's consider the problem:

$$\sup_{\substack{v \in \mathbb{R}^n \\ \|v\|_2 = 1 \\ v \perp u_1}} \|Xv\|_2$$

we still have $\sum_{i=1}^{m} \alpha_i^2 = 1$, $v^T X^T X v = \sum_{i=1}^{m} \alpha_i^2 \lambda_i$

But we also have that $\langle v, u_1 \rangle = \langle \sum_{i=1}^{m} \alpha_i u_i, u_1 \rangle = 0$

This means that $\alpha_1 = 0$.

Hence $v^T X^T X v = \sum_{i=2}^{m} \alpha_i^2 \lambda_i \leq \lambda_2 \sum_{i=2}^{m} \alpha_i^2 = \lambda_2 \quad \forall v \in \mathbb{R}^n$

we also have $u_2^T X^T X u_2 = \lambda_2$. $\Rightarrow \sup_{\substack{v \in \mathbb{R}^2 \\ \|v\|_2 = 1 \\ v \perp u_1}} v^T X^T X v = \lambda_2$

Therefore $\sup_{\substack{v \in \mathbb{R}^n \\ \|v\|_2 = 1 \\ v \perp u_1}} \|Xv\|_2 = \sigma_2$

We can iterate and easily get that:

$$\sup_{\substack{v \in \mathbb{R}^n \\ \|v\|_2 = 1 \\ v \perp u_i \ \forall i \in \{1, \dots, k\}}} \|Xv\|_2 = \sigma_k$$

Hence $\sum_{k=1}^{m} \sigma_k = \sup_{\substack{v \in \mathbb{R}^n \\ \|v\|_2 = 1}} \|Xv\|_2 + \sum_{k=1}^{m-1} \sup_{\substack{v \in \mathbb{R}^n \\ \|v\|_2 = 1 \\ v \perp u_i \ \forall i \in \{1, \dots, k\}}} \|Xv\|_2$

f is define as a sum of supremums of convex functions in X for all v. So each supremum is convex as pointwise supremum.

Finally f is a sum of convex functions hence

f is convex.

exercice 4 :

1) let's show that K_{m+} is a proper cone

i) K_{m+} is close.

$$K_{m+} = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$$
$$= \left[\bigcap_{i=1}^{m-1} \{x \in \mathbb{R}^n \mid x_i - x_{i-1} \geq 0\}\right] \cap \{x \in \mathbb{R}^n \mid x_n \geq 0\}$$

- $\forall i \in \{1, \dots, m\}$

$\{x \in \mathbb{R}^n \mid x_i - x_{i-1} \geq 0\} = f_i^{-1}([0, +\infty[)$ with $f_i : \begin{array}{l} \mathbb{R}^n \longrightarrow \mathbb{R} \\ x \longrightarrow x_i - x_{i-1} \end{array}$

f_i is continuous and $[0, +\infty[$ close, so $\{x \in \mathbb{R}^n \mid x_i - x_{i-1} \geq 0\}$ is close

- $\{x \in \mathbb{R}^m \mid x_n \geq 0\} = f_m^{-1}([0, +\infty[)$ with $f_n : \begin{array}{l} \mathbb{R}^n \to \mathbb{R} \\ x \to x_n \end{array}$, f_n is continuous and $[0, +\infty[$ close, so $\{x \in \mathbb{R}^n \mid x_n \geq 0\}$ is close

K_{m+} is close by intersection of close sets.

ii) K_{m+} is solid because the interior of K_{m+} is not empty. Indeed, there are points that statisfy strictly the inequalities.
For instance, the point $x = [n, m-1, m-2, \dots, 1] \in \mathring{K}_{m+}$.
Hence K_{m+} is solid.

iii) K_{m+} is pointed because it contains no line.
Let $x \in K_{m+}$ such that $-x \in K_{m+}$.

$x \in K_{m+} \Rightarrow x_1 \geq x_2 \geq \dots \geq x_n \geq 0$
$-x \in K_{m+} \Rightarrow -x_1 \geq -x_2 \geq \dots \geq -x_n \geq 0$

therefore $x_m = x_{n-1} = \dots = x_1 = 0$
That means that K_{m+} contains no line.

We show that K_{m+} is close, solid, pointed so K_{m+} is a proper cone

2) We want to find the dual cone $K_{m+}^* = \{ y \mid y^T x \geq 0 \ \forall x \in K_{m+} \}$

Let $y \in \mathbb{R}^n$ and $x \in K_{m+}$

$$y^T x = \sum_{i=1}^{m} y_i x_i \ , \quad y_i = \sum_{k=1}^{i} y_k - y_{k-1} \text{ by setting as convention that } y_0 = 0$$

$$= \sum_{i=1}^{m} x_i \sum_{k=1}^{i} y_k - y_{k-1}$$

$$= \sum_{i=1}^{m} x_i \left(\sum_{k=1}^{i} y_k - \sum_{k=0}^{i-1} y_k \right)$$

$$= \sum_{i=1}^{m} x_i \sum_{k=1}^{i} y_k - \sum_{i=0}^{m-1} x_{i+1} \sum_{k=0}^{i} y_k ; \text{ as } \sum_{i=0}^{m-1} x_{i+1} \sum_{k=0}^{i} y_k = \sum_{i=1}^{m-1} x_{i+1} \sum_{k=1}^{i} y_k$$

$$= x_n \sum_{k=1}^{m} y_k + \sum_{i=1}^{m-1} (x_i - x_{i+1}) \sum_{k=1}^{i} y_k \quad (1)$$

Let us show that $\forall i \in \{1, \dots, m\} \ \sum_{k=1}^{i} y_k \geq 0 \iff \forall x \in K_{m+} \ y^T x \geq 0$

($\Rightarrow$) We suppose $\forall i \in \{1, \dots, m\} \ \sum_{k=1}^{i} y_k \geq 0$, let $x \in K_{m+}$ then

$x_n \geq 0$ and $\forall i \in \{1, \dots, n-1\} \ x_i - x_{i+1} \geq 0$ so

$$\underbrace{\underbrace{x_n}_{\geq 0} \sum_{k=1}^{m} y_k}_{\geq 0} + \underbrace{\sum_{i=1}^{m-1} (x_i - x_{i+1})}_{\geq 0} \underbrace{\sum_{k=1}^{i} y_k}_{\geq 0} \geq 0$$ hence $\forall x \in K_{m+} \ y^T x \geq 0$.

($\Leftarrow$) We suppose that $\forall x \in K_{m+} \ y^T x \geq 0$

. $\forall k \in \{1, \dots, m\}$, we define $z_k = \begin{cases} [1, \dots, \underset{k^{th}}{1}, \underset{k+1^{th}}{0}, \dots, 0]^T & \text{if } k < m \\ [1, \dots\dots, 1]^T & \text{if } k = m \end{cases}$

Let $k \in \{1, \dots, m\}$, $z_k \in K_{m+}$ hence $y^T z_k \geq 0$ by hyp.

$$y^T z_k = \sum_{i=1}^{k} y_k \geq 0$$

Therefore $\forall k \in \{1, \dots, m\} \ \sum_{i=1}^{k} y_k \geq 0$

$$K_{m+}^* = \left\{ y \in \mathbb{R}^m \mid \sum_{i=1}^{k} y_i \geq 0 \ , k \in \{1, \dots, m\} \right\}$$

exercice 5:

1) $f(x) = \max_{i=1,\dots,m} x_i$ on $\mathbb{R}^n$

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(y^T x - \max_{i=1,\dots,n} x_i\right) = \sup_{x \in \mathbb{R}^n} \left(\sum_{i=1}^{m} x_i y_i - \max_{i=1,\dots,m} x_i\right)$$

We proceed by case disjunction:

- Let $i \in \{1,\dots,m\}$ such that $y_i < 0$, then we define:

$$x = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{th}, \text{ then } \sum_{i=1}^{m} y_i x_i - \max_{i=1,\dots,m} x_i = \underbrace{-y_i}_{\geq 0} t - \max(0, -t) \xrightarrow[t \to +\infty]{} +\infty$$

- let $y \geq 0$

  - if $\sum_{i=1}^{m} y_i > 1$ then we define $x = \begin{bmatrix} t \\ \vdots \\ t \end{bmatrix}$ then

$$\sum_{i=1}^{m} x_i y_i - \max_{i=1,\dots,m} x_i = t\left[\underbrace{\sum_{i=1}^{m} y_i - 1}_{>0}\right] \xrightarrow[t \to +\infty]{} +\infty$$

  - if $\sum_{i=1}^{m} y_i = 1$, by noticing that

$$\sum_{i=1}^{m} y_i x_i - \max_{i=1,\dots,m} x_i \leq \max_{i=1,\dots,m} x_i \left[\sum_{i=1}^{m} y_i - 1\right] = 0, \text{ so } f^*(y) \leq 0$$

and for $x = [0,\dots,0]$ this bound is reached so $f^*(y) = 0$

So:

$$\sup_{x \in \mathbb{R}^n} \left(\sum_{i=1}^{m} y_i x_i - \max_{i=1,\dots,m} x_i\right) = \begin{cases} 0 & \text{if } y \geq 0 \text{ and } \sum_{i=1}^{n} y_i = 1 \\ +\infty & \text{otherwise} \end{cases}$$

2) $f(x) = \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]}$$

- let $i \in \{1, \dots, n\}$ such that $y_i < 0$, then we define

$$x = \begin{bmatrix} 0 \\ \vdots \\ -t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{th}, \quad \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]} = \underbrace{-y_i}_{>0} t + \max(0, -t) \xrightarrow[t \to +\infty]{} +\infty$$

- let $i \in \{1, \dots, n\}$ such that $y_i > 1$ then we define

$$x = \begin{bmatrix} 0 \\ \vdots \\ t \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{th}, \quad \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]} = y_i t - \max(0, t)$$

for $t > 0$ $$\sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]} = t \underbrace{(y_i - 1)}_{>0} \xrightarrow[t \to +\infty]{} +\infty$$

- let $0 \leq y \leq 1$, let $x = [t \cdots t]^T$

$$\sum_{i=1}^{m} y_i t - rt = t\left[\sum_{i=1}^{m} y_i - r\right] \Rightarrow \begin{cases} \lim_{t \to +\infty} t\left[\sum_{i=1}^{m} y_i - r\right] = +\infty \text{ if } \sum_{i=1}^{m} y_i < r \\ \lim_{t \to -\infty} t\left[\sum_{i=1}^{m} y_i - r\right] = +\infty \text{ if } \sum_{i=1}^{m} y_i > r \end{cases}$$

- let $0 \leq y \leq 1$ and $\sum_{i=1}^{m} y_i = r$

$$\sum_{i=1}^{m} y_i x_i - \sum_{i=1}^{r} x_{[i]} = \sum_{i=1}^{m} y_{[i]} x_{[i]} - \sum_{i=1}^{r} x_{[i]} \quad \text{as } i \to [i] \text{ is a permutation}$$

$$= \sum_{i=1}^{r} \underbrace{(y_{[i]} - 1)}_{<0} x_{[i]} + \sum_{i=r+1}^{m} y_{[i]} x_{[i]}$$

$$\leq x_{[r]} \left(\sum_{i=1}^{r} y_{[i]} - 1 + \sum_{i=r+1}^{m} y_{[i]} \right)$$

$$\leq x_{[r]} \left[\sum_{i=1}^{m} y_{[i]} - r \right] = 0 \Rightarrow f^*(y) \leq 0$$

and for $x_1 = \dots = x_n = 0$, this bound is reached so $f^*(y) = 0$

So $$f^*(y) = \begin{cases} 0 & \text{if } 0 \leq y \leq 1 \text{ and } \sum_{i=1}^{m} y_i = r \\ +\infty & \text{otherwise} \end{cases}$$

3) $f(x) = \max_{i=1,\dots,m} (a_i x + b)$ on $\mathbb{R}$ with $\begin{cases} a_1 \leq \dots \leq a_m \\ a_i x + b_i \text{ are not redundant} \end{cases}$

$f^*(y) = \sup_{x \in \mathbb{R}} \left(xy - \max_{i=1,\dots,m} a_i x + b_i \right)$

As $(a_i)_{i \in \{1,\dots,m\}}$ are sorted and $(a_i x + b_i)_{i \in \{1,\dots,m\}}$ are not redundant we can define break points as

$a_{i+1} x + b_{i+1} = a_i x + b_i$ ie $x_i = \dfrac{b_i - b_{i+1}}{a_{i+1} - a_i}$ $\forall i \in \{1,\dots,m-1\}$

we set $x_0 = -\infty$ and $x_{m+1} = +\infty$

$g(x) := xy - \max_{i=1,\dots,m} (a_i x + b) = \sum_{i=1}^{m} xy \, 1\!\!1_{[x_{i-1}, x_i]}(x) - \sum_{i=1}^{m} (a_i x + b_i) \, 1\!\!1_{[x_{i-1}, x_i]}(x)$

$= \sum_{i=1}^{m} [x(y - a_i) + b_i] \, 1\!\!1_{[x_{i-1}, x_i]}(x) \quad \forall y \in \mathbb{R}$

- if $y \in [a_1, a_m]$, their exists i such that $a_i \leq y \leq a_{i+1}$ it means that $y - a_i \geq 0, \dots, y - a_1 \geq 0$, $y - a_{i+1} \leq 0, \dots, y - a_m \leq 0$ so g increases until x_i and then decreases. So g reaches it maximum on x_i

$f^*(y) = y \dfrac{b_i - b_{i+1}}{a_{i+1} - a_i} - a_i \dfrac{b_i - b_{i+1}}{a_{i+1} - a_i} + b_i = (y - a_i) \left[\dfrac{b_i - b_{i+1}}{a_{i+1} - a_i} \right] + b_i$

- if $y < a_1$, then $y - a_1 < 0, \dots, y - a_m < 0$, g is strictly decreasing on $\mathbb{R}$. $\lim_{x \to -\infty} g(x) = +\infty$

$f^*(y) = +\infty$

- if $y > a_m$, then $y - a_1 > 0, \dots, y - a_m > 0$ g is strictly increasing on $\mathbb{R}$: $\lim_{x \to +\infty} g(x) +\infty$

$f^*(y) = +\infty$

Finally: $f^*(y) = \begin{cases} (y - a_i) \left[\dfrac{b_i - b_{i+1}}{a_{i+1} - a_i} \right] + b_i & \text{if there exists } i \in \{1,\dots,m-1\} \text{ such that } y \in [a_i, a_{i+1}] \\ +\infty & \text{if } y \notin [a_1, a_m] \end{cases}$