Convex Optimization - HW1

Exercice 1

1) Un rectangle est <u>convexe</u> car un rectangle est une intersection de demi-espace.

$$\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i=1,\dots,n\}$$
$$= \bigcap_{i=1}^{n} \left(\{x \in \mathbb{R}^n,\ x_i \geq \alpha_i\} \cap \{x \in \mathbb{R}^n,\ x_i \leq \beta_i\} \right)$$

2) Par définition d'un espace convexe:

$$D = \{x \in \mathbb{R}_+^2,\ x_1 x_2 \geq 1\}$$

$x, y \in D$, $\theta \in [0,1]$, $(x = (x_1, x_2), y = (y_1, y_2))$

$$\theta x + (1-\theta) y \in D \iff (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) \geq 1$$
$$\iff \text{(A)}\ \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)[x_1 y_2 + y_1 x_2] \geq 1$$

Comme $x, y \in D$, $x_1 y_2 + y_1 x_2 \geq \frac{y_2}{x_2} + \frac{x_2}{y_2} \geq 2$

car $f : \begin{array}{l} \mathbb{R}_+^* \to \mathbb{R} \\ x \mapsto x + \frac{1}{x} \end{array}$ est minorée par 2.

Par conséquent,

$$\text{(A)}\ \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)[x_1 y_2 + y_1 x_2]$$
$$\geq \theta^2 + (1-\theta)^2 + 2\theta(1-\theta) \geq (\theta + (1-\theta))^2$$
$$\geq 1$$

d'où $\theta x + (1-\theta) y \in D$

Donc D est <u>convexe</u>.

1/4

3) Montrons que

$$D = \{x, \|x - x_0\|_2 \leq \|x-y\|_2 \ \forall y \in S\}, \quad S \subseteq \mathbb{R}^n, \ x_0 \in \mathbb{R}^n$$

est convexe.

$$D = \bigcap_{y \in S} \{x, \|x - x_0\|_2 \leq \|x - y\|\}, \text{ donc}$$

D est une intersection de demi espace.

En effet, $\forall y \in S$, $x \in \{x, \|x - x_0\|_2 \leq \|x-y\|\}$,

x s'écrit $x = x_0 + u$ avec $\left(\frac{y - x_0}{\|y - x_0\|}\right)^T u \leq \frac{\|y-x_0\|}{2}$

Ainsi, par intersection d'espace convexe, D est convexe.

4) Montrons que $D = \{x, \text{dist}(x,S) \leq \text{dist}(x,T)\}$, $S, T \subseteq \mathbb{R}^n$ n'est pas convexe.

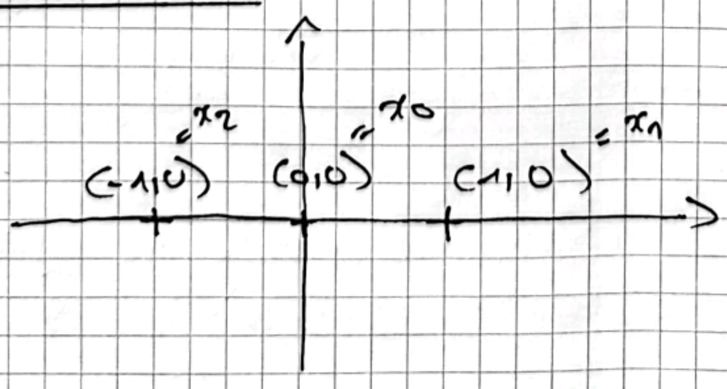
Dans $\mathbb{R}^2$, on considère : $T = \{(0,0)\}$

$S = \{(-1,0), (1,0)\}$

On a alors $(-1,0) \in D$, $(1,0) \in D$

mais $\frac{1}{2}(-1,0) + \frac{1}{2}(1,0) = (0,0) \notin D$

Donc avec ce contre exemple D n'est pas convexe.



$T = x_0$

$S = x_1 \cup x_2$

5) $S_1, S_2 \subseteq \mathbb{R}^n$ avec S_1 convexe,

$$D = \{ x \mid x + S_2 \subseteq S_1 \}.$$

Montrons que D est convexe.

Soit $D^* = \bigcap_{s_2 \in S_2} (S_1 - s_2)$, $D^* = D$.

En effet, $x \in D \Rightarrow \forall s_2 \in S_2,\ x + s_2 \in S_1$
$\Rightarrow \forall s_2 \in S_2,\ x \in S_1 - s_2$
$\Rightarrow x \in \bigcap_{s_2 \in S_2} (S_1 - s_2)$

$x \in D^* \Rightarrow \forall s_2 \in S_2 \quad x \in S_1 - s_2$
$\Rightarrow \forall s_2 \in S_2 \quad x + s_2 \in S_1$
$\Rightarrow x \in D$

D^* est une intersection de convexe donc D est bien convexe.

Exercice 2 -

Dans cet exercice, nous allons majoritairement utiliser la propriété suivante :
$f : \mathbb{R}^n \to \mathbb{R}$, si f est deux fois différentiables et $\mathrm{dom}(f)$ convexe, alors :

$$f \text{ convexe} \Leftrightarrow \nabla^2 f(x) \succeq 0 \quad \forall x \in \mathrm{dom} f$$

1) $f : \begin{cases} \mathbb{R}^2_{++} \to \mathbb{R} \\ (x_1, x_2) \mapsto x_1 x_2 \end{cases}$

f est deux-fois différentiable, $\mathrm{dom} f = \mathbb{R}^2_{++}$.

$\forall x \in \mathbb{R}^{2}_{++}$, $\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \nabla^2 f$

∇f est la matrice rotation.

Donc : $\nabla f^2 = I_2$. De plus $\nabla f \in S^2$ donc ∇f est diagonalisable.

$\Rightarrow \sigma(\nabla f) \subset \{-1, 1\}$.

Si $-1 \notin \sigma(\nabla f)$, alors ∇f est semblable à I_2, ce qui n'est pas le cas donc, par contradiction $-1 \in \sigma(\nabla f)$.

Par conséquent : $\pm \nabla f \notin S^2_+$ donc

<u>f n'est pas convexe ni concave.</u>

<u>option</u> : Montrons que f est quasiconcave.

On utilise le résultat de la question 2 de l'exercice 1.

f quasiconcave $\Leftrightarrow \forall \alpha \in \mathbb{R}$, $S_\alpha = \{x \in \mathbb{R}^2_{++}, -f(x) \leq \alpha\}$ est convexe.

Soit $\alpha \in \mathbb{R}$,

$$S_\alpha = \{x \in \mathbb{R}^2_{++}, -x_1 x_2 \leq \alpha\}$$
$$= \{x \in \mathbb{R}^2_{++}, x_1 x_2 \geq -\alpha\}$$

- cas 1 : $\alpha \geq 0$. alors $S_\alpha = \mathbb{R}^2_{++}$ donc S_α est convexe.
- cas 2 : $\alpha < 0$: $S_\alpha = \{x \in \mathbb{R}^2_{++}, x_1 x_2 \geq |\alpha|\}$

$$= \left\{x \in \mathbb{R}^2_{++}, \frac{x_1}{\sqrt{|\alpha|}} \frac{x_2}{\sqrt{|\alpha|}} \geq 1\right\}$$

Soit $D = \{x \in \mathbb{R}^2_{++}, x_1 x_2 \geq 1\}$

$f_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto \frac{1}{\sqrt{|\alpha|}} x$

D est convexe, f est linéaire donc $S_\alpha = f^{-1}(D)$ est

Ainsi, f est bien quasiconcave.

2) $f: \begin{array}{|l} \mathbb{R}^2_{++} \longrightarrow \mathbb{R} \\ (x_1,x_2) \longmapsto 1/x_1x_2 \end{array}$

On a: $\nabla^2 f(x) = \begin{bmatrix} \frac{2}{x^3 y} & \frac{1}{x^2y^2} \\ \frac{1}{x^2y^2} & \frac{2}{y^3x^3} \end{bmatrix}$

$x \in \mathbb{R}^2_{++}$, $x = (x,y)$

Soit λ_1, λ_2 les racines du polynôme caractéristique
On a :

$$\begin{cases} \lambda_1\lambda_2 = \det(\nabla^2 f(x)) = \frac{4}{x^4y^4} > 0 \\ \lambda_1+\lambda_2 = \mathrm{Tr}(\nabla^2 f(x)) = \frac{2}{x^3y} + \frac{2}{y^3x} > 0 \end{cases}$$

Les racines ont même signe et leur somme est positive.
Ainsi : $\lambda_1 > 0$, $\lambda_2 > 0$ donc $\nabla f^2(x) \succcurlyeq 0$,

f est convexe sur $\mathbb{R}^2_{++}$

▸ Montrons que f est quasiconvexe.
Soit $\alpha \in \mathbb{R}$, $S_\alpha = \{x \in \mathbb{R}^2_{++}, f(x) \leq \alpha\}$

$$= \{x \in \mathbb{R}^2_{++}, \frac{1}{x_1x_2} \leq \alpha\}$$

$$= \{x \in \mathbb{R}^2_{++}, x_1x_2 \geq 1/\alpha\}$$

$$\overset{\alpha \neq 0}{=} \{x \in \mathbb{R}^2_{++}, -x_1x_2 \leq -\frac{1}{\alpha}\}$$

$$= S^{g}_{-1/\alpha}, \text{ avec } g: x \longmapsto -x_1x_2$$

g est quasiconvexe, donc $S^g_{-1/\alpha}$ et donc S_α convexe &
Si $S_{0=\alpha} = \emptyset$. Ainsi: f est quasiconvexe.

3) $f : \begin{cases} \mathbb{R}^2_{+*} \longrightarrow \mathbb{R} \\ (x_1, x_2) \longmapsto x_1/x_2 \end{cases}$

$\forall x \in \mathbb{R}^2_{+*}, \nabla f^2(x) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$, $\lambda_1 \lambda_2$ racines du polynôme caractéristiques

$\begin{cases} \det(\nabla f^2(x)) = \lambda_1 \lambda_2 = -\frac{1}{x_2^4} < 0 \\ Tr(\nabla f^2(x)) = \lambda_1 + \lambda_2 = \frac{2x_1}{x_2^3} > 0 \end{cases}$

$\lambda_1 \lambda_2 < 0$ donc ∇f et $-\nabla f$ ne sont pas des matrices symétriques positives.

<u>f n'est ni convexe, ni concave.</u>

- Montrons que f est quasiconvexe.

Soit $\alpha \in \mathbb{R}$, $X_1, X_2 = (x_1, y_1), (x_2, y_2) \in S_\alpha$

$\theta \in]0,1[$,

$$f(\theta X_1 + (1-\theta)X_2) = \frac{\theta x_1 + (1-\theta)x_2}{\theta y_1 + (1-\theta)y_2} = \frac{\theta \frac{x_1}{y_1 y_2} + (1-\theta)\frac{x_2}{y_2 y_1}}{\frac{\theta}{y_2} + \frac{1-\theta}{y_1}}$$

$X_1, X_2 \in S_\alpha \Rightarrow \frac{x_1}{y_1} \leq \alpha, \frac{x_2}{y_2} \leq \alpha$

$$f(\theta X_1 + (1-\theta)X_2) \leq \alpha \frac{\frac{\theta}{y_2} + \frac{1-\theta}{y_1}}{\frac{\theta}{y_2} + \frac{1-\theta}{y_1}} = \alpha$$

Donc : S_α convexe, f est quasiconvexe.

4) $\alpha \in [0,1]$,

$$f_\alpha : \left| \begin{array}{l} \mathbb{R}^2_{++} \longrightarrow \mathbb{R} \\ (x_1,x_2) \longmapsto x_1^\alpha x_2^{1-\alpha} \end{array} \right.$$

si $\alpha = 0$ ou $\alpha = 1$, f est convexe car linéaire (et concave)

$\alpha \in]0,1[$, $x \in \mathbb{R}^2_{++}$

$$\nabla^2 f(x) = \begin{pmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & (1-\alpha)\alpha x_1^{\alpha-1}x_2^{-\alpha} \\ (1-\alpha)\alpha x_1^{\alpha-1}x_2^{-\alpha} & +\alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1} \end{pmatrix}$$

λ_1, λ_2 vp,

$$\begin{cases} \lambda_1\lambda_2 = \det(\nabla^2 f(x)) = \alpha^2(\alpha-1)^2 x_1^{2\alpha-2}x_2^{-2\alpha} - \alpha^2(\alpha-1)^2 x_1^{2\alpha-2}x_2^{-2\alpha} \\ \qquad = 0 \\ \lambda_1+\lambda_2 = \mathrm{Tr}[\nabla^2 f(x)] = \alpha\underbrace{(\alpha-1)}_{<0}\left(x_1^{\alpha-2}x_2^{1-\alpha} + x_1^{\alpha}x_2^{-\alpha-1}\right) \\ \qquad < 0 \end{cases}$$

Ainsi, $\nabla^2 f(x)$ a une valeur propre nulle et une positive.

On a donc $-\nabla^2 f(x) = \nabla^2(-f)(x) \succcurlyeq 0$,

donc f est concave pour $\alpha \in]0,1[$.

▸ f est aussi quasiconcave

$-f$ est convexe. Soit $\alpha \in \mathbb{R}$, $S_\alpha = \{x \in \mathbb{R}^2_{++}, -f(x) \leq \alpha\}$

$\theta \in]0,1[$, $x_1, x_2 \in S_\alpha$

$$-f(\theta x_1 + (1-\theta)x_2) \leq \theta(-f(x_1)) + (1-\theta)(-f(x_2))$$
$$\leq \alpha[\theta + (1-\theta)] = \alpha$$

Donc $-f$ est quasiconvexe, f est quasiconcave.

Exercice 3 :

1)

$$f : \left| \begin{array}{l} S_{++}^n \longrightarrow \mathbb{R} \\ X \longmapsto Tr[X^{-1}] \end{array} \right.$$

Soit $V \in S_{++}^n$, $X \in S_{++}^n$

$$g : t \longmapsto f(X+tV)$$

$t \in \mathbb{R}$, $g(t) = f(X+tV) = Tr[(X+tV)^{-1}]$

$X > 0$, donc $X = X^{1/2}X^{1/2}$, $X^{1/2} > 0$,

$$\longrightarrow = Tr[(X^{1/2}X^{1/2} + tV)^{-1}]$$

$$= Tr[X^{-1/2}[I_n + tX^{-1/2}VX^{-1/2}]^{-1}X^{-1/2}]$$

$$= Tr[X^{-1}[I_n + tX^{-1/2}VX^{-1/2}]^{-1}]$$

$X^{-1/2}VX^{-1/2} \in S^n$, donc d'après le théorème spectral $X^{-1/2}VX^{-1/2} = P^T diag(\sigma_1, \dots, \sigma_n)P$

$t \in \mathbb{R}$, $g(t) = Tr[X^{-1}[P^T[I_n + t\, diag(\sigma_i)]P]^{-1}]$

$$= Tr[X^{-1}P^T[I_n + t\, diag(\sigma_i)_i]^{-1}P]$$

$$= Tr[PX^{-1}P^T[I_n + t\, diag(\sigma_i)_i]^{-1}]$$

$$g(t) = \sum_{i=1}^{n} \alpha_i \frac{1}{1+\sigma_i t}, \quad (\alpha_i)_i \in \mathbb{R}^n$$

Par composition, $i \leq n$, $t \mapsto \frac{1}{1+\sigma_i t}$ est convexe car $t \mapsto \frac{1}{t}$ est convexe sur $\mathbb{R}_+^*$, décroissante et $t \mapsto 1+\sigma_i t$ est concave car linéaire.

Ainsi, g est une somme de fonction convexe donc g est convexe donc f est convexe.

précision : on a $\forall i \leq n, \sigma_i > 0$ donc on a pas de problème de définition de g.
En effet, $\forall x \in \mathbb{R}^n$,

$$\langle X^{-1/2} V X^{-1/2}, x \rangle = x^T X^{-1/2} V X^{-1/2} x$$
$$= (X^{-1/2} x)^T V (X^{-1/2} x)$$
$$> 0$$

donc $X^{-1/2} V X^{-1/2} \in S^n_{++}$ donc $\forall i \leq n, \sigma_i > 0$.

2) $f : \begin{array}{|l} S^n_{++}, \mathbb{R}^n \longrightarrow \mathbb{R} \\ (X, y) \longmapsto y^T X^{-1} y \end{array}$

On introduit :

$g : \begin{array}{|l} S^n_{++} \times \mathbb{R}^n \times \mathbb{R}^n \\ (X, y, z) \longmapsto y^T X^{-1} y + z^T X z - 2 z^T y \end{array}$

Soit $X \in S^n_{++}$, $y \in \mathbb{R}^n$,

$g_{X,y} : z \longmapsto g(X, y, z)$

$g_{X,y}$ est différentiable.

$z \in \mathbb{R}^n$, $\nabla g(z) = (X^T + X) z - 2y$
$= 2Xz - 2y$

$\nabla g(z^*) = 0 \iff z = X^{-1} y$

Comme $g_{X,y}$ est convexe en z sur $\mathbb{R}^n$, et infini à l'infini ($\|g_{X,y}(z)\| \xrightarrow[\|z\| \to +\infty]{} +\infty$) alors g atteint son minimum en $z^* = X^{-1} y$.

3/4

Par conséquent :

$\forall z \in \mathbb{R}^n,$

$$g_{X,y}(z) \geq g_{X,y}(z^*) = 0$$

$$\Leftrightarrow \quad y^T X^{-1} y + z^T X z \geq 2 z^T y$$

$$\Leftrightarrow \quad y^T X^{-1} y \geq 2 z^T y - z^T X z$$

On a égalité pour $z^* = X^{-1} y$,

$$2 z^{*T} y - z^T X z = 2 y^T X^{-1} y - y^T X^{-1} X X^{-1} y$$
$$= 2 y^T X^{-1} y - y^T X^{-1} y$$
$$= y^T X^{-1} y = f(X,y)$$

Ainsi : $f(X,y) = \sup_{z \in \mathbb{R}^n} \cancel{g_{X,y}(z)},\ 2 z^T y - z^T X z$

---

f est convexe comme supremum de fonctions convexes pour tout $z \in \mathbb{R}^n$.

3)

$$f : \left| \begin{array}{l} S^n \to \mathbb{R} \\ X \mapsto \sum_{i=1}^{n} \sigma_i(X) \end{array} \right.$$

On montre que $\forall X \in S^n_+,\ f(X) = \sup_{\sigma_m(Q) \leq 1} \langle Q, X \rangle$,

avec σ_m la norme spectrale.

$X = U \Sigma V^T$ (décomposition en valeur singulière)

$Q = U V^T,\ \sigma_m(Q) = 1.$

On a : $\langle Q, X \rangle = \mathrm{Tr}[V U^T U \Sigma V^T] = \mathrm{Tr}(V^T V U^T U \Sigma)$
$= \mathrm{Tr}(\Sigma) = f(X)$

Le sup est donc atteint. Montrons que c'est effectivement la plus grande valeur :

$$\sup_{\sigma_{m}(Q)\leq 1} \langle Q, X\rangle = \sup_{\sigma_{m}(Q)\leq 1} Tr[Q^T U\Sigma V^T]$$

$$= \sup_{\sigma_{m}(Q)\leq 1} Tr[U^T Q^T U \Sigma]$$

$$= \sup_{\sigma_{m}(Q)\leq 1} \langle U^T Q V, \Sigma\rangle$$

$$= \sup_{\sigma_{m}(Q)\leq 1} \sum_{i=1}^{n} \sigma_i (U^T Q V)_{ii}$$

$X = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$

$$= \sup_{\sigma_{m}(Q)\leq 1} \sum_{i=1}^{n} \sigma_i u_i^T Q v_i$$

$$\leq \sup_{\sigma_{m}(Q)\leq 1} \sum_{i=1}^{n} \sigma_i \sigma_{max}(Q) = \sum_{i=1}^{n} \sigma_i = f(X)$$

Par conséquent, on a bien :

$$f(X) = \sup_{\sigma_{m}(Q)\leq 1} \langle Q, X\rangle,$$

et $Q \mapsto \langle Q, X\rangle$ est linéaire donc convexe.

On peut donc conclure que f est convexe.

Exercice 4 :

1) On doit vérifier que :

i) K_n^+ est fermé

ii) K_n^+ est solide

iii) K_n^+ ne contient pas de ligne

i) $K_n^+ = \{x \in \mathbb{R}^n, x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$

$$= \bigcap_{i=1}^{n-1} \{x_i \geq x_{i+1}\} \cap \{x_n \geq 0\}$$

K_n^+ est une intersection de demi-espaces donc K_n^+ est fermé

ii) K_n^+ n'est pas l'ensemble vide

$x = (n, \dots, 1) \in K_n^+$

iii) si $x \in K_n^+$ et $-x \in K_n^+$, alors

$\forall i \leq n$, $x_i \geq 0$ et $-x_i \geq 0$ donc

$x = 0$

<u>Conclusion:</u> K_n^+ est un cône propre

2_ $K_n^{+*} = \{y, y^T x \geq 0 \quad \forall x \in K_n^+\}$

$= \{y, \sum_{i=1}^{n} y_i x_i \geq 0 \ \forall x \in K_n^+\}$

En prenant $x_r = (1, \dots, \underset{r}{1}, 0, \dots, 0) \in K_n^+$,

on observe que $y \in K_n^{+*} \Rightarrow \sum_{i=1}^{r} y_i \geq 0 \ \forall r \in [\![1, n]\!]$

Montrons que c'est une équivalence.

On peut montrer par récurrence que:

$$\sum_{i=1}^{n} x_i y_i = \sum_{k=1}^{n-1} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i + x_n \sum_{i=1}^{n} y_i$$

▸ $n = 1$: $x_1 y_1 = x_1 y_1$

▸ vrai pour n, $n+1$?

$$\sum_{i=1}^{n+1} x_i y_i = x_{n+1} y_{n+1} + x_n \sum_{i=1}^{n} y_i + \sum_{k=1}^{n-1} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i$$

$$= x_{n+1} \sum_{i=1}^{n+1} y_i - x_{n+1} \sum_{i=1}^{n} y_i + \sum_{k=1}^{n-1} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i + x_n \sum_{i=1}^{n} y_i$$

$$= x_{n+1} \sum_{i=1}^{n+1} y_i + \sum_{k=1}^{n} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i$$

Par conséquent :

$$\sum_{i=1}^{n} x_i y_i = \underbrace{x_n \sum_{i=1}^{n} y_i}_{\geq 0} + \sum_{k=1}^{n-1} \underbrace{(x_k - x_{k+1})}_{\geq 0} \underbrace{\sum_{i=1}^{k} y_i}_{\geq 0}$$

par les y candidats.

On a donc : $\boxed{K_n^+ = \left\{ \sum_{i=1}^{k} y_i \geq 0 \ , \ k \leq n \right\}}$

Exercice 5 :

1) $\underline{\operatorname{dom} f^* = \{ y, \ y \in \mathbb{R}_+^n \text{ et } \sum_{i=1}^{n} y_i = 1 \}}$.

En effet :

- si $y \notin \mathbb{R}_+^n$, soit $k \leq n$, $y_k < 0$.
  Alors, en choisissant $x = (0 - 0 - \underset{k}{t} \ 0 - 0)$,
  on a : $f^*(y) \geq y^T x - \max(x_i)_i$
  $\geq t|y_k| \xrightarrow[t \to +\infty]{} +\infty$

- si $\sum_{i=1}^{n} y_i > 1$, alors : $x = (t, \dots, t)$
  $f^*(y) \geq y^T x - \max(x_i)_i$
  $\geq t \underbrace{\sum_{i=1}^{n} y_i}_{> 1} - t \xrightarrow[t \to +\infty]{} +\infty$

- si $\sum_{i=1}^{n} y_i < 1$, avec $x = (-t, \dots, -t)$
  $f^*(y) \geq -t \sum_{i=1}^{n} y_i + t \xrightarrow[t \to +\infty]{} +\infty$

Avec $y \in \operatorname{dom} f^*$, on a :
$\forall x \in \mathbb{R}^n$, $x^T y - \max(x_i) \leq 0$, égalité en $x = 0_n$.

Donc : $f(y^*) = 0$