Convex Optimization : Homework n°1

$\boxed{\text{Exo 1}}$

1) Let $A := \{x \in \mathbb{R}^n : \alpha_i \leq x_i \leq \beta_i \ , i = 1 \ldots n\}$ with $\begin{cases} \alpha \in \mathbb{R}^n \\ \beta \in \mathbb{R}^n \end{cases}$.

* Case n°1 : $\exists i_0 \in [\![1, n]\!] : \alpha_{i_0} > \beta_{i_0}$. Then $A = \emptyset$ and A is convex

* Case n°2 : $\forall i \in [\![1,n]\!] , \ \alpha_i \leq \beta_i$

Let $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$.

Suppose $x \in A$ and $y \in A$ [$A \neq \emptyset$ because $\alpha \in A$]

$\forall i \in [\![1,n]\!]$, we have $\begin{cases} \theta\alpha_i \leq \theta x_i \leq \theta\beta_i \\ (1-\theta)\alpha_i \leq (1-\theta) y_i \leq (1-\theta)\beta_i \end{cases}$

By summation,

$\forall i \in [\![1,n]\!], \ \theta\alpha_i + (1-\theta)\alpha_i = \alpha_i \leq \theta x_i + (1-\theta) y_i \leq \beta_i = \theta\beta_i + (1-\theta)\beta_i$

This shows that $\theta x + (1-\theta) y \in A$

Cl <u>In all cases, A is convex</u>

---

2) Let $A := \{x \in \mathbb{R}^2_+ : x_1 x_2 \geq 1\}$. We have $A \neq \emptyset$ because $(1,1) \in A$

Let $\begin{cases} x, y \in A. \\ \theta \in [0,1] \end{cases}$ we have $x_1 x_2 \geq 1$ and $y_1 y_2 \geq 1$

$$(\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) = \theta^2 \overbrace{x_1 x_2}^{\geq 1} + (1-\theta)^2 \overbrace{y_1 y_2}^{\geq 1} + \theta(1-\theta)(x_1 y_2 + y_1 x_2)$$

$$\geq \theta^2 + (1-\theta)^2 + \theta(1-\theta)(x_1 y_2 + y_1 x_2)$$

As $x_1 x_2 \geq 1$, we have $x_2 \neq \emptyset$ and $y_2 x_1 \geq \frac{y_2}{x_2}$

Similarly, $y_1 x_2 \geq \frac{x_2}{y_2}$. we thus have

$$(\theta x_1+(1-\theta)y_1)(\theta x_2+(1-\theta)y_2) \geq 1 + 2\theta(\theta-1) + \theta(1-\theta)\left[\frac{y_2}{x_2}+\frac{x_2}{y_2}\right] \quad (*)$$

Let $f(x) = x + \frac{1}{x}$ for $x \in \mathbb{R}_+^*$ we have f derivable and

$\forall x \in \mathbb{R}_+^*,\ f'(x) = 1 - \frac{1}{x^2}$

Let $x \in \mathbb{R}_+^*$

$$f'(x) = 0 \iff x^2 = 1 \iff x = 1$$

Consequently, as $f' \leq 0$ on $]0,1]$ and $f' \geq 0$ on $[1,+\infty[$,
f reaches a minimum in 1 and $f(1) = 2$
we thus have $f \geq 1$

Going back to $(*)$, we thus obtain

$$(\theta x_1+(1-\theta)y_1)(\theta x_2+(1-\theta)y_2) \geq 1 + 2\theta(\theta-1) + \theta(1-\theta)\,2$$
$$\geq 1$$

This shows that $\theta x + (1-\theta)y \in A$

Cl. <u>A is convex</u>

---

3) Let $A = \{x : \|x - x_0\|_2 \leq \|x-y\|_2 \text{ for all } y \in S\}$

Let's show that A is not convex for all S

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

and $x_0 = (0,\ldots,0)$. we take $\begin{cases} x_1 = (0{,}5, 0, \ldots, 0) \\ x_2 = (-0{,}5, 0, \ldots, 0) \end{cases}$

we have $\text{dist}(x_1, S) = \|x_1 - (1,0,\ldots,0)\| = 0{,}5 = \text{dist}(x_1, x_0)$

$\text{dist}(x_2, S) = 0{,}5 = \text{dist}(x_2, x_0)$

so $x_1, x_2 \in A$. But $0{,}5x_1 + (1-0{,}5)x_2 = x_0 \notin A$

Cl. <u>A is not convex for all S</u>

4) let $A := \{x : \text{dist}(x,S) \leq \text{dist}(x,T)\}$ where $S,T \subseteq \mathbb{R}^n$ and $\text{dist}(x,S) = \inf\{\|x-z\|_2,\ z \in S\}$

if we take $S = \{x_0\}$ with $x_0 \in \mathbb{R}^n$, it is question 3); and we have shown that it is not necessary convex for all S,T

Ccl: A is not convex for all $S,T \subseteq \mathbb{R}^n$

---

5) let $A = \{x : x + S_2 \subseteq S_1\}$ where $S_1, S_2 \subseteq \mathbb{R}^n$ and S

Let $x \in A$, $y \in A$ and $\theta \in [0,1]$. Let $s_2 \in S_2$

we have

$$\theta x + (1-\theta)y + s_2 = \underbrace{\theta (\overbrace{x+s_2}^{\in S_1 \text{ because } x \in A}) + (1-\theta)(\underbrace{y+s_2}_{\in S_1 \text{ because } y \in A})}_{\in S_1 \text{ because } S_1 \text{ is convex}}$$

Then,

$\theta x + (1-\theta) y + s_2 \in S_1$, which shows that A is convex

Ccl A is convex

Exo 2

1) $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$

we have f differentiable and

$\forall (x_1, x_2) \in \mathbb{R}^2_{++}$, $\nabla f(x_1, x_2) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$ and $\nabla^2 f(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

As the eigenvalues of $\nabla^2 f$ are 1 and -1 (because the caracteristic polynom is $X^2 - 1$), we can conclude that f is not convex and not concave

Ccl: f is neither convex nor concave

2) $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$

we have f differentiable and

$\forall (x_1, x_2) \in \mathbb{R}^2_{++}$, $\nabla f(x_1, x_2) = \begin{pmatrix} -\dfrac{1}{x_1^2 x_2} \\ -\dfrac{1}{x_1 x_2^2} \end{pmatrix}$ and $\nabla^2 f(x_1, x_2) = \begin{pmatrix} \dfrac{2}{x_1^3 x_2} & \dfrac{1}{x_1^2 x_2^2} \\ \dfrac{1}{x_1^2 x_2^2} & \dfrac{2}{x_1 x_2^3} \end{pmatrix}$

ie $\nabla^2 f(x_1, x_2) = x_1^3 x_2^3 \begin{pmatrix} 2x_2^2 & x_1 x_2 \\ x_1 x_2 & 2x_1^2 \end{pmatrix}$

Let's find the eigenvalues of $\nabla^2 f(x_1, x_2)$ for $x_1, x_2 \in \mathbb{R}^2_{++}$

$$\chi_{\nabla^2 f} = 0 \Leftrightarrow (X - 2x_2^2)(X - 2x_1^2) - x_1^2 x_2^2 = 0$$
$$\Leftrightarrow X^2 - 2(x_1^2 + x_2^2)X + 3x_1^2 x_2^2 = 0$$

$$\Delta = 4(x_1^2 + x_2^2)^2 - 12 x_1^2 x_2^2 = 4x_1^4 + 4x_2^4 - 4x_1^2 x_2^2$$
$$= 4(x_1^4 + x_2^4 - 2x_1^2 x_2^2 + x_1^2 x_2^2)$$
$$= 4(\underbrace{x_1^2 x_2^2}_{\geq 0} + \underbrace{(x_1^2 - x_2^2)^2}_{\geq 0}) \geq 0$$

We thus have 2 solutions in $\mathbb{R}$, we now show the positivity.

$$\lambda_1 = \frac{2(x_1^2+x_2^2)+\sqrt{\Delta}}{2} \geq 0 \text{ as all the terms are positive}$$

aupor $b = -2(x_1^2+x_2^2)$ et $4ac = 12x_1^2x_2^2$

$$\lambda_2 = \frac{-b-\sqrt{b^2-4ac}}{2}$$

$$\lambda_2 \geq 0 \iff \underbrace{-b}_{\geq 0} \geq \underbrace{\sqrt{b^2-4ac}}_{\geq 0} \underset{\geq 0}{\iff} b^2 \geq b^2-4ac \iff 0 \geq -4ac$$

which is true because $-12x_1^2x_2^2 \leq 0$

Ccl $\lambda_{1,2} \geq 0$ and $\nabla^2 f \succcurlyeq 0$ so f is convex
As f is not concave, ($\nabla^2 f \underset{No}{\preccurlyeq} 0$), f is only convex

3) $f(x_1, x_2) = x_1/x_2$ on $\mathbb{R}^2_{++}$
f is differentiable and
$\forall (x_1,x_2) \in \mathbb{R}^2_{++}$, $\nabla f(x_1,x_2) = \begin{pmatrix} 1/x_2 \\ -\frac{x_1}{x_2^2} \end{pmatrix}$ and $\nabla^2 f(x_1,x_2) = \begin{pmatrix} 0 & \frac{-1}{x_2^2} \\ \frac{-1}{x_2^2} & \frac{2x_1}{x_2^3} \end{pmatrix}$

Let $(x_1,x_2) \in \mathbb{R}^2_{++}$,

$\nabla^2 f(x_1,x_2) \succcurlyeq 0 \iff$ eigenvalues of $\nabla^2 f(x_1,x_2) \geq 0$

we have

$$\chi_{\nabla^2 f} = X\left(X - \frac{2x_1}{x_2^3}\right) - \frac{1}{x_2^4}$$

$$\chi_{\nabla^2 f} = 0 \iff X^2 - \frac{2x_1}{x_2^3}X - \frac{1}{x_2^4} = 0$$

$$\Delta = \frac{4x_1^2}{x_2^6} + \frac{4}{x_2^4} = 4x_2^{-6}(x_1^2+x_2^2) \geq 0$$

$$\lambda_{1,2} = \frac{\frac{2x_1}{x_2^3} \pm \sqrt{4x_2^{-6}(x_1^2+x_2^2)}}{2}$$. $\lambda_1 \geq 0$ and

same reasoning as 2)

$\lambda_2 \geq 0 \iff 0 \geq \frac{4}{x_2^4}$ which is false

we have $\lambda_2 < 0 \iff 0 < \frac{4}{x_2^4}$ which is true

Cl.

we have $\lambda_1 > 0$ and $\lambda_2 < 0$; which shows that f is neither convex nor concave

4) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$ where $0 \leq \alpha < 1$ on $\mathbb{R}^2_{++}$. Suppose $\alpha \neq 0$ for now

f is differentiable and

$\forall (x_1, x_2) \in \mathbb{R}^2_{++}$, $\nabla f(x_1, x_2) = \begin{pmatrix} \alpha x_1^{\alpha-1} x_2^{1-\alpha} \\ (1-\alpha) x_1^{\alpha} x_2^{-\alpha} \end{pmatrix}$ and $\nabla^2 f(x_1, x_2) = \begin{pmatrix} \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} ; \alpha(1-\alpha) x_1^{\alpha} \ldots \\ \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} ; -\alpha(1-\alpha) x_1^{\alpha} \ldots \end{pmatrix}$

Let $(x_1, x_2) \in \mathbb{R}^2_{++}$

$\det(\nabla^2 f(x_1, x_2)) = \alpha^2 (\alpha-1)^2 x_1^{2\alpha-2} x_2^{-2\alpha} - \alpha^2 (1-\alpha)^2 x_1^{2\alpha-2} x_2^{-2\alpha}$

$= \alpha^2 (\alpha-1)^2 \left[x_1^{2\alpha-2} x_2^{-2\alpha} - x_1^{2\alpha-2} x_2^{-2\alpha} \right] = 0$

$\chi_{\nabla^2 f(x_1, x_2)} = (X - \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha})(X - \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1}) - \alpha^2 (1-\alpha)^2 x_1^{2\alpha-2} x_2^{-2\ldots}$

$= X^2 - \alpha(\alpha-1) \left[x_1^{\alpha-2} x_2^{1-\alpha} + x_1^{\alpha} x_2^{-\alpha-1} \right] X + 0$

$= X \left[X - \alpha(\alpha-1) \left[x_1^{\alpha-2} x_2^{1-\alpha} + x_1^{\alpha} x_2^{-\alpha-1} \right] \right]$

$\chi_{\nabla^2 f(x_1, x_2)} = 0 \iff \begin{cases} x = 0 \\ x = \underbrace{\alpha(\alpha-1)}_{< 0} \underbrace{\left[x_1^{\alpha-2} x_1^{1-\alpha} + x_1^{\alpha} x_2^{-\alpha+1} \right]}_{\geq 0} \end{cases}$

Cl. if $\begin{cases} \alpha \neq 0 \\ \alpha \neq 1 \end{cases}$, $\nabla^2 f \preceq 0$ and f is concave and not convex

if $\alpha = 0$, $f(x_1, x_2) = x_2^{1-\alpha} = x_2$ which is concave and convex

if $\alpha = 1$, $f(x_1, x_2) = x_1^{\alpha} = x_1$ which is concave and convex

$\boxed{\text{Exo 3}}$

1) $f(X) = Tr(X^{-1})$ on S^n_{++}

Let $\begin{cases} X \in S^n_{++} \\ H \in S^n_{++} \end{cases}$

$f(X+H) - f(X) = Tr((X+H)^{-1} - X^{-1})$

we have $(X+H)^{-1} = X^{-1/2}(I + \underbrace{X^{-1/2}HX^{-1/2}}_{Y := X^{-1/2}HX^{-1/2}})^{-1} X^{-1/2}$

$= X^{-1/2} \underbrace{(I - Y + Y^2 - Y^3 + Y^4 - Y^5 + Y^6 \dots)}_{\text{because } (I - Y + Y^2 - Y^3 + \dots)(I+Y) = I} X^{-1/2}$

So

$f(X+H) - f(X) = Tr(X^{-1/2}[I - Y + Y^2 + O(\|Y\|^3) - I] X^{-1/2})$

$= Tr(X^{-1/2}[-Y + Y^2 + O(\|H\|^3)] X^{-1/2})$

$= Tr(X^{-1/2}[-X^{-1/2}HX^{-1/2} + X^{-1/2}HX^{-1}HX^{-1/2}]X^{-1/2} + O(\|$

$= \underbrace{Tr(-X^{-1}HX^{-1})}_{\text{first derivative}} + \underbrace{Tr(X^{-1}HX^{-1}HX^{-1})}_{\text{second derivative}} + O(\|H\|^3)$

we have $Tr(X^{-1}HX^{-1}HX^{-1}) = Tr(X^{-1/2}\underbrace{(X^{-1/2}HX^{-1/2})^2}_{Y^2 \in S^n_+} X^{-1/2}) \geq 0$

Cl: f is convex

2) $f(X,y) = y^T X^{-1} y$ on $\text{dom} f = S^n_{++} \times \mathbb{R}^n$

we can write $f(X,y) = \sup_{z \in \mathbb{R}^n} [2y^T z - z^T X z]$

[as $z_{max} = X^{-1}y$ and $f(X,y) = 2y^T X^{-1} y - y^T X^{-1} y = y^T X^{-1} y$]

Now, we have to show that $f(X,y,z) = 2y^T z - z^T X z$ is convex in (X,y) for each $z \in \mathbb{R}^n$. This is true because f is linear in (X,y) for a fixed z.

$\mathscr{C}$ <u>f is convex</u>

3) $f(X) = \sum_{i=1}^{n} \sigma_i(X)$ on $\operatorname{dom} f = S^n$

Let's show that $f(X) = \sup_{\sigma_1(Q) \le 1} \langle Q, X \rangle$

we use the spectral theorem to write $X = PDP^{-1}$ (diagonal D, orthogonal P)

we have

$$f(X) = \sup_{\sigma_1(Q) \le 1} \langle Q, X \rangle = \sup_{\sigma_1(Q) \le 1} Tr(Q^T P D P^{-1}) = \sup_{\sigma_1(Q) \le 1} Tr(P^{-1} Q^T P D)$$

$$= \sup_{\sigma_1(Q) \le 1} Tr(P^T Q^T P D) = \sup_{\sigma_1(Q) \le 1} \langle P^T Q P, D \rangle = \sup_{\sigma_1(Q) \le 1} \langle P^{-1} Q P, D \rangle$$

Hence, the function $\begin{cases} \{\sigma_1(Q) \le 1\} \to \{\sigma_1(Q) \le 1\} \\ Q \mapsto P^{-1} Q P \end{cases}$ is bijective, so we can write

$$f(X) = \sup_{\sigma_1(Q) \le 1} \langle Q, D \rangle$$

if we take $Q = \begin{cases} sgn(d_i) & \text{for } i \in [\![1,n]\!] \\ 0 & \text{else} \end{cases}$, we have $\langle Q, D \rangle = \sum_i \sigma_i$

otherwise, we always have $\langle Q, D \rangle \le \sum_i \sigma_i$ as all the eigenvalues of Q are less than 1.

we thus have

$f(X) = \sup_{\sigma_1(Q) \le 1} Tr(Q^T X)$, which is evidently convex as the function $Tr(Q^T X)$ is convex as it is ~~def~~ linear in X.

Exo 5

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

we define

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

we have for $x, y \in \mathbb{R}^n$, if $y \succeq 0$ and $\max x_i \geq 0$

$$y^T x - \max_{i=1-n} x_i = \sum_{i=1}^n y_i x_i - \max_{i=1-n} x_i \leq \max_{i=1-n} x_i \left[\sum_{i=1}^n y_i - 1\right]$$

Let $y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$

If we take $x = \alpha . \mathbb{1}$, then

$$y^T x - \max x_i = \alpha \sum_{i=1} y_i - \alpha = \alpha\left(\sum_{i=1} y_i - 1\right)$$

- if $\sum_{i=1}^n y_i > 1$, then with $\alpha \to -\infty$, $f^*(y) = +\infty$
- if $\sum_{i=1}^n y_i < 1$, we also have $f^*(y) = +\infty$
- if $\sum_{i=1}^n y_i = 1$ if we find i_0 such that $y_{i_0} < 0$, then we take $x_{i_0} = -|\alpha| < 0$ and $x_i = 0$ otherwise and we have $f^*(y) = +\infty$ with $\alpha \to -\infty$
- if $\sum_{i=1}^n y_i = 1$ and $y \succeq 0$, we have

$$\sum_{i=1}^n x_i y_i - \max_{i=1-n} x_i \leq \sum_{i=1}^n y_i \times \max_{i=1-n} x_i - \max_{i=1-n} x_i$$

$$\leq 0 \text{ with equality if } x = 0.$$

we thus have $f^*(y) = 0$

Ccl. $f^*(y) = \begin{cases} 0 & \text{if } y \succeq 0 \text{ and } \sum y_i = 1 \\ +\infty & \text{otherwise} \end{cases}$

$\boxed{\text{Exo } 2}$ Quasi concavity / vexity

1) $f(x_1, x_2)$ on $\mathbb{R}^2_{++}$ is quasiconcave. $\alpha \in \mathbb{R}^*_+$ [the other case, $S_\alpha = \ldots$

$S_\alpha := \{x \in \mathbb{R}^2_{++} : f(x) \geq \alpha\}$.

$x, y \in S_\alpha$; $\lambda \in [0,1]$. we have

$$(\lambda x_1 + (1-\lambda)y_1)(\lambda x_2 + (1-\lambda)y_2) \geq \lambda^2 \alpha + (1-\lambda)^2 \alpha + \lambda(1-\lambda)\big[\underbrace{x_1 y_2 + x_2 y_1}_{\geq \frac{\alpha x_1}{y_1} + \alpha \text{ as in ex}}\big]$$

$$\geq \lambda^2 \alpha + (1-\lambda)^2 \alpha + \lambda(1-\lambda)\big[\underbrace{\alpha \frac{x_1}{y_1} + \alpha \frac{y_1}{x_1}}_{\geq 2\alpha \text{ as shown i}}\big]$$

$$\geq \alpha + 2\lambda^2\alpha - 2\lambda\alpha + 2\alpha\lambda - 2\lambda^2\alpha$$

$$\geq \alpha.$$

Cl: f is quasiconcave

2) $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$ on $\mathbb{R}^2_{++}$

Let $\alpha \in \mathbb{R}^*_+$ [otherwise $S_\alpha = \emptyset$]

$$S_\alpha = \{x \in \mathbb{R}^2_{++} : \frac{1}{x_1 x_2} \leq \alpha\} = \{x \in \mathbb{R}^2_{++} : x_1 x_2 \geq \frac{1}{\alpha}\}.$$

we have shown in the precedent question that S_α is convex

Cl: f is quasi convex.