Convex Optimization
Homework 1

Eya Ghamgui

Exercice n 1: Which of the following sets are convex?

1) Rectangle: $R = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, \quad i = 1, \dots, m\}$

Let's have $x, y \in \mathbb{R}^n$ such that $x, y \in R$.

Let's take $\theta \in [0,1]$.

We have $\alpha_i \leq x_i \leq \beta_i$ (1) for all $i = 1, \dots, m$

and $\alpha_i \leq y_i \leq \beta_i$ (2) for all $i = 1, \dots, m$.

Multiplying (1) with $\theta \geq 0$ gives: $\theta \alpha_i \leq \theta x_i \leq \theta \beta_i$

Multiplying (2) with $(1-\theta) \geq 0$ gives: $(1-\theta)\alpha_i \leq (1-\theta) y_i \leq (1-\theta)\beta_i$

Summing the two previous results, we get:

$$\theta \alpha_i + (1-\theta)\alpha_i \leq \theta x_i + (1-\theta) y_i \leq (1-\theta)\beta_i + \theta \beta_i$$

$$\alpha_i \leq \theta x_i + (1-\theta) y_i \leq \beta_i \quad \forall i = 1, \dots, m$$

Thus, we get $\theta x + (1-\theta) y \in R$.

Conclusion: $\left.\begin{array}{l} x, y \in R \\ \theta \in [0,1] \end{array}\right\} \Rightarrow \theta x + (1-\theta) y \in R$

We can conclude that R is a convex set.

2) Hyperbolic set: $H = \{x \in \mathbb{R}_+^2 \mid x_1 x_2 \geq 1\}$

Let's take $x, y \in H$ and $\theta \in [0,1]$

We can write $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

We get $x_1 x_2 \geq 1$ and $y_1 y_2 \geq 1$.

$$\Rightarrow \theta x + (1-\theta) y = \begin{pmatrix} \theta x_1 + (1-\theta) y_1 \\ \theta x_2 + (1-\theta) y_2 \end{pmatrix}$$

$$\Rightarrow [\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2]$$

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

Now, we have $x_1 x_2 \geq 1 \Rightarrow x_1 \geq \frac{1}{x_2}$

$$\Rightarrow \boxed{x_1 y_2 \geq y_2 / x_2} \quad (1)$$

$$y_1 y_2 \geq 1 \Rightarrow y_1 \geq \frac{1}{y_2} \Rightarrow y_1 x_2 \geq \frac{x_2}{y_2}$$

$$\geq \frac{1}{y_2/x_2}$$

$$\Rightarrow \boxed{y_1 x_2 \geq \frac{1}{y_2/x_2}}$$

① + ② gives : $x_1 y_2 + y_1 x_2 \geq y_2/x_2 + \frac{1}{y_2/x_2}$ ②

$$\Rightarrow x_1 y_2 + y_1 x_2 - 2 \geq y_2/x_2 + \frac{1}{y_2/x_2} - 2$$

Now we can write: $y_2/x_2 + \frac{1}{y_2/x_2} - 2$

$$= \sqrt{y_2/x_2}^2 + \frac{1}{\sqrt{y_2/x_2}^2} - 2\frac{\sqrt{y_2/x_2}}{\sqrt{y_2/x_2}}$$

$$= \left(\sqrt{\frac{y_2}{x_2}} - \frac{1}{\sqrt{\frac{y_2}{x_2}}}\right)^2 \geq 0$$

We can conclude : $x_1 y_2 + y_1 x_2 \geq 2$

We have $\theta(1-\theta) \geq 0 \Rightarrow \theta(1-\theta)(x_1 y_2 + y_1 x_2) \geq 2(\theta(1-\theta))$

Now we come back to the previous expression:

$$[\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2]$$

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

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

$$\geq \theta^2 + (1-\theta)^2 + 2\theta(1-\theta)$$

$$\geq \cancel{\theta^2} + 1 + \theta^2 - \cancel{2\theta} + \cancel{2\theta} - 2\theta^2 = 1$$

Conclusion :

$$\left.\begin{array}{r} x, y \in H \\ \theta \in [0,1] \end{array}\right\} \Rightarrow \theta x + (1-\theta) y \in H$$

We can conclude that H is a convex set.

3) The set of points closer to a given point than a given set.

$$A = \{ x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S \}$$

where $S \subseteq \mathbb{R}^n$.

Let's take $x \in A$, we have:

$\| x - x_0 \|_2 \leq \| x - y \|_2$ for all $y \in S$.

$\Rightarrow (x - x_0)^T (x - x_0) \leq (x - y)^T (x - y)$

$\Rightarrow (x^T - x_0^T)(x - x_0) \leq (x^T - y^T)(x - y)$

$\Rightarrow \cancel{x^T x} - x^T x_0 - x_0^T x + x_0^T x_0 \leq \cancel{x^T x} - x^T y - y^T x + y^T y$

$\Rightarrow -2 x_0^T x + x_0^T x_0 \leq -2 y^T x + y^T y$

$\Rightarrow 2(y^T - x_0^T) x \leq y^T y + x_0^T x_0$

$\Rightarrow (y - x_0)^T x \leq \dfrac{y^T y + x_0^T x_0}{2}$

This is the equation of a half space written as $a^T x \leq b$ with:

$$\begin{cases} a = y - x_0 \\ b = \dfrac{y^T y + x_0^T x_0}{2} \end{cases}$$

Conclusion: $A = \bigcap_{y \in S} \{ x \mid a^T x \leq b \}$: the intersection of halfspaces

In addition, halfspaces are convex and the intersection of convex sets is convex, that's why A is a convex set as it is considered as the intersection of halfspaces.

4) The set of points closer to one set than another:

$$\mathcal{C} = \{ x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T) \}$$

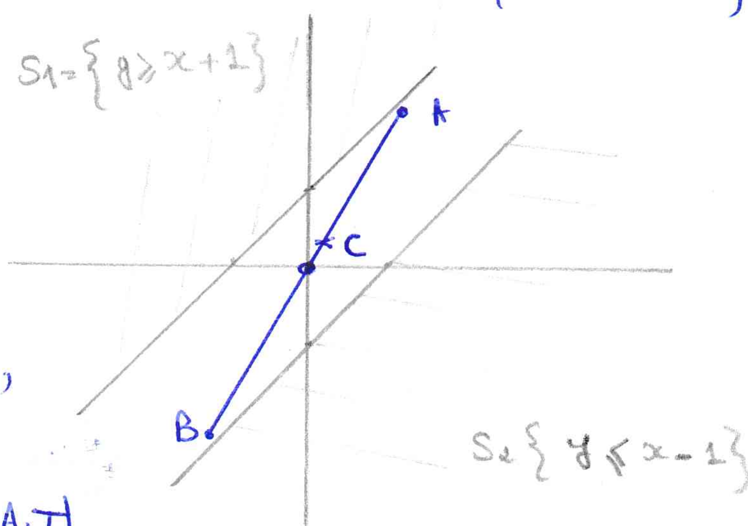
where $S, T \subseteq \mathbb{R}^n$ and $\operatorname{dist}(x, S) = \inf \{ \| x - z \|_2 , z \in S \}$

In this example, we will work with:

$n = 2$

$T = \{ \vec{0} \}$

$S = S_1 \cup S_2$



For the point $A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $B \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, it is clear that:

$\operatorname{dist}(A, S) < \operatorname{dist}(A, T)$

$\operatorname{dist}(B, S) \leq \operatorname{dist}(B, T)$

However, for the point C which belong to the segment $[A, B]$:

$\operatorname{dist}(C, T) \leq \operatorname{dist}(C, S)$.

This result let us to conclude that $\mathcal{G}$ is not a convex set because taking A and B in $\mathcal{G}$, the point C which belong to the segment $[A,B]$ is not in $\mathcal{G}$.

5) The set $A = \{x \mid x + S_2 \subseteq S_1\}$ where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 is a convex set.

We have for $x \in A \Rightarrow x + S_2 \subseteq S_1$

$\Leftrightarrow x + y \in S_1$ for all $y \in S_2$

From this result, we can write:

$$A = \{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\}$$

$$= \bigcap_{y \in S_2} \{x \mid x \in S_1 - y\} = \bigcap_{y \in S_2} \{S_1 - y\}$$

We have S_1 is a convex set:

$$\left.\begin{array}{l} x_1, x_2 \; S_1 \\ \theta \in [0,1] \end{array}\right\} \begin{array}{l} \theta x_1 + (1-\theta) x_2 \in S_1 \\ \Rightarrow \theta x_1 + (1-\theta) x_2 - y \in S_1 - y \\ \Rightarrow \theta (x_1 - y) + (1-\theta)(x_2 - y) \in S_1 - y \end{array}$$

$\Rightarrow x_1 - y, \; x_2 - y \in \{S_1 - y\}$

$\theta \in [0,1] \Rightarrow$ we have $\theta(x_1 - y) + (1-\theta)(x_2 - y) \in S_1 - y$

Thus, $\{S_1 - y\}$ is a convex set.

The intersection of convex sets is a convex set.

Conclusion: $A = \bigcap_{y \in S_2} \{S_1 - y\}$ is a convex set.

Exercice n°2: Convex, Concave, quasi convex, quasiconcave.

1) $f(x_1, x_2) = x_1 x_2$ $\quad \mathrm{dom}\, f = \mathbb{R}^2_{++}$.

- $\mathrm{dom}\, f = \mathbb{R}^2_{++}$ is a convex set.
- we have f is twice differentiable.

$$\begin{cases} \dfrac{\partial f}{\partial x_1} = x_2 \\ \dfrac{\partial f}{\partial x_2} = x_1 \end{cases} \Rightarrow \begin{cases} \dfrac{\partial^2 f}{\partial^2 x_1} = 0 & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} = 1 \\ \dfrac{\partial^2 f}{\partial^2 x_2} = 0 & \dfrac{\partial^2 f}{\partial x_2 \partial x_1} = 1 \end{cases}$$

$$\Rightarrow \nabla^2 f(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\det(\nabla^2 f - \lambda I_d) = \det\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$$

we can deduce that the eigen values of the hessian matrix are $\{-1, 1\}$.

$\Rightarrow \nabla^2 f(x_1, x_2) \not\succeq 0$ and $\nabla^2 f(x_1, x_2) \not\preceq 0$

$\Rightarrow$ we can conclude that f is not a convex function nor a concave function.

- Now, let's take $S_\alpha = \{ x \in \mathbb{R}^2_{++} ;\ f(x) \geq \alpha \}$

$$= \{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} ;\ f(x_1, x_2) \geq \alpha \}$$
$$= \{ x_1 \in \mathbb{R}_{++} ;\ x_2 \in \mathbb{R}_{++} ;\ x_1 x_2 \geq \alpha \}$$

To show that f is a quasiconcave, we need to show that S_α is convex.

Let's take $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in S_\alpha$; $\theta \in [0,1]$

Such that $x_1 x_2 \geq \alpha$ and $y_1 y_2 \geq \alpha$.

$$\theta x + (1-\theta) y = \begin{pmatrix} \theta x_1 + (1-\theta) y_1 \\ \theta x_2 + (1-\theta) y_2 \end{pmatrix}$$

$$\Rightarrow [\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2]$$
$$= \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)[x_2 y_1 + x_1 y_2]$$

$$\begin{cases} \theta^2 x_1 x_2 \geq \alpha \theta^2 \\ (1-\theta)^2 y_1 y_2 \geq (1-\theta)^2 \alpha \\ \theta(1-\theta)[x_2 y_1 + x_1 y_2] \geq 2\alpha\theta(1-\theta) \end{cases}$$

In fact, the last inequality is determined as follow:

$$x_1 x_2 \geq \alpha \Rightarrow x_1 y_2 \geq \alpha \frac{y_2}{x_2}$$

$$y_1 y_2 \geq \alpha \Rightarrow y_1 x_2 \geq \alpha \frac{x_2}{y_2} = \alpha \left(\frac{1}{y_2/x_2}\right)$$

$$\Rightarrow x_1 y_2 + y_2 x_2 \geq \alpha\, y_2/x_2 + \alpha \frac{1}{y_2/x_2}$$

$$\geq \alpha \left(\frac{y_2}{x_2} + \frac{1}{y_2/x_2}\right)$$

$$\geq 2\alpha$$ (same demonstration as question 2 exercice 1)

$$\Rightarrow \theta(1-\theta)\left[x_1 y_2 + y_2 x_1\right] \geq 2\alpha\,\theta(1-\theta)$$

$$\Rightarrow \theta x + (1-\theta) y \geq \underbrace{\alpha\theta^2 + \alpha(1-\theta)^2 + 2\alpha\,\theta(1-\theta)}_{=\alpha}$$

$\Rightarrow \theta x + (1-\theta) y \geq \alpha \Rightarrow \theta x + (1-\theta) y \in S_\alpha \ \forall \alpha$

$\Rightarrow$ we can conclude that S_α is a convex set and thus f is a quasiconcave function.

Conclusion: f is not convex nor concave
f is quasiconcave but not quasiconvex.

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

- dom $f = \mathbb{R}^2_{++}$ is a convex set.
- f is twice differentiable on dom f:

$$\begin{cases} \dfrac{\partial f}{\partial x_1} = -\dfrac{1}{x_1^2} \times x_2 = -\dfrac{x_2}{x_1^2} \\[2ex] \dfrac{\partial f}{\partial x_2} = -\dfrac{x_1}{x_2^2} \end{cases}$$

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \dfrac{2}{x_1^3 x_2} & \dfrac{1}{x_1^2 x_2^2} \\[3ex] \dfrac{1}{x_1^2 x_2^2} & \dfrac{2}{x_1 x_2^3} \end{pmatrix}$$

$$\begin{cases} \det(\nabla^2 f) = \dfrac{4}{x_1^4 \, x_2^4} - \dfrac{1}{x_1^4 \, x_2^4} = \dfrac{3}{x_1^4 \, x_2^4} > 0 \\ \mathrm{Trace}(\nabla^2 f) = \dfrac{2}{x_1 x_2}\left(\dfrac{1}{x_1^2} + \dfrac{1}{x_2^2}\right) > 0 \end{cases}$$

From these two results, we can notice that the eigenvalues of the hessian matrix have the same sign as the determinant is positive and positive because the trace is positive.
So, the hessian matrix is a positive define matrix.
Conclusion: f is a convex function on $\mathbb{R}^2_{++}$.

- Now, let's take $S_\alpha = \{ x \mid f(x) \leq \alpha \}$.

$$= \left\{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} \mid \frac{1}{x_1 x_2} \leq \alpha \right\}$$

for $\alpha = 0 \Rightarrow S_0 = \left\{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} \mid \frac{1}{x_1 x_2} \leq 0 \right\}$

$= \phi$ which is a convex set

for $\alpha < 0 \Rightarrow S_\alpha = \left\{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} \mid \frac{1}{x_1 x_2} \leq \alpha < 0 \right\}$

$= \phi$ which is a convex set.

for $\alpha > 0 \Rightarrow S_\alpha = \left\{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} \mid \frac{1}{x_1 x_2} \leq \alpha \right\}$

$$= \left\{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} \mid x_1 x_2 \geq \frac{1}{\alpha} \right\}$$

from the previous question, $\{ x_1 \in \mathbb{R}_{++},\ x_2 \in \mathbb{R}_{++} \mid x_1 x_2 \geq \alpha' \}$ is a convex set for all α' in particular $\alpha' = \frac{1}{\alpha}$.

$\Rightarrow$ S_α is a convex set for all $\alpha > 0$.

Conclusion: S_α is a convex set for all $\alpha \Rightarrow f$ is quasiconvex.

Conclusion: f is a convex function and a quasi convex function.

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

- dom f is a convex set.
- f is twice differentiable:

$$\begin{cases} \dfrac{\partial f}{\partial x_1} = \dfrac{1}{x_2} \\ \dfrac{\partial f}{\partial x_2} = -\dfrac{x_1}{x_2^2} \end{cases}$$

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

$$\begin{cases} \det \nabla^2 f(x_1, x_2) = 0 - 1/x_2^4 = -\frac{1}{x_2^4} < 0 \\ \operatorname{trace} \nabla^2 f(x_1, x_2) = \frac{2x_1}{x_2^3} > 0 \end{cases}$$

from these two results, we found that, the eigenvalues of the hessian matrix have different sign (i.e one is positive and the other is negative). Thus $\nabla^2 f \not\succeq 0$ and $\nabla^2 f \not\preceq 0$.
So, f is not a convex function nor a concave function.

- Let's have $S_\alpha = \{x \in \mathbb{R}^2_{++} \mid f(x) \leq \alpha\}$

$$= \{x_1 \in \mathbb{R}_{++}, x_2 \in \mathbb{R}_{++} \mid x_1/x_2 \leq \alpha\}$$
$$= \{x_1 \in \mathbb{R}_{++}, x_2 \in \mathbb{R}_{++} \mid x_1 \leq \alpha x_2\}$$

Now, let's take $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in S_\alpha \Rightarrow \begin{cases} x_1 \leq \alpha x_2 \\ y_1 \leq \alpha y_2 \end{cases}$

$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in S_\alpha \Rightarrow$

and $\theta \in [0,1] \Rightarrow \theta x + (1-\theta) y = \begin{pmatrix} \theta x_1 + (1-\theta) y_1 \\ \theta x_2 + (1-\theta) y_2 \end{pmatrix}$

we have $\begin{cases} \theta x_1 \leq \theta \alpha x_2 \\ (1-\theta) y_1 \leq (1-\theta) \alpha y_2 \end{cases}$

$$\Rightarrow \theta x_1 + (1-\theta) y_1 \leq \alpha [\theta x_2 + (1-\theta) y_2]$$

$$\Rightarrow \frac{\theta x_1 + (1-\theta) y_1}{\theta x_2 + (1-\theta) y_2} \leq \alpha$$

So, S_α is a convex set.

Now, let's take $S'_\alpha = \{x \in \mathbb{R}^2_{++} \mid f(x) \geq \alpha\}$

$$= \{x_1 \in \mathbb{R}^2_{++}, x_2 \in \mathbb{R}^2_{++} \mid \frac{x_1}{x_2} \geq \alpha\}$$
$$= \{x_1 \in \mathbb{R}^2_{++}, x_2 \in \mathbb{R}^2_{++} \mid x_1 \geq \alpha x_2\}$$

Let's take $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in S'_\alpha \Rightarrow x_1 \geqslant \alpha x_2$

$y \in S'_\alpha$; $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow y_1 \geq \alpha x_2$

$\theta \in [0,1] \Rightarrow \theta x + (1-\theta) y = \begin{pmatrix} \theta x_1 + (1-\theta) y_1 \\ \theta x_2 + (1-\theta) y_2 \end{pmatrix}$

$$\Rightarrow \begin{cases} \theta x_1 \geqslant \alpha \theta x_2 \\ (1-\theta) y_1 \geqslant \alpha (1-\theta) y_2 \end{cases}$$

$$\Rightarrow \theta x_1 + (1-\theta) y_1 \geqslant \alpha (\theta x_2 + (1-\theta) y_2)$$

$$\Rightarrow \frac{\theta x_1 + (1-\theta) y_1}{\theta x_2 + (1-\theta) y_2} \geqslant \alpha$$

So, S'_α is a convex set.

$\Rightarrow$ f is a quasi convex and quasiconcave function, so, it is a quasi linear function.

Conclusion, f is not a convex function nor a concave function.

f is a quasi linear function (quasiconvex and quasiconcave).

4) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$ where $0 \leqslant \alpha \leqslant 1$ on $\mathbb{R}^2_{++}$.

- $\text{dom} f = \mathbb{R}^2_{++}$ is a convex set.
- f is a twice differentiable function.

$$\begin{cases} \dfrac{\partial f}{\partial x_1} = \alpha x_1^{\alpha-1} x_2^{1-\alpha} \\ \dfrac{\partial f}{\partial x_2} = (1-\alpha) x_1^{\alpha} x_2^{-\alpha} \end{cases}$$

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

$\det \nabla^2 f = \alpha^2(\alpha-1)^2 x_1^{2\alpha-2} x_2^{-2\alpha} - \alpha^2(\alpha-1)^2 x_1^{2\alpha-2} x_2^{-2\alpha} = 0$

$\text{trace } \nabla^2 f = \alpha \underbrace{(\alpha-1)}_{\leqslant 0} \left[x_1^{\alpha} x_2^{-\alpha-1} + x_1^{\alpha-2} x_2^{1-\alpha} \right] \leqslant 0$

The two eigenvalues are : $\{0, \lambda \leqslant 0\}$ because $\det = 0$
and $tr \leqslant 0$

So, f is a concave function.

- To show that f is a quasiconcave function, we need to demonstrate the following result:

f is concave $\Rightarrow$ f is quasiconcave.

Let's take $S_\alpha = \{ x \in \mathbb{R}^n \mid f(x) \geqslant \alpha \}$

$\left.\begin{array}{l} x \in S_\alpha \\ y \in S_\alpha \\ \theta \in [0,1] \end{array}\right\}$ this show that $f(x) \geqslant \alpha$ and $f(y) \geqslant \alpha$

f is concave so :

$\theta f(x) + (1-\theta) f(y) \leqslant f(\theta x + (1-\theta) y)$

or $\begin{cases} f(x) \geqslant \alpha \Rightarrow \theta f(x) \geqslant \theta\alpha \\ f(y) \geqslant \alpha \Rightarrow (1-\theta) f(y) \geqslant (1-\theta)\alpha \end{cases}$

$\Rightarrow f(\theta x + (1-\theta) y) \geqslant \cancel{\theta}\alpha + (1-\cancel{\theta})\alpha = \alpha$

$\Rightarrow \theta x + (1-\theta) y \in S_\alpha$

So, S_α is a convex set and thus f is concave

$\Rightarrow$ f is quasiconcave.

Conclusion : $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$ is a concave function and quasiconcave function.

Exercice n3 : Convexity of functions:

1) $f(X) = Tr(X^{-1})$ on $dom\, f = S_n^{++}$

- we have $dom\, f = S_n^{++}$ is a convex set.
- f is a twice differentiable. To show the convexity of f, let's take $S(t) = A + tB$ where A is a symmetric positive define and B is a symmetric matrix.

$\Rightarrow$ to show the convexity of f, it is enough to show that:

$$\frac{d^2}{dt^2} Tr\left(S(t)^{-1}\right)\Big|_{t=0} \geq 0.$$

Now, $S(t) = A + tB$.

$$S(t)^{-1} = (A + tB)^{-1} = \left[A\left(I + tA^{-1}B\right)\right]^{-1}$$

$$= \left(I + tA^{-1}B\right)^{-1} A^{-1}$$

We know that : $(A+B)^{-1} = A^{-1} - (I + A^{-1}B)^{-1} A^{-1}BA^{-1}$

$$\overset{\text{(reccursivity)}}{=} A^{-1} - \left[I - (I + A^{-1}B)^{-1}A^{-1}B\right] A^{-1}BA^{-1}$$

$$= A^{-1} - \left[I - \left[I - (I + A^{-1}B)^{-1}A^{-1}B\right]A^{-1}B\right]A^{-1}BA^{-1}$$

$$= A^{-1} - A^{-1}BA + A^{-1}BA^{-1}BA^{-1} - (I + A^{-1}B)^{-1}A^{-1}BA^{-1}BA^{-1}BA^{-1}$$

Now let's take $\boxed{A = I}$ and $\boxed{B = tA^{-1}B}$

$$S(t)^{-1} = \left(I + tA^{-1}B\right)^{-1} A^{-1}$$

$$= \left[I - tA^{-1}B + t^2 A^{-1}BA^{-1}B + \cdots\right] A^{-1}$$

$$= A^{-1} - tA^{-1}BA^{-1} + t^2 A^{-1}BA^{-1}BA^{-1} + \cdots$$

$$\Rightarrow Tr\left(S(t)^{-1}\right) = Tr\left(A^{-1} - tA^{-1}BA^{-1} + t^2A^{-1}BA^{-1}BA^{-1} + \cdots\right)$$

The previous function is twice differentiable with respect to the variable t.

$$Tr\left(S(t)^{-1}\right) = Tr\left(A^{-1}\right) - t\,Tr\left(A^{-1}BA^{-1}\right) + t^2 tr\left(A^{-1}BA^{-1}BA^{-1}\right) + \cdots$$

$$\frac{d\,Tr\left(S(t)^{-1}\right)}{dt} = -Tr\left(A^{-1}BA^{-1}\right) + 2t\,Tr\left(A^{-1}BA^{-1}BA^{-1}\right) + \cdots$$

$$\frac{d^2 Tr(S(t)^{-1})}{dt^2} = 2\, Tr\left(A^{-1}BA^{-1}BA^{-1}\right) + \cdots$$

$$\left.\frac{d^2 Tr\left(S(t)^{-1}\right)}{d^2t}\right|_{t=0} = 2Tr\left(A^{-1}BA^{-1}BA^{-1}\right)$$

- A is a symmetric positive define matrix $\Rightarrow (A^{-1})^T = A^{-1}$
- B is a symmetrix matrix $\Rightarrow B^T = B$

Let's denote $C = A^{-1}B \Rightarrow C^T = B^T(A^{-1})^T = BA^{-1}$

$\Rightarrow Tr\left(A^{-1}BA^{-1}BA^{-1}\right) = Tr\left(CA^{-1}C^T\right)$

Let's take x a vector $\Rightarrow x^T CA^{-1}C^T x = x^T C A^{-1/2}A^{-1/2}C^T x$

$= \left(A^{-1/2}C^Tx\right)^T\left(A^{-\frac{1}{2}}C^Tx\right)$

$= \| A^{-1/2}C^Tx\|^2 \geq 0$

We found that $CA^{-1}C$ is a positive semidefine matrix.

and therefore $Tr\left(CA^{-1}C\right) \geq 0$

$\Rightarrow Tr\left(A^{-1}BA^{-1}BA^{-1}\right) \geq 0$

$\Rightarrow \frac{d^2 Tr\left(S(t)^{-1}\right)}{d^2t} \geq 0$

Conclusion: we found that $\left.\frac{d^2 Tr\left(S(t)^{-1}\right)}{d^2t}\right|_{t=0} \geq 0$

so $f(X) = Tr\left(X^{-1}\right)$ is a convex function.

2) $f(X,y) = y^T X^{-1} y$ on $\mathrm{dom}\, f = S_n^{++} \times \mathbb{R}^n$.

To show that f is convex, we need to show that its epigraph is convex.

- we have $\mathrm{dom}\, f = S_n^{++} \times \mathbb{R}^n$ which is a convex set.
- $\mathrm{epi}\, f = \left\{(y, X, t) \mid X \succ 0 \quad y^T X^{-1} y \leq t\right\}$

using schur complement condition

$= \left\{ (y, X, t) \mid \begin{bmatrix} X & y \\ y^T & t \end{bmatrix} \text{ is positive semi define and } X \text{ is positive define} \right\}$

We note $G(X, y, t) = \begin{bmatrix} X & y \\ y^T & t \end{bmatrix}$

$\mathrm{epi}\, f = G^{-1}\left(S_+^{n+1}\right)$

We know that the inverse image of the positive semidefine

Cone S_+^{n+1} under affine mapping $G(X, y, t)$ is convex.
Thus, epi f is a convex set and then f is a convex function.

Second Method:

According to the lecture note:

$$\frac{1}{2} y^T Q^{-1} y = \sup_x \left(y^T x - \frac{1}{2} x^T Q x \right)$$

$$\Rightarrow \quad y^T X^{-1} y = \sup_x \left(2y^T x - x^T Q x \right)$$

We note $g(X, y) = 2y^T x - x^T Q x$

It is clear that the function g is a linear fuction for both variables X and y.

$\Rightarrow$ g is a convex function in (X, y)

We know that the superemum of a set of convex functions is convex.

$$\Rightarrow \quad f(X, y) = y^T X^{-1} y = \sup_x \left(2y^T x - x^T X x \right) \text{ is convex.}$$

3) $f(X) = \sum_{i=1}^{n} \sigma_i(X)$ on $\text{dom} f = S^n$; $\sigma_1(X), \dots, \sigma_n(X)$ are the singular value of X.

- $\text{dom} f = S^n$ which is a convex set.
- to show that f is a convex, we need to show it is equal to a norm.

To show that f is a norm, we need to verify 4 properties:

* $\|A\| = 0 \Leftrightarrow A = 0$
* $\|tA\| = |t| \, \|A\|$
* $\|A\| \geq 0$
* $\|A + B\| \leq \|A\| + \|B\|$

Let's start with the first one:

$$\begin{cases} \|A\| = 0 \Rightarrow \sum_{i=1}^{n} \underbrace{\sigma_i(A)}_{\geq 0} = 0 \Rightarrow \forall i \ \sigma_i(A) = 0 \Rightarrow A = 0 \\ A = 0 \Rightarrow \sigma_i(A) = 0 \Rightarrow \|A\| = 0 \end{cases}$$

$$\Rightarrow \boxed{\|A\| = 0 \Leftrightarrow A = 0}$$

$$\|tA\| = \sum_{i=1}^{n} \sigma_i(tA) \overset{A \in S^n}{=} \sum_{i=1}^{n} \sqrt{|t|^2 \lambda_i^2}$$

$$= |t| \sum_{i=1}^{n} \sqrt{|\lambda_i|^2}$$

$$= |t| \sum_i \sigma_i(A) = |t| \, \|A\|$$

$$\|A\| = \sum_{i=1}^{n} \underbrace{\sigma_i(A)}_{\geq 0} \geq 0$$

* to show that $\|A + B\| \leq \|A\| + \|B\|$, we need to show that $\|A\| = \sup_{\sigma_1(A) \leq 1} \langle Q, A \rangle = \sup_{\sigma_1(Q) \leq 1} \text{Tr}(Q^T A)$

where $\sigma_1(Q)$ is the maximal singular value of Q which is a norm.

* Let's decompose A according to singular values decomposition:

$$A = U \Sigma V^T \quad ; \quad Q = U V^T$$

$$\Rightarrow \langle Q, A \rangle = \langle U V^T, U \Sigma V^T \rangle$$

$$= \text{Tr}(V U^T U \Sigma V^T)$$

$$= \text{Tr}(V V^T U^T U \Sigma) = \text{Tr}(\Sigma) = \sum_{i=1}^{n} \sigma_i(A)$$

for all Q

$$\Rightarrow \sup_{\sigma_1(Q) \leq 1} \langle Q, A \rangle \geq \sum_{i=1}^{n} \sigma_i(A)$$

* $\sup_{\sigma_1(Q) \leq 1} \langle Q, A \rangle = \sup_{\sigma_1(Q) \leq 1} \text{Tr}(Q^T U \Sigma V^T)$

$$\sup_{\sigma_1(Q)\leq 1} \langle Q, A\rangle = \sup_{\sigma_1(Q)\leq 1} Tr\left(V^T Q^T U \Sigma\right) \quad \left(\text{because } Tr(ABC) = Tr(CAB)\right)$$

$$= \sup_{\sigma_1(Q)\leq 1} \langle U^T Q V, \Sigma\rangle$$

$$= \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{m} \sigma_i \left(U^T Q V\right)_{ii}$$

$$= \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{n} \sigma_i \, u_i^T Q v_i$$

$$\leq \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{n} \sigma_i \, \sigma_{max}(Q)$$

$$\leq \sup_{\sigma_1(Q)\leq 1} \sum_{i=1}^{m} \sigma_i = \sum_{i=1}^{m} \sigma_i$$

We have proven both the $\leq$ and $\geq$ cases, so we have the equality: $\displaystyle \sup_{\sigma_1(Q)\leq 1} \langle Q, A\rangle = \|A\| = \sum_{i=1}^{m} \sigma_i(A).$

Using the triangle inequality, we have:

$$\|A+B\| = \sup_{\sigma_1(Q)\leq 1} \langle Q, A+B\rangle$$

$$\leq \sup_{\sigma_1(Q)\leq 1} \langle Q, A\rangle + \sup_{\sigma_1(Q)\leq 1} \langle Q, B\rangle$$

$$= \|A\| + \|B\|$$

Thus $f(X) = \sum_{i=1}^{m} \sigma_i(X)$ is a norm

So, f is a convex function.

Exercice n 4 :

Monotone non negative cone:

$$K_{m+} = \{ x \in \mathbb{R}^n \mid x_1 \geq x_2 \cdots \geq x_n \geq 0 \}$$

1) To show that K_{m+} is a proper cone, we need to verify:

- K_{m+} is closed
- K_{m+} has a solid interior
- K_{m+} is pointed : $\left.\begin{matrix} x \in C \\ -x \in C \end{matrix}\right\} x = 0$

1 - K_{m+} is closed:

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

We have $\left.\begin{matrix} \{ x_i \geq x_{i+1} \} \ \forall i \\ \text{and} \ \{ x_n \geq 0 \} \end{matrix}\right\}$ are halfspaces

$\Rightarrow$ K_{m+} is the intersection of halfspaces $\Rightarrow$ K_{m+} is closed.

2 - K_{m+} has a solid interior:

the point $x = (m, m-1, \ldots, 1) \in K_{m+}$

$K_{m+} \neq \phi \Rightarrow K_{m+}$ has a solid interior

3 - K_{m+} is pointed:

$x \in K_{m+} \Rightarrow x_1 \geq x_2 \cdots \geq x_n \geq 0 \Rightarrow x_i \geq 0 \ \forall i$

$-x \in K_{m+} \Rightarrow -x_1 \geq -x_2 \cdots \geq -x_n \geq 0 \Rightarrow -x_i \geq 0 \ \forall i$

$\Rightarrow x_i = 0 \ \forall i \Rightarrow x = 0$

$\Rightarrow K_{m+}$ is pointed

Conclusion : K_{m+} is proper cone.

2) Dual cone of K_{m+} :

$$K_{m+}^* = \{ y \mid y^T x \geq 0 \ \forall x \in K_{m+} \} = \{ y \mid \sum_{i=1}^{n} y_i x_i \geq 0 \ \forall x \in K_{m+} \}$$

We have $y \in K_{m+}^* \Rightarrow$ for a given $z_k = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ } k element

it is clear that $z_k \in K_{m+}$; we get:

$$y^T z_k = \sum_{i=1}^{k} y_i \geq 0 \Rightarrow \sum_{i=1}^{k} y_i \geq 0 \text{ for all } k.$$

Reciprocally, we have $\sum_{i=1}^{k} y_i \geqslant 0 \quad \forall\, k \in 1, \dots m$

Let's show that $\boxed{y^T x \geqslant 0}$ with x in K_{m+}.

To obtain this result, we need to show the following equality:

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

With recurrence:

for $n = 1 \Rightarrow x_1 y_1 = \underbrace{\sum_{k=1}^{0} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i}_{=0} + x_1 y_1$

$= x_1 y_1$

We suppose it is true until n, we must show it for $n+1$:

$$\begin{aligned}
\sum_{i=1}^{n+1} x_i y_i &= \sum_{i=1}^{n} x_i y_i + x_{n+1} y_{n+1} \\
&= \sum_{k=1}^{n-1} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i + x_n \sum_{i=1}^{n} y_i + x_{n+1} y_{n+1} \\
&= \sum_{k=1}^{n-1} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i + x_n \sum_{i=1}^{n} y_i + x_{n+1} \left(- \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} y_i \right) + x_{n+1} y_{n+1} \\
&= \sum_{k=1}^{n-1} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i + (x_n - x_{n+1}) \sum_{i=1}^{n} y_i + x_{n+1} \sum_{i=1}^{n+1} y_i \\
&= \sum_{k=1}^{n} (x_k - x_{k+1}) \sum_{i=1}^{k} y_i + x_{n+1} \sum_{i=1}^{n+1} y_i
\end{aligned}$$

$\Rightarrow$ we have now $\sum_{i=1}^{n} x_i y_i = y^T x$

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

So, $y^T x \geqslant 0 \Rightarrow y \in K_{m+}^{*}$

Conclusion: $K_{m+}^{*} = \left\{ y \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} y_i x_i \geqslant 0 \quad \forall\, x \in K_{m+} \right\}$

$$\boxed{K_{m+}^{*} = \left\{ \sum_{i=1}^{k} y_i \geqslant 0 \quad \text{for } k = 1, \dots, m \right\}}$$

Exercice n°5 : Conjugate of functions.

1) Max function $f(x) = \max_{1\leq i\leq n} x_i$ on $\mathbb{R}^n$.

$$f^*(y) = \max_x \; y^T x - f(x)$$

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

– Let's take y_k a negative component of y.

if $x_k = -t$ and $x_i = 0 \; \forall i \neq k$

$\Rightarrow \; y^T x - \max x_i = -t y_k \xrightarrow[t\to\infty]{} +\infty$

– Let's take $y \geq 0$: $y^T x - \max_{1\leq i\leq n} x_i = \sum_{i=1}^n x_i y_i - \max_{1\leq i\leq n} x_i$

$\leq \max_{1\leq i\leq n} x_i \left(\sum_{i=1}^n y_i - 1\right)$

* if $\sum_{i=1}^n y_i > 1$, we take x : a constant vector.

$$x = \begin{pmatrix} t \\ \vdots \\ t \end{pmatrix}$$

$$\Rightarrow \; y^T x - \max x_i = \sum_{i=1}^n t y_i - t$$

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

* if $\sum_{i=1}^n y_i < 1$, we take the same x

$$y^T x - \max x_i = \sum_{i=1}^n t y_i - t$$

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

* if $\sum_{i=1}^n y_i = 1$

$$y^T x - \max_{1\leq i\leq n} x_i \leq \max_{1\leq i\leq n} x_i \times \left(\underbrace{\sum_i y_i - 1}_{=0}\right)$$

$$y^T x - \max_{1\leq i\leq n} x_i \leq 0$$

$$\Rightarrow \; f^*(y) = \max_x \left(y^T x - \max_{1\leq i\leq n} x_i \right) = 0$$

$$\Rightarrow \; \boxed{f^*(y) = \begin{cases} 0 & \text{if } y \geq 0 \;\; \sum_{i=1}^n y_i = 1 \\ +\infty & \text{otherwise} \end{cases}}$$

2) Sum of largest elements : $f(x) = \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$.

- Suppose y has a negative component : ... $y_k < 0$

take $x_k = -t$ and $x_i = 0 \quad \forall k \neq i$

$y^T x - f(x) = -t y_k \xrightarrow[t \to \infty]{} +\infty$

- Now, we take $y \geq 0$, if $y_k > 1$; x verify $\begin{cases} x_k = t \\ x_i = 0 & \forall i \neq k \end{cases}$

$y^T x - f(x) = t y_k - t = t(\underbrace{y_k - 1}_{>0}) \xrightarrow[t \to \infty]{} +\infty$.

- Thus $0 \leq y \leq 1$;

$y^T x - f(x) = \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]}$

* if $\sum_{i=1}^{n} y_i \neq r \Rightarrow$ take x a constant vector $x = \begin{pmatrix} t \\ \vdots \\ t \end{pmatrix}$

$\Rightarrow y^T x - f(x) = \sum_{i=1}^{n} t y_i - r t$

$= t \left(\sum_{i=1}^{n} y_i - r \right) \xrightarrow[t \to \infty]{} +\infty$

* if $\sum_{i=1}^{n} y_i = r \Rightarrow y^T x - f(x)$

here we take $(x_1 \geq \dots \geq x_n)$

$= \sum_{i} x_i y_i - \sum_{i=1}^{r} x_{[i]}$

$= \sum_{i=r+1}^{n} x_i y_i + \sum_{i=1}^{r} (y_i - 1) x_{[i]}$

$\leq x_r \left(\sum_{i=r+1}^{n} y_i + \sum_{i=1}^{r} (y_i - 1) \right)$

$\leq x_r \left(\underbrace{\sum_{i=1}^{n} y_i - r}_{=0} \right)$

$\Rightarrow f^*(y) = 0$

$$\Rightarrow \boxed{f^*(y) = \begin{cases} 0 & \text{if } 0 \leq y \leq 1 \; ; \; \sum_{i=1}^{n} y_i = r \\ +\infty & \text{otherwise} \end{cases}}$$

3) Piece wise - linear function on $\mathbb{R}$: $f(x) = \max_{1 \leq i \leq m} (a_i x + b_i)$

$\begin{cases} * \ a_i \text{ are sorted in increasing order.} \\ (a_1 \leq \dots \leq a_m) \\ * \ a_i x + b_i \text{ are not redundant: for each } k \text{; there is at} \\ \text{least one } x \text{ with } f(x) = a_k x + b_k. \end{cases}$

$\Rightarrow$ f is a piece wise linear with breack point $x_j = \dfrac{b_i - b_{i+1}}{a_{i+1} - a_i}$.

- If $y \in [a_1, a_m]$ $\exists j \ / \ a_j \leq y \leq a_{j+1}$

then $\Rightarrow \begin{cases} y - a_1, \dots y - a_j \geq 0 \\ y - a_{j+1}, \dots y - a_m \leq 0 \end{cases}$

This may indicate that $\{y x - \max_i a_i x + b_i\}$ achieves its maximum between the two point $[x_j ; x_{j+1}]$ and exactely at the point x_j.

$$f^*(y) = y x_j - a_j x_j - b_j$$
$$= y \frac{b_j - b_{j+1}}{a_{j+1} - a_j} - a_j \frac{b_j - b_{j+1}}{a_{j+1} - a_j} - b_j$$
$$= (y - a_j)\left(\frac{b_j - b_{j+1}}{a_{j+1} - a_j}\right) - b_j$$

- If $y < a_1 \Rightarrow \forall i \quad y - a_i < 0$

then $f^*(y) = \max_x \left(y x - \max_i (a_i x + b_i) \right)$ has negative slopes and then it tends to $+\infty$ when x tends to $-\infty$.

- If $y > a_m \Rightarrow \forall i \quad y - a_i > 0$

then $f^*(y) = \max_x \left(y x - \max_i (a_i x + b_i) \right)$

has positive slopes and then it tends to $+\infty$ when x tends to $+\infty$

Thus : $f^*(y) = \begin{cases} (y - a_j)\left(\dfrac{b_j - b_{j+1}}{a_{j+1} - a_j}\right) - b_j & \text{if } y \in [a_j ; a_{j+1}] \\ +\infty & \text{other wise} \end{cases}$

$$f^*(y) = \begin{cases} (y - a_j)\dfrac{b_j - b_{j+1}}{a_{j+1} - a_j} - b_j & \begin{array}{l} a_1 \leq y \leq a_m \\ a_j \leq y \leq a_{j+1} \end{array} \\ +\infty & \text{otherwise} \end{cases}$$