Convex Optimization - HW1

Théotime de Charrin

Monday 16e October, 2023

1 Exercise 1 : Which of the following sets are convex?

1) The rectangle set defined as $\left\{x \in \mathbb{R}^n \mid \forall i \in [\![1, n]\!], \alpha_i \leq x_i \leq \beta_i\right\}$
For a given i, a $\{x_i\}$ is the intersection of two halfspaces $\left\{x \in \mathbb{R}^n | x_i \leq \beta_i\right\}$ and $\left\{x \in \mathbb{R}^n | x_i \geq \alpha_i\right\}$
As i is finite, a rectangle is a finite intersection of halfspaces so it is a convex set.

2) The hyperbolic set defined as $\left\{H : x \in R_+^2 | x_1 x_2 \geq 1\right\}$.
We take $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in R^2$ such that $x, y \in H$ and $\theta \in [0;1]$ Hence, we have $x_1 x_2 \geq 1$ and the same for y.
Let's look at the convex combination of x and y :

$$\begin{aligned} \theta x + (1-\theta) y &= \begin{pmatrix} \theta x_1 + (1-\theta) y_1 \\ \theta x_2 + (1-\theta) y_2 \end{pmatrix} \\ &= \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \end{aligned}$$

$$z_1 z_2 = \theta^2 \overbrace{x_1 x_2}^{\geq 1} + (1-\theta)^2 \overbrace{y_1 y_2}^{\geq 1} + \theta(1-\theta)(x_1 y_2 + y_1 x_2)$$

Let's have a look at $x_1 y_2 + y_1 x_2$: we have $x_1 x_2 \geq 1 \Leftrightarrow x_1 y_2 \geq \frac{y_2}{x_2}$. In the same manner, $y_1 x_2 \geq \frac{x_2}{y_2}$.
We now have :

$$\begin{aligned} x_1 y_2 + y_1 x_2 &\geq \frac{y_2}{x_2} + \frac{x_2}{y_2} \\ &\geq \frac{y_2}{x_2} + \frac{1}{\frac{y_2}{x_2}} \\ \Leftrightarrow x_1 y_2 + y_1 x_2 - 2 &\geq \frac{y_2}{x_2} + \frac{1}{\frac{y_2}{x_2}} - 2 \\ &\geq \sqrt{\frac{y_2}{x_2}}^2 + \frac{1}{\sqrt{\frac{y_2}{x_2}}}^2 - 2 \frac{\sqrt{\frac{y_2}{x_2}}}{\sqrt{\frac{y_2}{x_2}}} \\ &\geq \left(\sqrt{\frac{y_2}{x_2}} - \frac{1}{\sqrt{\frac{y_2}{x_2}}} \right)^2 \geq 0 \ \forall \left\{(x_1, x_2), (y_1, y_2)\right\} \end{aligned}$$

We can conclude that $x_1 y_2 + y_1 x_2 \geq 2$
$\theta(1-\theta) \geq 0 \ \forall \theta \in [0,1]$ so $\theta(1-\theta)(x_1 y_2 + y_1 x_2) \geq 2\theta(1-\theta)$ and :

$$\begin{aligned} z_1 z_2 &\geq \theta^2 + (1-\theta)^2 + 2\theta(1-\theta) \\ &\geq (\theta + 1 - \theta)^2 \geq 1 \end{aligned}$$

We have shown that $z_1 z_2 \geq 1$, i.e. $\theta x + (1-\theta) y \in H$
A convex combination of two elements in H is also in H, so the hyperbolic set is convex as well.

3) The set of points closer to a given point than a given set, *i.e.*

$$A = \left\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \ \forall y \in S\right\} \text{where} S \subseteq \mathbb{R}^n.$$

Let $y \in S$ and $x \in A$. Then we must have :

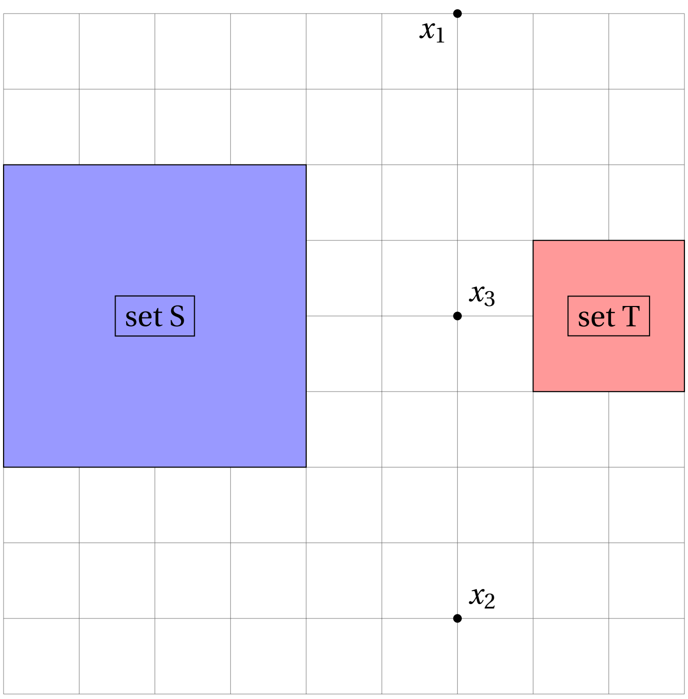
$$\begin{aligned} \|x - x_0\|_2 &\leq \|x - y\|_2 \\ (x - x_0)^T (x - x_0) &\leq (x - y)^T (x - y) \\ (x^T - x_0{}^T)(x - x_0) &\leq (x^T - y^T)(x - y) \\ x^T x - x^T x_0 - x_0{}^T x + x_0{}^T x_0 &\leq x^T x - x^T y - y^T x + y^T y \\ -2x^T x_0 + \|x_0\|_2^2 &\leq -2x^T y + \|y\|_2^2 \\ 2x^T (y - x_0) &\leq \|x_0\|_2^2 + \|y\|_2^2 \\ (y - x_0)^T x &\leq \frac{\|x_0\|_2^2 + \|y\|_2^2}{2} \end{aligned}$$

This is the equation of an halfspace ($a^T x \leq b$), which is a convex set. S is the intersection over $y \in S$ of these halfspaces, hence it's also a convex set.

4) The set of points closer to one set than another, *i.e.*

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

A visual example should better explain why this is not always a convex set :



Here, we see that x_1 and x_2 are closer to the set S rather than the set T. Therefore, they would be part of the set A. However, x_3, a convex combination of $x_1,\ x_2$, *i.e.* on the same segment, is closer to the set T than the set S.
Therefore, A is not a convex set.

5) The set

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

where $S_1,\ S_2 \ \in \mathbb{R}^n$, S_1 convex.

$$x + S_2 \subseteq S_1 \Leftrightarrow \forall y \ \in S_2,\ x + y \in S_1$$

Let's look at the convex combination of $u,\ v \in A,\ \theta \in [0,1]$, and see if it's still in A :

$$\theta u + (1-\theta)v + y = \theta(u+y) + (1-\theta)(v+y) \in S_1$$

By definition, $u + y,\ v + y \in S_1$ because $u,\ v \in A$ and $y \in S_2$.
As S_1 is convex, any convex combination of $z \in S_1$ is also in S_1. We conclude that A is a convex set.

2 Exercise 2 : For each of the following functions determine whether it is convex or concave or not.

1) $f(x_1,x_2) = x_1x_2$ on $\mathbb{R}^2_{++}$
- dom $f = \mathbb{R}^2_{++}$ which is convex.
- f is twice differentiable, let's look at the second-order conditions :

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

Hence, we have the following hessian : $\Delta^2 f(x_1,x_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

With -1 and 1 as eigen values. $\Delta^2 f \notin S_n^+$, so f is not convex. In the same manner, $\Delta^2 f \not\preceq 0$
We conclude that F is neither convex nor concave

2) $f(x_1,x_2) = \frac{1}{x_1x_2}$ on $\mathbb{R}^2_{++}$
- dom $f = \mathbb{R}^2_{++}$ which is convex.
- f is twice differentiable, let's look at the second-order conditions :

$$\begin{cases} \dfrac{\partial f}{\partial x_1} & = -\dfrac{1}{x_1^2 x_2} \\ \dfrac{\partial f}{\partial x_2} & = -\dfrac{1}{x_2^2 x_1} \end{cases} \Rightarrow \begin{cases} \dfrac{\partial^2 f}{\partial^2 x_1} = \dfrac{2}{x_1^3 x_2} & \dfrac{\partial^2 f}{\partial^2 x_1 x_2} = \dfrac{1}{x_1^2 x_2^2} \\ \dfrac{\partial^2 f}{\partial^2 x_2} = \dfrac{2}{x_2^3 x_1} & \dfrac{\partial^2 f}{\partial^2 x_2 x_1} = \dfrac{1}{x_1^2 x_2^2} \end{cases}$$

Hence, we have the following hessian : $\Delta^2 f(x_1, x_2) = \begin{pmatrix} \dfrac{2}{x_1^3 x_2} & \dfrac{1}{x_1^2 x_2^2} \\ \dfrac{1}{x_1^2 x_2^2} & \dfrac{2}{x_2^3 x_1} \end{pmatrix}$

$\begin{cases} \mathbf{det}(\Delta^2 f) = \dfrac{4}{x_1^4 x_2^4} - \dfrac{1}{x_1^4 x_2^4} > 0 \\ \mathrm{Tr}(\Delta^2 f) = \dfrac{2x_2^2 + 2x_1^2}{x_1^3 x_2^3} > 0 \end{cases} \Rightarrow \lambda_1 \lambda_2 > 0,\ \lambda_1 + \lambda_2 > 0$:All diagonal values are positive, the determinant is positive so all eigenvalues are positive, hence $\Delta^2 f$ is a positive definite matrix, $\Delta^2 f \succ 0$. We can say that f is a convex function on $\mathbb{R}^2_{++}$.

3) $f = \frac{x_1}{x_2}$ on $\mathbb{R}^2_{++}$

— dom $f = \mathbb{R}^2_{++}$ which is convex.

— f is twice differentiable, let's look at the second-order conditions :

$$\begin{cases} \dfrac{\partial f}{\partial x_1} & = \dfrac{1}{x_2} \\ \dfrac{\partial f}{\partial x_2} & = -\dfrac{x_1}{x_2^2} \end{cases} \Rightarrow \begin{cases} \dfrac{\partial^2 f}{\partial^2 x_1} = 0 & \dfrac{\partial^2 f}{\partial^2 x_1 x_2} = -\dfrac{1}{x_2^2} \\ \dfrac{\partial^2 f}{\partial^2 x_2} = -\dfrac{2x_1}{x_2^3} & \dfrac{\partial^2 f}{\partial^2 x_2 x_1} = \dfrac{1}{x_2^2} \end{cases}$$

Hence, we have the following hessian : $\Delta^2 f(x_1, x_2) = \begin{pmatrix} 0 & -\dfrac{1}{x_2^2} \\ -\dfrac{1}{x_2^2} & \dfrac{2x_1}{x_2^3} \end{pmatrix}$

$\begin{cases} \mathbf{det}\ (\Delta^2 f) = \overbrace{-\dfrac{2x_1}{x_2^3}}^{<0,\ \mathrm{dom}\ f = \mathbb{R}^2_{++}} \quad -\dfrac{1}{x_2^2}^{2} < 0 \\ \mathrm{Tr}\ (\Delta^2 f)\ = \dfrac{2x_1}{x_2^3} > 0 \end{cases} \Rightarrow \lambda_1 \lambda_2 < 0,\ \lambda_1 + \lambda_2 > 0$:All diagonal values are positive, the determinant is negative so one eigenvalue is positive an the other negative, hence $\Delta^2 f \nsucceq 0$, $\Delta^2 f \npreceq 0$.

We can conclude that f is neither convex nor concave.

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

— dom $f = \mathbb{R}^2_{++}$ which is convex.

— f is twice differentiable, let's look at the second-order conditions :

$$\begin{cases} \dfrac{\partial f}{\partial x_1} & = \alpha x_1^{\alpha-1} x_2^{1-\alpha} \\ \dfrac{\partial f}{\partial x_2} & = (1-\alpha) x_1^{\alpha} x_2^{-\alpha} \end{cases} \quad \Rightarrow \quad \begin{cases} \dfrac{\partial^2 f}{\partial^2 x_1} = \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} & \dfrac{\partial^2 f}{\partial^2 x_1 x_2} = \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \\ \dfrac{\partial^2 f}{\partial^2 x_2} = \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1} & \dfrac{\partial^2 f}{\partial^2 x_2 x_1} = \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \end{cases}$$

Hence, we have the following hessian : $\Delta^2 f(x_1, x_2) = \begin{pmatrix} \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} & \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \\ \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} & \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1} \end{pmatrix}$

$$\begin{cases} \textbf{det}\,(\Delta^2 f) & = \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} \times \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1} - \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \times \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \\ & = \alpha^2(\alpha-1)^2 \left[x_1^{2\alpha-2} x_2^{-2\alpha} + x_1^{2\alpha-2} x_2^{-2\alpha} \right] \\ & = \alpha^2(\alpha-1)^2 \left(\dfrac{x_1^{\alpha-1}}{x_2^{\alpha}} \right)^2 \geq 0 \\ \mathrm{Tr}\,(\Delta^2 f) & = \underbrace{\alpha(\alpha-1)}_{\leq 0} \left(x_1^{\alpha-2} x_2^{1-\alpha} + x_1^{\alpha} x_2^{-\alpha-1} \right) \leq 0 \end{cases}$$

$\Rightarrow \quad \lambda_1 \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 \leq 0$: All diagonal values are negative, the determinant is positive so all eigenvalues are negatives, hence $\Delta^2 f \preceq 0$.
We can conclude that f is concave.

3 Exercise 3 : Show that following functions are convex

1) $f(X) = \mathbf{Tr}(X^{-1})$ on **dom** $f = S^n_{++}$.
- dom $f = S^2_{++}$ which is convex.
- f is twice differentiable, the second-order conditions might apply :

Let's look at the function $g(t) = f(X + tV)$, $\mathrm{dom}X = \{t \mid X + tV \in \mathrm{dom}f\}$.
Let's pose $X \in S^n_{++}$and$V \in S_n$. If for all X and V in this domain g is convex, then f(X) must be convex as well.

$$\begin{aligned} g(t) &= \mathrm{Tr}((X + tV)^{-1}) \\ \text{Let's look inside the Trace :}\ &= \mathrm{Tr}(X(I + tX^{-1}V))^{-1}) \\ \text{Let's make a Taylor expansion in t :}\ &= \mathrm{Tr}(X^{-1} - tX^{-1}VX^{-1} + t^2 + t^2 X^{-1}VX^{-1}VX^{-1} + \dots) \end{aligned}$$

To show that $\mathrm{Tr}(X^{-1})$ is convex, we can look at the second derivative, when t=0 :

$$\left.\frac{\partial^2 g}{\partial t^2}\right|_{t=0} = 2\mathrm{Tr}(X^{-1}VX^{-1}VX^{-1})$$

We can write this as $PX^{-1}P^T$, with $P = X^{-1}V$, and $X^{-1} \in S^n_{++}$. Hence $PX{-}1P^T$ is positive definite as well, so its trace must be positive.
g is a linear function of y and X, hence a convex function in (X,y)
The supremum of a set of convex functions (every g(X,y), with any x in $\mathbb{R}^n$)is a convex function.
We conclude that f(X,y) is a convex function.

3) $f(X) = \sum_{i=1}^{n} \sigma_i(X)$ on **dom** f $= S^n$, where $\sigma_1(X), \ldots, \sigma_n(X)$ are the singular values of a matrix $X \in \mathbb{R}^{n\times n}$

— dom $f = S^n$ which is convex.

f is also known as the nuclear norm. A norm is always convex, so we just have to show that f is a norm.

We want to verify the definition of norm :

— Homogeneity : $\|\alpha A\| = |\alpha| \times \|A$ for $\alpha \in \mathbb{R}$

— Triangular Inequality : $\|A+B\| \leq \|A\| + \|B\|$

— Positive definiteness : if $\|A\| = 0$, then A=0

— Positivity : $\|A\| \geq 0$

Regarding the positivity :

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

Regarding the positive definiteness :

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

Regarding the homogeneity, for t a scalar :

$$\begin{aligned} \|tA\| = \sum_{i=1}^{n} \sigma_i(tA) \underbrace{=}_{\text{As A is in } S^n} & \sqrt{\|t^2\|\lambda_i}^2 \\ &= |t| \sum_{1}^{n} \sqrt{\lambda_i{}^2} \\ &= |t| \sum_{1}^{n} \sigma_i(A) \\ &= |t| \|A\| \end{aligned}$$

Regarding the triangular inequality :

Using the supremum, we will try to prove that

$$\sup_{\sigma_1(Q)\leq 1} \langle Q, A\rangle = \sup_{\sigma_1(Q)\leq 1} \mathrm{Tr}(Q^T A) = \sum_{1}^{n} \sigma_i(A) = \|A\|$$

Where $\sigma_i(\cdot)$ is itself a norm, and σ_1 is the maximum singular value of Q.

Let $A = U\Sigma V^T = \sum \sigma_i u_i v_i^T$ be the singular value decomposition of A, and pose $Q = UV^T = UIV^T$. Hence, $\sigma_1(Q) = 1$.

$$\begin{aligned} \langle Q, A\rangle &= \langle UV^T, U\Sigma V^T\rangle \\ &= \mathrm{Tr}(VU^T U\Sigma V^T) \\ &= \mathrm{Tr}(V^T V U^T U\Sigma) \qquad = \mathrm{Tr}(\Sigma) \end{aligned}$$

Hence, it comes that

$$\sup_{\sigma_1(Q)\leq 1} \langle Q, A\rangle \geq \sum_1^n \sigma_i(A)$$

Let's prove the other direction and we will have an equality.

$$\begin{aligned}
\sup_{\sigma_1(Q)\leq 1} \langle Q, A\rangle &= \sup_{\sigma_1(Q)\leq 1} \mathrm{Tr}(Q^T A) \\
&= \sup_{\sigma_1(Q)\leq 1} \mathrm{Tr}(Q^T U \Sigma V^T) \\
&= \sup_{\sigma_1(Q)\leq 1} \mathrm{Tr}(V^T Q^T U \Sigma) \\
&= \sup_{\sigma_1(Q)\leq 1} \langle U^T Q V, \Sigma\rangle \\
&= \sup_{\sigma_1\leq 1} \sum_1^n \sigma_i(U^T Q V) \\
&= \sup_{\sigma_1\leq 1} \sum_1^n \sigma_i u_i^T Q v_i \\
&\leq \sup_{\sigma_1\leq 1} \sum_1^n \sigma_i \sigma_{max}(Q) \\
&= \sum_i^n \sigma_i(A) = \|A\|
\end{aligned}$$

We have proven the equality $\sup_{\sigma_1(Q)\leq 1} = \|A\|$. We can now show the triangular inequality :

$$\begin{aligned}
\|A+B\| &= \sup_{V, \sigma_1(V)\leq 1} \langle V, A+B\rangle \\
&\leq \sup_{V,\ \sigma_1(V)\leq 1} \langle V, A\rangle + \langle V, B\rangle = \|A\| + \|B\|.
\end{aligned}$$

Hence, we've shown that the nuclear norm is a norm, so f is convex.