Homework 1

Exercise 1.

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

Let $x, y \in S$. So for every $i \in \{1, \dots, n\}$, $\begin{cases} \alpha_i \leq x_i \leq \beta_i \\ \alpha_i \leq y_i \leq \beta_i \end{cases}$

Let $\theta \in [0,1]$, and show that $\theta x + (1-\theta) y \in S$

~~Let θ~~ Let $i \in \{1, \dots, n\}$.

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

So $\theta x + (1-\theta) y \in S$

By definition, <u>S is convex</u>

2) $S = \{x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1\}$

As $(0,0) \notin S$, we can write equivalently:

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

$$= \left\{(x,y) \in \mathbb{R}^2_{++} \mid y \geq \frac{1}{x}\right\}.$$

Let $f(x) = \frac{1}{x}$ with $x \in \mathbb{R}_{++}$. $S = \{(x,y) \in \mathbb{R}^2_{++} \mid y \geq f(x)\}$

f is convex and $\operatorname{dom} f = \mathbb{R}_{++}$ is a convex set

Let $x = (x_1, y_1), y = (x_2, y_2) \in S$ and $\theta \in [0,1]$.

$$\theta y_1 + (1-\theta) y_2 \underset{\substack{y_1 \geq f(x_1) \\ y_2 \geq f(x_2)}}{\geq} \theta f(x_1) + (1-\theta) f(x_2) \underset{f \text{ is convex}}{\geq} f(\theta x_1 + (1-\theta) x_2)$$

So by definition, S is convex

3. $E = \{x \mid \|x - x_0\| \leq \|x - y\| \text{ for all } y \in S\}$

Let $x \in E$ and $y \in S$. We have:

$$\|x - x_0\|_2^2 \leq \|x - y\|_2^2$$

$$\Leftrightarrow \langle x - x_0, x - x_0\rangle \leq \langle x - y, x - y\rangle$$

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

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

$$\Leftrightarrow 2(y - x_0)^T x \leq \|y\|_2^2 - \|x_0\|_2^2$$

So $E = \{x \mid 2(y - x_0)^T x \leq \|y\|_2^2 - \|x_0\|_2^2 \text{ for all } y \in S\}$.

That set is an half-space for all $y \in S$, which is convex

By an intersection, $E = \bigcap_{y \in S} \{x \mid 2(y - x_0)^T x \leq \|y\|_2^2 - \|x_0\|_2^2\}$

So E is convex by operation that preserves convexity

4) $E = \{x \mid \text{dist}(x, S) \leq \text{dist}(x, T)\}$, $S, T \subseteq \mathbb{R}^n$

Let us show with a counterexample that E is not convex

5) $E = \{x \mid x + S_2 \subseteq S_1\}$ and $x + S_2 = \{x + y \mid y \in S_2\}$

So, if we note $f_{S_2}(x_1) = x + y$ with $y \in S_2$, we have an affine function

And affine functions preserve convexity, we have

$E = \{x \mid f(x) \in S_1\}$ that is convex (inverse image of f_{S_2})

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

Let us compute the hessian $\nabla^2 f$ of because is twice differentiable:

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

This definite positive because for all $x \in \mathbb{R}^2_{++}$, $x^T \nabla^2 f x = 2x_1 x_2 > 0$

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

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

Let us compute the hessian $\nabla^2 f$ of f because f is twice differentiable by composition

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

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = -\frac{x_2}{(x_1 x_2)^2} = -\frac{1}{x_1^2 x_2}$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = -\frac{1}{x_1 x_2^2}$$

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

As every element of the matrix is strictly positive, for all $x \in \mathbb{R}^2_{++}$, $x^T \nabla^2 f x > 0$

~~So~~ Thus, f is strictly convex

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

Let us compute the hessian $\nabla^2 f$ as f is twice differentiable to show that f is neither convex nor concave

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

$\frac{\partial f}{\partial x_1}(x_1, x_2) = \frac{1}{x_2}$

$\frac{\partial f}{\partial x_2}(x_1, x_2) = -\frac{x_1}{x_2^2}$

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

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

$$X^T \nabla^2 f(x_1, x_2) X = -\frac{2xy}{x_2^2} + \frac{2x_1 x_2 y^2}{x_2^3}$$

And the sign of this expression depends on all variable so

The hessian is neither semi definite positive nor negative

Thus f is neither convex nor concave.

Let us show that f is quasi linear, meaning quasi convex and quasi concave.

Let $\alpha \in \mathbb{R}$. $S_\alpha^- = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid \frac{x_1}{x_2} \leq \alpha\}$

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

This defines an half space of $\mathbb{R}^2_{++}$, so S_α^- is convex

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

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

defines an half space of $\mathbb{R}^2_{++}$ so S_α^+ is ~~concave~~ convex.

So f is both quasi convex and quasi concave thus f is quasi linear

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

Let us compute the hessian $\nabla^2 f$ as f is twice differentiable

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

$\frac{\partial f}{\partial x_1} = \alpha x_1^{\alpha-1} x_2^{1-\alpha} \Rightarrow \frac{\partial^2 f}{\partial x_1^2} = \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha}$

$\frac{\partial f}{\partial x_2} = (1-\alpha) x_1^\alpha x_2^{-\alpha} \Rightarrow \frac{\partial f}{\partial x_2^2} = -\alpha(1-\alpha) x_1^\alpha x_2^{-\alpha-1}$

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

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

To show that $\nabla^2 f$ is semi-definite ~~positive~~ negative we will that $\det(\nabla^2 f)$ is ~~positive~~ negative and the leading principal minors is ~~positive~~ negative:

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

and $\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} \leqslant 0$ because $\alpha \leqslant 1$

So $\nabla^2 f$ is semi definite negative meaning that f is concave

Exercise 3. 1) $f(X) = \mathrm{Tr}(X^{-1})$ on $\mathrm{dom} f = S^n_{++}$

To show that this function is convex, we will show that:

$$\forall X, Y \in S^n_{++},\ (X+Y)^{-1} \preceq X^{-1} + Y^{-1}$$

Let $X, Y \in S^n_{++}$ and $g(t) = (X + tY)^{-1}$ on $\mathbb{R}_+$

g is differentiable and $\frac{dg}{dt}(t) = -(X+tY)^{-1}\, Y\, (X+tY)^{-1}$

Since Y and $X+tY$ are definite positives, $\frac{dg}{dt}$ is definit negative.

This means that g is a decreasing function of t, in particular:

$$g(1) \leq g(0) \iff (X+Y)^{-1} \leq X^{-1}$$

Similarly, we can show that $(X+Y)^{-1} \leq Y^{-1}$ with the function $h(t) = (Y+tX)^{-1}$ on $\mathbb{R}_+$

We just showed that $(X+Y)^{-1} \leq X^{-1} + Y^{-1}$

Moreover, the operator Tr is increasing so

$$\forall X, Y \in S_{++}^n, \quad \mathrm{Tr}\left((X+Y)^{-1}\right) \leq \mathrm{Tr}\left(X^{-1} + Y^{-1}\right)$$

$$\iff \mathrm{Tr}\left((X+Y)^{-1}\right) \leq \mathrm{Tr}\left(X^{-1}\right) + \mathrm{Tr}\left(Y^{-1}\right)$$

Let $X, Y \in \mathrm{dom} f$ and $\theta \in [0,1]$. Let us show that

$$f(\theta X + (1-\theta) Y) \leq \theta f(X) + (1-\theta) f(Y)$$

We apply the inequality we showed before:

$$\mathrm{Tr}\left((\theta X + (1-\theta) Y)^{-1}\right) \leq \mathrm{Tr}\left((\theta X)^{-1}\right) + \mathrm{Tr}\left(((1-\theta) Y)^{-1}\right)$$

$$\leq \theta\, \mathrm{Tr}\left(X^{-1}\right) + (1-\theta)\, \mathrm{Tr}\left(Y^{-1}\right)$$

So f is convex

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

~~Annexes~~

Let $(X, y) \in \text{dom } f$. As $X \in S^n_{++}$, X is diagonalizable

So : $\exists (V, D) \in S^n_{++}$, $X = VDV^T$ with $D = \text{diag}(\lambda_1, \dots, \lambda_n)$

$\lambda_1, \dots, \lambda_n$ the eigenvalues of X

~~An~~ and V the matrix composed of eigenvectors

As $X \in S^n_{++}$, all eigenvalues are strictly positives so

$$X^{-1} = (VDV^T)^{-1} = (V^T)^{-1} D^{-1} V^{-1}$$

$$= V D^{-1} V^T$$

$$= \sum_{i=1}^{n} \frac{1}{\lambda_i} v_i v_i^T$$

V is orthogonal so $V^T = V^{-1}$

with v_i the eigenvector associated to the eigenvalue λ_i

So $f(X, y) = y^T X^{-1} y$

$$= \sum_{i=1}^{n} \frac{1}{\lambda_i} y^T v_i v_i^T y$$

$$= \sum_{i=1}^{n} \frac{1}{\lambda_i} z_i^2$$

$v_i^T y = (y^T v_i)^T$ and $z_i = y^T v_i$

We can see that sum as a ~~su~~ supremum because z_i represents the projection of y into the space where X is diagonal but scaled with the corresponding eigenvalue. This means:

$$f(X, y) = \sum_{i=1}^{n} \frac{1}{\lambda_i} z_i^2 = \sup_{\lambda_i, i \in \{1, \dots, n\}} \left(\frac{1}{\lambda_i} \|y\|_2^2 \right)$$

As the norms are convex and the squared functions too, f is convex by applying a supremum to a convex function

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

Let $X \in S^n$. X can be decomposed on singlevalue:
So:

$$\exists (U, V, D) \in S^n, \ X = UDV^T \text{ with } \begin{cases} D = \text{diag}(\sigma_1(X), \dots, \sigma_n(X)) \\ U \text{ and } V \text{ are orthogonal} \end{cases}$$

Every single value of a matrix ~~can be~~ is a maximum:

For every $i \in \{1, \dots, n\}$, $\sigma_i(X) = \max_{\|y\|_2 = 1} \|Xy\|_2$

So $f(X) = \sum_{i=1}^{n} \max_{\|y\|_2=1} \langle Xy, e_i \rangle$ with e_i the i-th singular basis vector of $\mathbb{R}^n$

$$= \sup_{\substack{Y \in S^n \\ \|Y\|_2 = 1}} \sum_{i=1}^{n} \langle XY, e_i \rangle$$

~~So~~ Thus f can be expressed as the supremum of an affine function, which is convex

Exercise 4 1) To show that $K_{m+} = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$ is a proper cone we need to show that is non-empty, convex, closed and pointed.

This set is non empty because the vector $(1, 1, \dots, 1) \in K_{m+}$.

Let us show its convexity: Let $x, y \in K_{m+}$ and $\theta \in [0,1]$.

We need to show that $\theta x + 1 - \theta y \in K_{m+}$.

But $\begin{cases} x_1 \geq x_2 \geq \dots \geq x_n \geq 0 \Rightarrow \theta x_1 \geq \theta x_2 \geq \dots \geq \theta x_n \geq 0. \\ y_1 \geq y_2 \geq \dots \geq y_n \geq 0 \Rightarrow (1-\theta) y_1 \geq (1-\theta) y_2 \geq \dots \geq (1-\theta) y_n \geq 0 \end{cases}$

So $\theta x_1 + (1-\theta) y_1 \geq \theta x_2 + (1-\theta) y_2 \geq \dots \geq \theta x_n + (1-\theta) y_n \geq 0$

K_{m+} is thus convex

This set is closed because it contains all its limit points as it is defined with inequalities that are continuous conditions.

Finally, ~~the~~ K_{m+} is pointed. to show that let $x \in K_{m+}$ such that $-x \in K_{m+}$. We need to show that $x = 0$.

We ~~have~~ have $\begin{cases} x_1 \geq x_2 \geq \dots \geq x_n \geq 0 \text{ and} \\ -x_1 \geq -x_2 \geq \dots \geq -x_n \geq 0 \end{cases}$

So $x_1 \leq x_2 \leq \dots \leq x_n \leq 0 \leq x_n \leq \dots \leq x_2 \leq x_1$

The only vector that satisfies this condition is the null vector.

That shows that K_{m+} is pointed.

As K_{m+} is pointed, convex, non-empty and closed, <u>K_{m+} is a proper cone</u>

2) The dual cone of K_{m+} is defined as :

$$K^*_{m+} = \{ y \in \mathbb{R}^n \mid \langle x, y \rangle \geq 0 \text{ for all } x \in K_{m+} \}$$

We are going to show that $K^*_{m+} = K_{m+}$.

Let $y \in \mathbb{R}^n$. For all $x \in K_{m+}$, $\langle y, x \rangle \geq 0$.

In particular, for $x = (1, 0, \dots 0)$ give us $y_1 \geq 0$.

~~a~~ Similarly, we obtain these conditions:

- For $x = (1, 1, 0, \dots, 0)$, $y_1 + y_2 \geq 0$
- For $x = (1, 1, \dots, 1)$, $\sum_{i=1}^{n} y_i \geq 0$
- For $x = (\cancel{1}, \cancel{1}, \dots, \cancel{0})$, $\sum_{i=1}^{n-1} y_i \geq 0$

From all those condition on y, we can see appear that all component need to be non negative and their sums also need to be positives.

We can then conclude that the component need to be ordered as their sum need to remain positive.

Thus, we conclud that ~~all~~ y must be non negative and each conditions on the sum indicates that each component cannot exceed its predecessor, ~~so~~ it comes:

$$y_1 \geq y_2 \geq \dots \geq y_n \geq 0.$$

We showed that $K_{m+}^* = K_{m+}$.