DAVID SOTO

CONVEX OPTIMIZATION

HOMEWORK 1

Ex 1 :

1) Let R be a rectangle, a set of the form
$R = \{ x \in \mathbb{R}^n, \forall i = 1, \dots, n, \quad \alpha_i \leq x_i \leq \beta_i \}$ where $(\alpha_i)_i \in \mathbb{R}^n$ and $(\beta_i)_i \in \mathbb{R}^n$.

Let x and $y \in R$. Let $\lambda \in [0,1]$. By definition of x and y, we have that $\forall i = 1, \dots, n$, $\alpha_i \leq x_i \leq \beta_i$ and $\alpha_i \leq y_i \leq \beta_i$.
Since $\lambda \in [0,1]$, we have that $1-\lambda \in [0,1]$ and:

$$\forall i = 1, \dots, n, \quad \lambda \alpha_i \leq \lambda x_i \leq \lambda \beta_i$$
$$\text{and } (1-\lambda)\alpha_i \leq (1-\lambda) y_i \leq (1-\lambda)\beta_i .$$

$\Rightarrow \forall i = 1, \dots, n, \quad \lambda \alpha_i + (1-\lambda)\alpha_i \leq \lambda x_i + (1-\lambda) y_i \leq \lambda \beta_i + (1-\lambda)\beta_i$

That is $\forall i = 1, \dots, n, \quad \alpha_i \leq \lambda x_i + (1-\lambda) y_i \leq \beta_i$.
Hence, we can conclude that $\lambda x + (1-\lambda) y \in R$.
By definition R is a convex set.

2) Let $S = \{ (x_1, x_2) \in \mathbb{R}_+^2, \quad x_1 x_2 \geq 1 \}$ be a hyperbolic set of $\mathbb{R}^2$.
Let $(x_1, x_2) \in S$ and $(y_1, y_2) \in S$.
$\Rightarrow \quad x_1 x_2 \geq 1 \quad$ and $\quad y_1 y_2 \geq 1$.
Let $\lambda \in [0,1]$, therefore $(1-\lambda) \in [0,1]$ and we have:
$\circledast \quad \lambda^2 x_1 x_2 \geq \lambda^2 \quad$ and $\quad (1-\lambda)^2 y_1 y_2 \geq (1-\lambda)^2$.

$\lambda (x_1, x_2) + (1-\lambda)(y_1, y_2) = (\lambda x_1 + (1-\lambda) y_1, \lambda x_2 + (1-\lambda) y_2)$.

$$\Rightarrow \quad A = (\lambda x_1 + (1-\lambda) y_1) \times (\lambda x_2 + (1-\lambda) y_2)$$
$$A = \lambda^2 x_1 x_2 + \lambda(1-\lambda) x_1 y_2 + \lambda(1-\lambda) y_1 x_2 + (1-\lambda)^2 y_1 y_2$$
$$\circledast \Rightarrow \quad A \geq \lambda^2 + (1-\lambda)^2 + \lambda(1-\lambda)(y_1 x_2 + x_1 y_2)$$

In particular, $y_1 x_2 + x_1 y_2 = \dfrac{y_1 y_2 x_2^2}{y_2 x_2} + \dfrac{x_1 x_2 y_2^2}{y_2 x_2}$

$$y_1 x_2 + x_1 y_2 = \frac{1}{y_2 x_2}\left(\overbrace{y_1 y_2}^{\geqslant 1} x_2^2 + \overbrace{x_1 x_2}^{\geqslant 1} y_2^2\right) \geqslant 2$$

$\Rightarrow \quad x_1 y_2 + y_1 x_2 \geqslant 2$. We obtain:

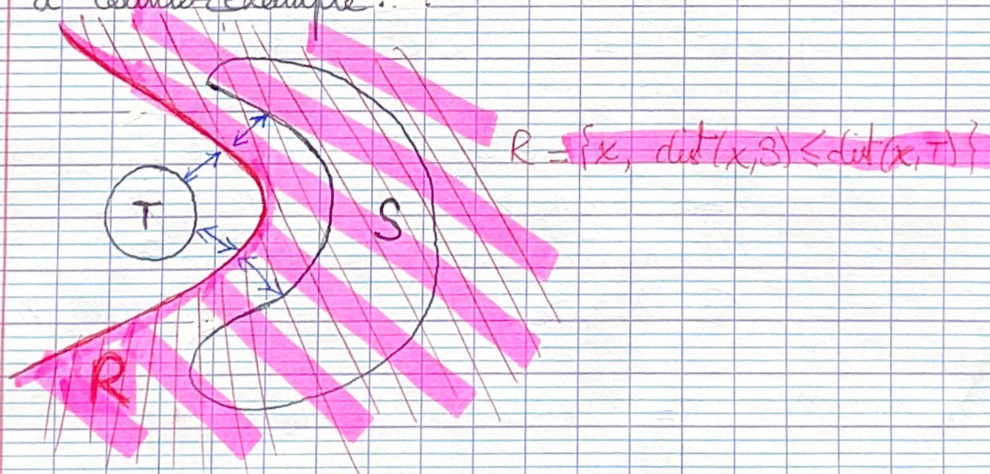
$A \geqslant \lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda)$

$A \geqslant (\lambda + (1-\lambda))^2$

$A \geqslant 1 \quad \Rightarrow$ By definition ~~we can conclude~~ we have $\lambda(x_1, x_2) + (1-\lambda)(y_1, y_2) \in S$ and we can conclude by definition that S is a convex set.

4) Let S and $T \subset \mathbb{R}^n$. Let $R = \{x \in \mathbb{R}^n,\ dist(x,S) \leq dist(x,T)\}$ be the set of points closer to S than T.

R is not necessarily a convex set. Here is a counter example:



R is clearly not a convex set here!

5) Let $S_1 \subset \mathbb{R}^n$ and $S_2 \subset \mathbb{R}^n$ be 2 subsets of $\mathbb{R}^n$ such that S_1 is convex.

Let $R = \{x \in \mathbb{R}^n,\ x + S_2 \subset S_1\}$

$R = \{x \in \mathbb{R}^n,\ \forall a \in S_2,\ a + x \in S_1\}$ be a subset of $\mathbb{R}^n$.

Let x and y be 2 points of R and let $\lambda \in [0,1]$.
By definition of x and y, we have:

$\forall a \in S_2, \quad a + x \in S_1$

$\forall a \in S_2, \quad a + y \in S_1$.

$\Rightarrow \quad \forall a \in S_2, \quad \lambda(a+x) \in \lambda S_1 = \{\lambda b,\ b \in S_1\}$

$\forall a \in S_2, \quad (1-\lambda)(a+y) \in (1-\lambda) S_1 = \{(1-\lambda) b,\ b \in S_1\}$.

Therefore, we have: $\forall a \in S_2$:

$\lambda(a+x) + (1-\lambda)(a+y) \in \lambda S_1 + (1-\lambda) S_1 = \{\lambda a + (1-\lambda) b,\ a \in S_1 \ \&\ b \in S_1\}$.

However S_1 is convex $\Rightarrow \forall (a,b) \in S_1^2,\ \forall \lambda \in [0,1]$
$\lambda a + (1-\lambda) b \in S_1$.

$\Rightarrow \lambda S_1 + (1-\lambda) S_1 = S_1$. Therefore we obtain:

$\forall a \in S_2, \quad \lambda a + \lambda x + (1-\lambda) a + (1-\lambda) y \in S_1$, that is

$\forall a \in S_2 \quad \lambda x + (1-\lambda) y + a \in S_1$, that is

$\lambda x + (1-\lambda) y + S_2 \subset S_1$. By definition

$\lambda x + (1-\lambda) y \in R \Rightarrow R$ is a convex set.

3) Let S be a subset of $\mathbb{R}^n$. Let $x_0 \in \mathbb{R}^n$.
Let $R = \{a \in \mathbb{R}^n, \|a - x_0\|_2 \leq \|a - y\|_2 \quad \forall y \in S\}$
be the set of points closer to x_0 than S.
R can be expressed in an other way:

$$R = \bigcap_{y \in S} \{x \in \mathbb{R}^n, \|x - x_0\|_2 \leq \|x - y\|_2\}$$, that is

R is an intersection of halfspaces (for every y in S, the set $\{x \in \mathbb{R}^n, \|x - x_0\|_2 \leq \|x - y\|_2\}$ is a halfspace of $\mathbb{R}^n$, therefore it is a convex set of $\mathbb{R}^n$).

$\Rightarrow$ R is the intersection of convex sets, we can conclude by the course that R is a convex set.

THM: $f: \mathbb{R}^n \to \mathbb{R}$ is convex iff $\forall x \in \operatorname{dom} f$, $\forall v \in \mathbb{R}^n$, the function $g: \mathbb{R} \to \mathbb{R}$ st $g(t) = f(x + tv)$ is convex (in t)
$\operatorname{dom} g = \{t \in \mathbb{R}, \ x + tv \in \operatorname{dom} f\}$.

3

Exercise 2 :

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

We define the function $g : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $g(t) = f((x_1, x_2) + t(v_1, v_2))$ $\forall (x_1, x_2) \in \mathbb{R}^2_{++}$, $\forall (v_1, v_2) \in \mathbb{R}^2$ and $\forall t \in \mathbb{R}$ such that $(x_1, x_2) + t(v_1, v_2) \in \mathbb{R}^2_{++}$

Let $(x_1, x_2) \in \mathbb{R}^2_{++}$, let $(v_1, v_2) \in \mathbb{R}^2$. $\forall t \in \mathbb{R}$ such that $(x_1, x_2) + t(v_1, v_2) \in \mathbb{R}^2_{++}$, we have :

$g(t) = f(x_1 + tv_1 , x_2 + tv_2)$

$g(t) = (x_1 + tv_1)(x_2 + tv_2) = x_1 x_2 + t x_1 v_2 + t v_1 x_2 + t^2 v_1 v_2$

$\Rightarrow$ $g'(t) = 2t\, v_1 v_2 + x_1 v_2 + v_1 x_2$

$\Rightarrow$ $g''(t) = 2 v_1 v_2$ $\forall t \in \mathbb{R}$ such that $x + tv \in \mathbb{R}^2_{++}$.

$(v_1, v_2) \in \mathbb{R}^2$, therefore $g''(t)$ could be positive as well as negative

We can conclude that g is neither convex nor concave and by using a theorem of the course we can conclude that f is neither convex nor concave on $\mathbb{R}^2_{++}$.

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

Let $g : \mathbb{R} \longrightarrow \mathbb{R}$ be a function defined by $\forall (x_1, x_2) \in \mathbb{R}^2_{++}$ $\forall (v_1, v_2) \in \mathbb{R}^2$, $\forall t \in \mathbb{R}$ such that $(x_1, x_2) + t(v_1, v_2) \in \mathbb{R}^2_{++}$

$g(t) = f(x_1 + tv_1, x_2 + tv_2)$

$g(t) = \dfrac{x_1 + tv_1}{x_2 + tv_2}$.

$\Rightarrow$ $g'(t) = \dfrac{v_1(x_2 + tv_2) - v_2(x_1 + tv_1)}{(x_2 + tv_2)^2}$

$$g'(t) = \frac{v_1 x_2 + \cancel{t v_1 v_2} - v_2 x_1 - \cancel{t v_1 v_2}}{(x_2 + t v_2)^2} = \frac{v_1 x_2 - v_2 x_1}{(x_2 + t v_2)^2}$$

$$\Rightarrow g''(t) = (v_1 x_2 - v_2 x_1) \times (-1) \times \frac{2 v_2 (x_2 + t v_2)}{(x_2 + t v_2)^4}$$

$$g''(t) = \frac{(v_1 x_2 - v_2 x_1)(-2 v_2 x_2 - 2 t v_2^2)}{(x_2 + t v_2)^4}$$ always positive

$\Rightarrow g''(t)$ has the same sign as $(v_1 x_2 - v_2 x_1)(-2 v_2 x_2 - 2 t v_2^2)$

for $x = (1,1)$, $v = (2,1)$ and $t = 1$

$$\Longleftrightarrow g''(1) = \frac{-4}{1} < 0$$

and for $x = (4,1)$, $v = (2,1)$ and $t = 1$

$$\Longleftrightarrow g''(1) = \frac{8}{1} > 0$$

Hence, we can conclude that g is neither convex nor concave, and by using a theorem of the course we can conclude that
f is neither convex nor concave.

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

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

$$\Rightarrow \text{Hess} f(x_1, x_2) = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix} \quad \forall (x_1, x_2) \in \mathbb{R}^2_{++}$$

$$\Rightarrow \det(\text{Hess} f(x_1, x_2)) = \frac{4}{x_1^4 x_2^4} - \frac{1}{x_1^4 x_2^4}$$

$$\det(\text{Hess} f(x_1, x_2)) = \frac{3}{x_1^4 x_2^4} > 0$$

(4)

$$Tr(Hess f(x_1,x_2)) = \frac{2}{x_1^3 x_2} + \frac{2}{x_1 x_2^3} > 0 \quad \text{since } (x_1,x_2) \in \mathbb{R}_{++}^2$$

We have that $\det(Hess f(x_1,x_2)) > 0$ and $Tr(Hess f(x_1,x_2)) > 0$

$\Rightarrow$ the eigenvalues of $Hess f(x_1,x_2)$ are positive $\forall (x_1,x_2) \in \mathbb{R}_{++}^2$.

$\Rightarrow \forall (x_1,x_2) \in \mathbb{R}_{++}^2$, $Hess f(x_1,x_2)$ is positive definite

$\Rightarrow$ f is convex on $\mathbb{R}_{++}^2$.

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

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

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

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

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

$$\det(Hess f(x_1,x_2)) = 0$$

and $$Tr(Hess f(x_1,x_2)) = \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} + \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1}$$

$$= \underbrace{\alpha(\alpha-1)}_{\substack{\leq 0 \\ \text{since } \alpha \in [0,1]}} \underbrace{\left(x_1^{\alpha-2} x_2^{1-\alpha} + x_1^{\alpha} x_2^{-\alpha-1} \right)}_{\geq 0 \text{ since } (x_1,x_2) \in \mathbb{R}_{++}^2}$$

$\Rightarrow Tr(Hess f(x_1,x_2)) \leq 0$.

We have $\det(Hess f(x_1,x_2)) = 0$ and $Tr(Hess f(x_1,x_2)) \leq 0$

$\Rightarrow$ one of the eigenvalues of $Hess f(x_1,x_2)$ is equal to 0

and the other eigenvalue of $Hess\, f(x_1, x_2)$ is negative $\forall (x_1, x_2) \in \mathbb{R}_{++}^2$

$\Rightarrow$ $\forall (x_1, x_2) \in \mathbb{R}_{++}^2$, $Hess\, f(x_1, x_2)$ is negative semi definite

$\Rightarrow$ f is concave on $\mathbb{R}_{++}^2$.

5

Exercise 3:

2) $f(X,y) = y^T X^{-1} y$ defined on $S^n_{++} \times \mathbb{R}^n$.
Let $(X,y) \in S^n_{++} \times \mathbb{R}^n$. ~~[illegible]~~:

$f(X,y) = y^T X^{-1} y = 2 \times \frac{1}{2} y^T X^{-1} y$. We know from the course that:

$$f(X,y) = 2 \sup_{x \in \mathbb{R}^n} \left(y^T x - \frac{1}{2} x^T X x \right)$$

$$f(X,y) = \sup_{x \in \mathbb{R}^n} \left(\boxed{2y^T x - x^T X x} \right)$$

The function $[(X,y) \longmapsto 2y^T x - x^T X x]$ is clearly a linear function! Therefore it is an convex function ~~and we then know from the course that it is a convex function~~.

From the course we know that pointwise supremum preserves the convexity of a function, and since $[(X,y) \longmapsto 2y^T x - x^T X x]$ is a convex function, we then deduce that

$f(X,y) = \sup_{x \in \mathbb{R}^n} \left(2y^T x - x^T X x \right)$ is a convex function on $S^n_{++} \times \mathbb{R}^n$

That is $f(X,y) = y^T X^{-1} y$ is a convex function on $S^n_{++} \times \mathbb{R}^n$.

3) $f(X) = \sum_{i=1}^{n} \sigma_i(X)$ defined on S^n. where $\sigma_1(X), \dots, \sigma_n(X)$ are singular values of $X \in S^n$.
$\forall X \in S^n$, since X has only real coefficients, we have:

$f(X) = Tr\left(\left(X^T X\right)^{1/2}\right) = \|X\|_{Ky\,Fan}$

$\hookrightarrow$ f is equal to the Ky Fan norm! (the Trace norm)

We can conclude by the course that f is convex on S^n

(Indeed, we know from the course that any norm is a convex function).

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