MATHÉMATIQUES
VISION
APPRENTISSAGE

Convex Optimization - Homework 1

Louis Martinez
louis.martinez@telecom-paris.fr

Exercise 1

1)

Let x and y two elements of the rectangle, and $\lambda \in [0,1]$. By definition of the rectangle we have the following inequalities:

$$\begin{cases} \alpha \leq x \leq \beta \\ \alpha \leq y \leq \beta \end{cases}$$

By summing both lines we get:

$$\lambda\alpha + (1-\lambda)\alpha \leq \lambda x + (1-\lambda)y \leq \lambda\beta + (1-\lambda)\beta$$
$$\alpha \leq \lambda x + (1-\lambda)y \leq \beta$$

So $\lambda x + (1-\lambda)y$ is an element of the rectangle.

$\boxed{\text{The rectangle is convex.}}$

2)

We define the real function $f : x \mapsto \frac{1}{x}$. Let (x_1, x_2) and (y_1, y_2) two elements of the hyperbolic set. By definition of the set we have:

$$\begin{cases} x_1 \geq f(x_2) \\ y_1 \geq f(y_2) \end{cases}$$

Let $\lambda \in [0,1]$ and $z = \lambda x + (1-\lambda)y$.

$$\begin{aligned} z_1 = \lambda x_1 + (1-\lambda)y_1 &\geq \lambda f(x_2) + (1-\lambda)f(y_2) && \text{Definition of the set} \\ &\geq f(\lambda x_2 + (1-\lambda)y_2) = f(z_2) && \text{Because } f \text{ is convex} \end{aligned}$$

So z is an element of the hyperbolic set.

$\boxed{\text{The hyperbolic set is convex.}}$

3)

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2, \forall y \in S\} = \bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$$

For a fixed y, these sets define half spaces, which are convex sets. So we have an intersection of convex sets.

$\boxed{\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2, \forall y \in S\} \text{ is convex.}}$

4)

Let's fix $n \in \mathbb{N}$, $S = \{0, 1\} \subset \mathbb{R}$ and $T = \left\{\frac{1}{2}\right\} \subset \mathbb{R}$. We can rewrite the set as:

$$\{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\} = \left\{x \mid x \leq \frac{1}{4} \text{ or } x \geq \frac{3}{4}\right\}$$

$\boxed{\text{This set is not convex.}}$

5)

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

For $y \in S_2$ fixed, we can define $f_y : x \mapsto x - y$, affine in x. So

$$\forall y \in S_2, f_y(S_1) = S_1 - y$$

is convex.

$\boxed{\{x \mid x + S_2 \subseteq S_1\} \text{ is convex as an intersection of convex sets.}}$

Exercise 2

1)

f is twice differentiable on $\mathrm{dom} f = \mathbb{R}^2_{++}$. The Hessian matrix of f is:

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

This matrix is neither positive semi-definite (for instance $\begin{pmatrix} -1 & 1 \end{pmatrix} \nabla^2 f(x) \begin{pmatrix} -1 & 1 \end{pmatrix}^T = -1 < 0$) nor negative semi-definite

$\boxed{f \text{ is neither convex nor concave.}}$

2)

f is twice differentiable on $\mathrm{dom} f = \mathbb{R}^2_{++}$. The Hessian matrix of f is:

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

This matrix is semi definite positive. Indeed, for any $u = (u_1, u_2) \in \mathbb{R}^2$ we have:

$$u^T \nabla^2 f(x) u = 2 \frac{(u_1 x_2 + u_2 x_1)^2}{x_1^3 x_2^3} > 0 \qquad \text{because } (x_1, x_2) > 0$$

$\boxed{f \text{ is convex.}}$

3)

f satisfies the same conditions as the previous questions. Its Hessian matrix is:

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

This matrix is neither positive nor negative semi definite. Indeed:

$$\frac{x_2^2}{2} \left(u^T \nabla^2 f(x) u\right) = u_2^2 \frac{x_1}{x_2} - u_1 u_2$$

. With $(u_1, u_2) = (0, 1)$ this quantity is positive. But with $u_1 = 1$ and $u_2 < \frac{x_2}{x_1}$, it's negative.

$\boxed{f \text{ is neither convex nor concave.}}$

4)

f satisfies the same conditions as the previous questions. Its Hessian matrix is:

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

For any $u = (u_1, u_2) \in \mathbb{R}^2$, we get:

$$\frac{1}{\alpha(1-\alpha)}\left(u^T \nabla^2 f(x) u\right) = -\left(u_1 x_1^{\frac{\alpha}{2}-1} x_2^{\frac{1-\alpha}{2}} - u_2 x_1^{\frac{\alpha}{2}} x_2^{\frac{-\alpha-1}{2}}\right)^2 \leq 0$$

So $u^T \nabla^2 f(x) u$ is negative semi-definite.

$\boxed{f \text{ is concave}}$

Exercise 3

1)

Let $V \in S^n$ and $X \succ 0$. We define $g : t \mapsto f(X + tV)$

$$\begin{aligned} f(X+tV) &= \operatorname{Tr}\left((X+tV)^{-1}\right) \\ &= \operatorname{Tr}\left(\left[X^{\frac{1}{2}}\left(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}\right)X^{\frac{1}{2}}\right]^{-1}\right) && \text{(Factorization)} \\ &= \operatorname{Tr}\left(X^{-\frac{1}{2}}\left(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}\right)^{-1}X^{-\frac{1}{2}}\right) && \text{(Properties of inverse)} \\ &= \operatorname{Tr}\left(X^{-1}\left(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}\right)^{-1}\right) && \text{(Properties of Tr)} \\ &= \operatorname{Tr}\left(X^{-1}\left(I + tP\Lambda P^T\right)^{-1}\right) && \text{(Since } X^{-\frac{1}{2}}VX^{-\frac{1}{2}} \text{ is symmetric)} \end{aligned}$$

with $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$ where $(\lambda_i)_{1 \leq i \leq n}$ are the eigenvalues of $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$.

$$\begin{aligned} f(X+tV) &= \operatorname{Tr}\left(X^{-1}P(I+t\Lambda)^{-1}P^T\right) && \text{(Factorization with } I = PP^T\text{)} \\ &= \operatorname{Tr}\left((P^TX^{-1}P)(I+t\Lambda)^{-1}\right) && \text{(Factorization with } I = PP^T\text{)} \end{aligned}$$

$$f(X+tV) = \sum_{i=1}^n \left(P^TX^{-1}P\right)_{ii}\left(\frac{1}{1+t\lambda_i}\right)$$

$\forall 1 \leq i \leq n, t \mapsto \frac{1}{1+\lambda_i}$ is convex so g is convex as a weighted sum of convex functions.

$\boxed{f \text{ is convex}}$

2)

Let $X \in S_{++}^n(\mathbb{R})$. $\forall (y,z) \in \mathbb{R}^n$ we have:

$$\|X^{-\frac{1}{2}}y - X^{\frac{1}{2}}z\|^2 = y^TX^{-1}y + z^TXz - 2y^Tz \geq 0$$

Hence,

$$\frac{1}{2}y^TX^{-1}y \geq y^Tz - \frac{1}{2}z^TXz$$

We conclude that $\frac{1}{2}y^TX^{-1}y$ is an upper bound of $y^Tz - \frac{1}{2}z^TXz$.
this upper bound is reached by choosing $z = X^{-1}y$. Hence:

$$\frac{1}{2}y^TX^{-1}y = \sup_{z \in \mathbb{R}^n} y^Tz - \frac{1}{2}z^TXz$$

The right hand-side of the equality is affine in y and X.

$\boxed{f \text{ is convex}}$

3)

Let $X \in S^n$. We denote by $\|X\|_* = \sigma_{\max}(X)$ the spectral norm of X where $\sigma_{\max}(X)$ is the largest singular value of X. We can show that $X \mapsto \sum_{i=1}^n \sigma_i(X)$ is the dual norm of the spectral norm: Formally speaking, let's show that:

$$\sup_{\sigma_{\max}(Y)\le 1} |\langle X, Y\rangle| = \sum_{i=1}^n \sigma_i(X)$$

- First we define $X = U\Sigma V^T$ the SVD (U and V are unitary matrices and $\Sigma = \operatorname{diag}\left(\sigma_1(X), ..., \sigma_{n(X)}\right)$) of X, and $Y = UV^T = UIV^T$. With this definition, Y is unitary so $\sigma_{\max}(Y) = 1$. We have:

$$\begin{aligned} \langle Y, X\rangle &= \langle UV^T, U\Sigma V^T\rangle \\ &= \operatorname{Tr}\left(VU^TU\Sigma V^T\right) \qquad \text{(by definiton)} \\ &= \operatorname{Tr}\left(V^TVU^TU\Sigma\right) \qquad (\operatorname{Tr}(AB) = \operatorname{Tr}(BA)) \\ &= \sum_{i=1}^n \sigma_i(X) \qquad\qquad (UU^T = I \text{ and} VV^T = I) \end{aligned}$$

So

$$\sup_{\sigma_{\max}(\tilde{Y})\le 1} |\langle \tilde{Y}, X\rangle| \ge \sum_{i=1}^n \sigma_i(X)$$

- We now prove the other inequality:

$$\begin{aligned} \sup_{\sigma_{\max}(\tilde{Y})\le 1} \langle \tilde{Y}, X\rangle &= \sup_{\sigma_{\max}(\tilde{Y})\le 1} \operatorname{Tr}\left(\tilde{Y}^T U\Sigma V^T\right) \\ &= \sup_{\sigma_{\max}(\tilde{Y})\le 1} \operatorname{Tr}\left(V^T\tilde{Y}^TU\Sigma\right) \qquad (\operatorname{Tr}(AB) = \operatorname{Tr}(BA)) \\ &= \sup_{\sigma_{\max}(\tilde{Y})\le 1} \langle U^T\tilde{Y}V, \Sigma\rangle \\ &= \sup_{\sigma_{\max}(\tilde{Y})\le 1} \sum_{i=1}^n \left(U^T\tilde{Y}V\right)_{ii}\sigma_i(X) \\ &= \sup_{\sigma_{\max}(\tilde{Y})\le 1} \sum_{i=1}^n u_i^T\tilde{Y}v_i\sigma_i(X) \\ &\le \sup_{\sigma_{\max}(\tilde{Y})\le 1} \sum_{i=1}^n \sigma_{\max}\left(\tilde{Y}\right)\sigma_i(X) \\ &= \sum_{i=1}^n \sigma_i(X) \end{aligned}$$

So

$$\sup_{\sigma_{\max}(\tilde{Y})\le 1} |\langle \tilde{Y}, X\rangle| \le \sum_{i=1}^n \sigma_i(X)$$

With both previous points we finally have:

$$\sup_{\sigma_{\max}(\tilde{Y})\le 1} |\langle \tilde{Y}, X\rangle| = \sum_{i=1}^n \sigma_i(X)$$

This shows that $f : X \mapsto \sum_{i=1}^n \sigma_i(X)$ is the dual norm of the spectral, so it's a norm.

$\boxed{f \text{ is convex.}}$
