Convex optimization: homework 1

Exercise 1

① **CONVEXE** en tant qu'intersection des demi-espaces $\{x_i \le \beta_i\}$ et $\{x_i \ge \alpha_i\}$ pour $i=1,\dots,n$ (et un demi-espace est bien convexe)

② **CONVEXE** car $\{x \in \mathbb{R}^2_+ \mid x_1 x_2 \ge 1\} = \{x \in \mathbb{R}^2_+ \mid \frac{1}{x_1} \le x_2\}$ ($x_1 \neq 0$ car sinon $x_1 x_2 = 0 \not\ge 1$) est l'épigraphe de la fonction convexe $x_1 \mapsto \frac{1}{x_1}$; c'est donc un ensemble convexe.

③ **CONVEXE** car $\{x \mid \|x-x_0\|_2 \le \|x-y\|_2 \ \forall y \in S\} = \bigcap_{y\in S} \underbrace{\{x \mid \|x-x_0\|_2 \le \|x-y\|_2\}}_{=: A_{x_0,y}}$ est une intersection d'ensembles convexes. En effet:

$x \in A_{x_0,y}$ ssi $\|x-x_0\|_2^2 \le \|x-y\|_2^2$

ssi $\cancel{\|x\|_2^2} - 2x_0^T x + \|x_0\|_2^2 \le \cancel{\|x\|_2^2} - 2y^T x + \|y\|_2^2$

ssi $2(y-x_0)^T x \le \|y\|_2^2 - \|x_0\|_2^2$

donc $A_{x_0,y}$ est un demi-espace (et est donc convexe)

④ **NON CONVEXE**: il suffit de choisir S non convexe et T son complémentaire dans $\mathbb{R}^n$, on a alors $\{x \mid \mathrm{dist}(x,S) \le \mathrm{dist}(x,T)\} = S$ non convexe.

⑤ **CONVEXE**: $\{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x+y \in S_1\} = \bigcap_{y\in S_2} \underbrace{(S_1 - y)}_{\text{ensemble convexe car } S_1 \text{ est convexe}}$,

est convexe en tant qu'intersection de convexes.

Exercice 2

① **NI CONVEXE NI CONCAVE**: f est deux fois différentiable et $\nabla^2 f(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ a deux valeurs propres 1 et -1 donc $\nabla^2 f(x)$ n'est ni positive semi-définie ni semi-définie négative (pour $x = \binom{0}{0}$ par exemple).

② **NON CONCAVE** sans quoi, par composition avec la fonction convexe $x \mapsto \frac{1}{x}$ définie sur $\mathbb{R}_{++}$ et non-croissante, on aurait $h(x,y) = xy$ convexe, ce qui est absurde d'après ①.

CONVEXE car f est deux fois différentiable et: $\nabla^2 f(x_1,x_2) = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix}$

pour $(x_1,x_2) \in \mathbb{R}^2_{++}$. Soit $(y_1,y_2) \in \mathbb{R}^2$. On a:

$$(y_1, y_2)\, \nabla^2 f(x_1,x_2) \binom{y_1}{y_2} = \left(\frac{2y_1}{x_1^3 x_2} + \frac{y_2}{x_1^2 x_2^2},\ \frac{y_1}{x_1^2 x_2^2} + \frac{2y_2}{x_1 x_2^3} \right) \binom{y_1}{y_2}$$

$$= 2\left(\frac{y_1^2}{x_1^3 x_2} + \frac{y_1 y_2}{x_1^2 x_2^2} + \frac{y_2^2}{x_1 x_2^3} \right)$$

$$= \frac{2}{x_1^2 x_2^2} \underbrace{\left(y_1^2 \frac{x_2}{x_1} + y_2^2 \frac{x_1}{x_2} + y_1 y_2 \right)}_{= \left(y_1 \sqrt{\frac{x_2}{x_1}} + \frac{1}{2} y_2 \sqrt{\frac{x_1}{x_2}} \right)^2 + \frac{3}{4} y_2^2 \frac{x_1}{x_2} \ge 0} \ge 0$$

Donc $\nabla^2 f(x_1,x_2)$ est semi-définie positive $\forall (x_1,x_2) \in \mathbb{R}^2_{++}$ donc f est convexe.

③ NI CONVEXE NI CONCAVE : f est deux fois différentiable et $\nabla^2 f\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3}\end{pmatrix}$

Soit $(y_1,y_2)\in\mathbb{R}^2$ et $(x_1,x_2)\in\mathbb{R}^2_{++}$. Alors:

$$(y_1,y_2)\,\nabla^2 f(x_1,x_2)(y_1,y_2)^T = \left(-\frac{y_2}{x_2^2},\ \frac{-y_1}{x_2^2}+\frac{2x_1y_2}{x_2^3}\right)\begin{pmatrix}y_1\\y_2\end{pmatrix}$$

$$= \frac{2}{x_2^2}\left(\frac{x_1}{x_2}y_2^2 - y_1y_2\right) \quad =: F_{x_1,x_2}(y_1,y_2)$$

en $(x_1,x_2)=(1,1)$, on a $F_{1,1}(0,1) = 2>0$

et $F_{1,1}(2,1) = -2<0$

donc $\nabla^2 f(1,1)$ n'est ni positive ni négative.

④ $f(x_1,x_2)=x_1^\alpha x_2^{1-\alpha} = x_2\left(\frac{x_1}{x_2}\right)^\alpha$ est la fonction perspective de $\cdot^\alpha$ qui est concave pour $\alpha\in[0,1]$ et convexe pour $\alpha=0$ ou $\alpha=1$.

Donc : f est CONCAVE pour $0\le\alpha\le1$

CONVEXE pour $\alpha\in\{0,1\}$

Exercice 3

① Soit $X\in S^n_{++}$ et $V\in S^n$. On considère la fonction $g: t\mapsto f(X+tV)$.

Si $X+tV\in S^n_{++}$, on a:

$$g(t) = tr\left((X+tV)^{-1}\right)$$

$$= tr\Big(X^{-1}\big(I + t\underbrace{X^{-1/2}VX^{-1/2}}_{=:\,ODO^T}\big)^{-1}\Big) \quad \text{avec } D = diag(\lambda_i)_{i=1,\dots,n}$$

$$= tr\left(O^TX^{-1}O(I+tD)^{-1}\right)$$

$$= \sum_{i=1}^n\sum_{k=1}^n (O^TX^{-1}O)_{i,k}\,\underbrace{(I+tD)^{-1}_{k,i}}_{=\,\delta_{k,i}\left(\frac{1}{1+t\lambda_i}\right)}$$

$$= \sum_{i=1}^n (O^TX^{-1}O)_{i,i}\,\frac{1}{1+t\lambda_i}$$

Pour montrer que g est convexe, il suffit de vérifier que (1) $(O^TX^{-1}O)_{i,i}>0\ \forall i$

(2) $t\mapsto\frac{1}{1+t\lambda_i}$ est convexe.

(1) $X\in S_n^{++}$ donc $O^TX^{-1}O\in S_n^{++}$ puis $(O^TX^{-1}O)_{i,i} = x^TO^TX^{-1}Ox > 0$ car $O^TX^{-1}O\in S^n_{++}$

avec $x=(0,\dots,0,1,0,\dots,0)^T$
1 en i-ème position

(2) il suffit de vérifier que $I+tX^{-1/2}VX^{-1/2}>0$. Soit $z\in\mathbb{R}^n$.

$z^T(I+tX^{-1/2}VX^{-1/2})z = (X^{-1/2}z)^T(X+tV)(X^{-1/2}z)>0$ car $X+tV>0$

donc $1+t\lambda_i>0\ \forall t$ puis $t\mapsto\frac{1}{1+t\lambda_i}$ est convexe car $t\mapsto 1+t\lambda_i$ est affine et $x\mapsto\frac{1}{x}$ est convexe et non croissante sur $\mathbb{R}_{++}$

② l'épigraphe de f est $\{(X,y,t) \mid y^T X^{-1} y \leq t\} = \{(X,y,t) \mid \underbrace{t - y^T X^{-1} y}_{} \geq 0$

on reconnaît le complément de Schur de $\begin{bmatrix} X & y \\ y^T & t \end{bmatrix}$ et comme $X \succ 0$ alors on a $t - y^T X^{-1} y \geq 0$ ssi $\begin{bmatrix} X & y \\ y^T & t \end{bmatrix} \succeq 0$

donc $\operatorname{epi} f = \{(X,y,t) \mid \begin{bmatrix} X & y \\ y^T & t \end{bmatrix} \succeq 0\}$

et on reconnaît l'ensemble de faisabilité d'un problème convexe à contraintes d'inégalités généralisées, qui est donc convexe. Donc f est convexe.

③ il suffit de montrer que f est une norme. La seule chose non triviale à démontrer est l'inégalité triangulaire. Montrons pour cela que $f(X) = \sup_{\sigma_{max}(Q) \leq 1} \langle Q, X \rangle_F$

(≤) SVD de X : $X = U \Sigma V^T$

Prenons $Q = UV^T$; Alors $QQ^T = I$ donc $\sigma_{max}(Q) = 1 \leq 1$ et

$\langle Q, X \rangle_F = Tr(Q^T X) = Tr(V \underbrace{U^T U}_{=I} \Sigma V) = Tr(\underbrace{V V^T}_{=I} \Sigma) = Tr(\Sigma) = f(X)$

(≥) $\sup_{\sigma_{max}(Q) \leq 1} \langle Q, X \rangle_F = \sup Tr(Q^T X)$

$= \sup Tr(Q^T U \Sigma V^T)$

$= \sup Tr(V^T Q^T U \Sigma)$

$= \sup \underbrace{Tr((U^T Q V)^T \Sigma)}$

$= \sum_{i=1}^{n} ((U^T Q V)^T \Sigma)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} (U^T Q V)^T_{ij} \underbrace{\Sigma_{ij}}_{= \delta_{ij} \sigma_i(X)}$

$= \sup \sum_{i=1}^{n} \sigma_i(X) \underbrace{(U^T Q V)_{ii}}_{= u_i^T Q v_i \leq \sup_{\|u\|=1, \|v\|=1} u^T Q v = \sigma_{max}(Q)}$

$\leq \sup \sum_{i=1}^{n} \sigma_i(X) \underbrace{\sigma_{max}(Q)}_{\leq 1}$

$= f(X)$

donc, en conclusion, on a $f(X+Y) = \sup \langle Q, X+Y \rangle \leq \sup \langle Q, X \rangle + \sup \langle Q, Y \rangle = f(X) + f(Y)$

donc f est bien une norme et donc f est convexe.