Optimisation convexe

Simon QUERIC.

Devoir Maison 1

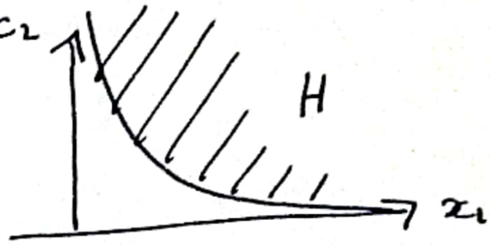
Exercice 1

1) On note $(e_i)_{1\leq i\leq n}$ la base canonique de $\mathbb{R}^n$.

$$R = \{ x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i \ \forall i \in [n] \}$$
$$= \bigcap_{i=1}^{n} \{ x \in \mathbb{R}^n \mid x^T e_i \geq \alpha_i \} \cap \{ x \in \mathbb{R}^n \mid x^T e_i \leq \beta_i \}$$

Le rectangle R est convexe en tant qu'intersection de demi-plans affines.

2) $$H = \{ x \in \mathbb{R}_+^2 \mid x_1 x_2 \geq 1 \}$$
$$= \{ (x,t) \in \mathbb{R}_{++}^2 \mid t \geq \frac{1}{x} \}$$
$$= \text{epi}\left(x \in \mathbb{R}_{++} \longmapsto \frac{1}{x} \right)$$



H est convexe car c'est l'épigraphe de la fonction inverse qui est convexe sur $\mathbb{R}_{++}$.

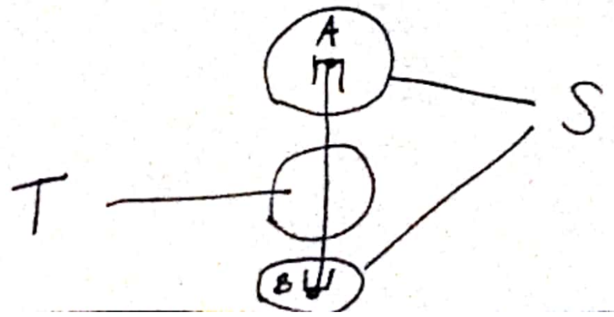
3) Soit $y \in S$. $\|x - x_0\|_2 \leq \|x - y\|_2$

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

L'ensemble considéré est $\bigcap_{y \in S} \{ x \in \mathbb{R}^n \mid 2x^T(y - x_0) \leq \|y\|^2 - \|x_0\|^2 \}$

Il est bien convexe en tant qu'intersection de demi-plans affines.

4) On considère les ensembles S et $T \subseteq \mathbb{R}^2$:



Il est clair que $\{ x \mid \text{dist}(x,S) \leq \text{dist}(x,T) \}$ n'est pas convexe, ~~Simon~~ car $[A,B]$ n'y appartient pas.

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

Soit $t \in [0,1]$ et $s \in S_2$

Alors : $tx + (1-t)y + s = \underbrace{t(x+s)}_{\in S_1} + (1-t)\underbrace{(y+s)}_{\in S_2} \in S_1$

donc $tx + (1-t)y \in E$

E est convexe.

Exercice 2

1) On a : $f(x_1, x_2) = x^T A x$ où $A = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$.

$\det A = -1/4 < 0$ donc $A \notin S^n_+$. Or $\nabla^2 f(x_1, x_2) = A$ pour tout x_1, x_2. f n'est donc ni convexe ni concave.

Montrons sa quasiconcavité.

Soit $\alpha \in \mathbb{R}$. $C_\alpha = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid x_1 \geq \frac{\alpha}{x_2}\}$ est vide si $\alpha < 0$ et sinon il s'agit de l'épigraphe de la fonction convexe $t \in \mathbb{R}_{++} \mapsto \frac{\alpha}{t}$ qui est convexe.

f est bien quasiconcave.

2) Montrons que f est convexe sur $\mathbb{R}^2_{++}$.

Soit $x, y \in \mathbb{R}^2_{++}$, $t \in [0,1]$.

$$f(tx + (1-t)y) = \frac{1}{(tx_1 + (1-t)y_1)(tx_2 + (1-t)y_2)}$$

$$\leq \frac{1}{x_1^t y_1^{(1-t)}} \frac{1}{x_2^t y_2^{1-t}}$$ par inégalité arithmético géométrique

$$= \left(\frac{1}{x_1 x_2}\right)^t \left(\frac{1}{y_1 y_2}\right)^{1-t}$$ idem.

$$\leq t f(x_1, x_2) + (1-t) f(y_1, y_2)$$

f est donc convexe □

3) On calcule la hessienne de f.

$$\nabla^2 f = \begin{pmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2 \end{pmatrix}, \quad \det \nabla^2 f = -\frac{1}{x_2^2} < 0 \text{ car } x_2 \in \mathbb{R}_{++}$$

donc f n'est ni convexe ni concave

Soit $\alpha \in \mathbb{R}$. $C_\alpha = \{(x_1, x_2) \mid \frac{x_1}{x_2} \overset{(*)}{\leq} \alpha\}$

$= \{(x_1, x_2) \mid 0 \leq (-1 + \alpha)^T \binom{x_1}{x_2}\}$

Les C_α sont des demi-plans affines intersectés avec $\mathbb{R}^2_{++}$, donc sont convexes. Si on change l'inégalité (*) de sens on obtient des ensembles également convexe. f est donc quasilinéaire.

4) On calcule la hessienne de f. $\nabla^2 f = *$ ~~$\alpha(\alpha-1) x_1^{\alpha-1} x_2^{1-\alpha}$~~

$* \quad \alpha(\alpha-1)\begin{pmatrix} x_1^{\alpha-2} x_2^{1-\alpha} & x_1^{\alpha-1} x_2^{-\alpha} \\ x_1^{\alpha-1} x_2^{-\alpha} & x_1^{\alpha} x_2^{-\alpha-1} \end{pmatrix}$ d'où $\det \nabla^2 f = 0$. ~~$= \alpha$~~

donc : $Sp(\nabla^2 f) = \{0, Tr(\nabla^2 f)\}$ et $Tr(\nabla^2 f) = \alpha(\alpha-1)(x_1^{\alpha-1} x_2^{1-\alpha} + x_1^{\alpha} x_2^{-\alpha-1})$

≤ 0

donc, si $\alpha = 0$ ou 1 f est linéaire

sinon f est concave.

<u>Exercice 3</u>. Soit $X, V \in S^n_{++}$ et t tq $X + tV \in \text{dom} f$.

On note $P^T D_t P = I + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$ avec $\begin{cases} P \in O_n \\ D_t \text{ diagonale} \end{cases}$

On remarque que $Sp(D_t) = \{1 + t\lambda_i, \lambda_i \in Sp(\sqrt{X} V \sqrt{X})\}$

et que chaque $\lambda_i \geq 0$.

Ainsi $Tr((X + tV)^{-1}) = Tr(X^{-1} P^T D_t P)$

$= Tr(P X^{-1} P^T D_t)$

$= \sum_{i=1}^{n} \frac{(P X^{-1} P^T)_{ii}}{1 + t\lambda_i}$ $\quad (PX^{-1}P^T)_{ii} \geq 0$

Chaque $t \longmapsto \frac{(P X^{-1} P^T)_{ii}}{1 + t\lambda_i}$ est bien convexe sur $\{t \mid X + tV \in \text{dom} f\}$

donc $t \longmapsto Tr((X + tV)^{-1})$ est convexe et f l'est sur S^n_{++}.

2) On remarque que $f(X,y) = y^T X y$ $\overbrace{g(u,X,y)}$

$$= \sup_{u \in \mathbb{R}^n} \left(\cdot \frac{u^T X u}{4} - u^T y \right)$$

pour tout $u \in \mathbb{R}^n$, $g(u,\cdot,\cdot)$ est linéaire en (X,y)
f est donc convexe sur $S_n^{++} \times \mathbb{R}^n$.

3) Montrons que $f(X) = \sup_{M \mid M^T M = I_n} Tr(MX)$.

$X \in S^n$ est diagonalisable : $X = P^T D P$ où $\begin{cases} P \in O_n \\ D \text{ diag.} \end{cases}$

d'où : $Tr(MX) = Tr(MP^T DP)$
$= Tr(PMP^T D)$

et $M \in O_n \mapsto \underset{\in O_n}{PMP^T}$ est bijective donc

$$\sup \{Tr(MX),\ M \in O_n\} = \sup \{Tr(MD),\ M \in O_n\}$$

Or pour tout $i,j \in [n]$: $|M_{ij}| \leq 1$ pour $M \in O_n$.

donc $Tr(MD) = \sum_{i=1}^n M_{ii} D_{ii} \leq \sum_{i=1}^n |D_{ii}| = \sum_{i=1}^n \sigma_i(X)$
$= f(X)$

donc : $f(X) \geq \sup \{Tr(M\cancel{D}),\ M^T M = I\}$

cette borne est atteinte pour $M = \operatorname{diag}(\operatorname{sgn}(D_{ii})_i)$

où $\operatorname{sgn}(x) = \mathbb{1}(x \geq 0) - \mathbb{1}(x < 0)$. Ainsi f est convexe comme sup de fonctions linéaires en X.

<u>Ex 4</u>. 1) Montrons que K_{m+} est un cône.

Soit $x, y \in K_{m+}$ et $\theta_1, \theta_2 \geq 0$.

$\forall i \leq j$: $\theta_1 x_i + \theta_2 y_i \geq \theta_1 x_j + \theta_2 y_j$ donc $\theta_1 x + \theta_2 y \in K_{m+}$.

Il est clair que K_{m+} est fermé par caracterisation sequentielle.

Montrons que K_{m+} est pointé.

Soit $x \in K_{m+}$, $x \neq 0$. Alors $-x_1 \leq -x_2 \leq \dots \leq -x_n$ avec une égalité stricte. Donc $-x \notin K_{m+}$.

Montrons que $\mathrm{int}(K_{m+}) \neq \emptyset$.

Soit $x_1 > x_2 > \dots > x_m$. Posons $r = \inf\{x_i - x_{i+1}, 1 \leq i \leq n-1\}$

Soit $\varepsilon \in \mathbb{R}^n$ tq $\|\varepsilon\|_2 < r/2$. On a donc $x + \varepsilon \in B(x, \frac{r}{2})$

et $\forall i$: $x_i - x_{i+1} \geq r \geq \varepsilon_{i+1} - \varepsilon_i$

d'où $x_i + \varepsilon_i \geq \cancel{x}_{i+1} + \varepsilon_{i+1}$ donc $x + \varepsilon \in K_{m+}$.

d'où $B(x, \frac{r}{2}) \subseteq K_{m+}$. donc $x \in \mathrm{int}(K_{m+})$

et $\mathrm{int}(K_{m+}) \neq \emptyset$.

K_{m+} est un cône propre.

2) Soit $(x, y) \in K_{m+} \times K_{m+}^*$. Posons $x_{n+1} = 0$.

$$x^T y = \sum_{i=1}^{n} (x_i - x_{i+1}) \sum_{k=1}^{i} y_k.$$

En choisissant $e_i = (1 \dots \underset{\underset{i}{\uparrow}}{1}\ 0 \dots 0)^T$

on obtient pour $i \in [n]$: $\sum_{k=1}^{i} y_k \geq 0$.

donc $K_{m+}^* \subseteq \{ y \in \mathbb{R}^n \mid \forall i \in [n] : \sum_{k=1}^{i} y_k \geq 0 \}$

L'inclusion réciproque se vérifie facilement.

D'où $K_{m+}^* = \{ y \in \mathbb{R}^n \mid \forall i \in [n] : \sum_{k=1}^{i} y_k \geq 0 \}$.