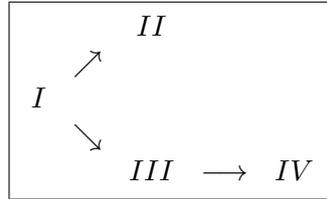


## X-ENS 2019 : épreuve C

L'objet de ce problème est la minimisation sur un sous-domaine  $K \subset \mathbb{R}^n$  d'une fonction  $f$  définie sur  $\mathbb{R}^n$ . Nous proposons plusieurs approches pour trouver des conditions nécessaires d'optimalité, et obtenir des approximations des minimiseurs de  $f$  dans des cas particuliers. Les dépendances entre les parties du problème sont données par le schéma suivant :



### Notations et Définitions

Pour tout entier  $k \in \mathbb{N}^*$ , on désignera le produit scalaire usuel sur  $\mathbb{R}^k$  par  $\langle \cdot, \cdot \rangle$ , et la norme euclidienne sur  $\mathbb{R}^k$  par  $\|\cdot\|$ .

Dans tout le sujet, on se place sur  $\mathbb{R}^n$ , où  $n \in \mathbb{N}^*$ .

— On dit qu'une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est **convexe** si pour tous  $x, y \in \mathbb{R}^n$  et tout  $\lambda \in [0, 1]$ ,

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$

— On dit qu'une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est **coercive** si  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ , autrement dit :

$$\forall M > 0, \exists R > 0, \forall x \in \mathbb{R}^n, \|x\| > R \Rightarrow f(x) > M$$

## 1 Préliminaires

### Fonctions convexes

1. On considère une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

(a) Pour tous  $x, y \in \mathbb{R}^n$ , soit  $\varphi_{x,y} : \mathbb{R} \rightarrow \mathbb{R}$  la fonction définie par

$$\forall t \in \mathbb{R}, \varphi_{x,y}(t) = f(x + t(y-x))$$

Montrer que  $f$  est convexe si et seulement si pour tous  $x, y \in \mathbb{R}^n$ ,  $\varphi_{x,y}$  est convexe.

(b) On suppose que  $f$  est différentiable sur  $\mathbb{R}^n$ . Montrer que pour tous  $x, y \in \mathbb{R}^n$ , la fonction  $\varphi_{x,y}$  est dérivable, et montrer que

$$\forall t, \varphi'_{x,y}(t) = \langle \nabla f(x + t(y-x)), y-x \rangle$$

(c) En déduire que si  $f$  est différentiable sur  $\mathbb{R}^n$ , alors  $f$  est convexe si et seulement si pour tous  $x, y \in \mathbb{R}^n$ ,

$$f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle$$

(d) Montrer que si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est différentiable sur  $\mathbb{R}^n$ , alors  $f$  est convexe si et seulement si pour tous  $x, y \in \mathbb{R}^n$

$$\langle \nabla f(y) - \nabla f(x), y-x \rangle \geq 0$$

2. Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction convexe et différentiable sur  $\mathbb{R}^n$ , et  $x^* \in \mathbb{R}^n$ . Montrer que si  $\nabla f(x^*) = 0$  alors  $f$  admet un minimum global en  $x^*$ .

**Définition.** Soit  $\alpha \in \mathbb{R}_+^*$ . On dit que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  différentiable sur  $\mathbb{R}^n$  est  $\alpha$ -convexe si pour tous  $x, y \in \mathbb{R}^n$ ,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2} \|y - x\|^2$$

3. On considère un réel  $\alpha > 0$  et une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  différentiable sur  $\mathbb{R}^n$ .

(a) On considère la fonction  $g_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$  définie par

$$\forall x \in \mathbb{R}^n, g_\alpha(x) = f(x) - \frac{\alpha}{2} \|x\|^2$$

Calculer  $\nabla g_\alpha(x)$  pour tout  $x \in \mathbb{R}^n$ , et montrer que  $f$  est  $\alpha$ -convexe si et seulement si  $g_\alpha$  est convexe.

(b) En déduire que  $f$  est  $\alpha$ -convexe si et seulement si pour tous  $x, y \in \mathbb{R}^n$ ,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \alpha \|y - x\|^2$$

### Fonctions coercives

4. Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue et coercive. Montrer que si  $K$  est un fermé non vide de  $\mathbb{R}^n$ , alors il existe  $x^* \in K$  tel que  $f(x^*) = \inf_{x \in K} f(x)$ .

5. Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable sur  $\mathbb{R}^n$  et  $\alpha$ -convexe où  $\alpha \in \mathbb{R}_+^*$ . Montrer que si  $K$  est un convexe fermé non vide de  $\mathbb{R}^n$ , alors  $f$  admet un unique minimum sur  $K$ .

### Projection sur un convexe fermé

6. Soient  $C$  un convexe fermé non vide de  $\mathbb{R}^n$  et  $x \in \mathbb{R}^n$ .

(a) Montrer qu'il existe un unique point  $P_C(x) \in C$  tel que  $\|P_C(x) - x\| = \inf_{y \in C} \|y - x\|$ .

(b) Soit  $\bar{x} \in C$ . Montrer que  $\bar{x} = P_C(x)$  si et seulement si

$$\forall y \in C, \langle x - \bar{x}, y - \bar{x} \rangle \leq 0$$

*Indication : on pourra considérer la fonction  $\psi_y : t \in \mathbb{R} \mapsto \|x - (\bar{x} + t(y - \bar{x}))\|^2$  où  $y \in C$ .*

(c) En déduire que si  $x, y \in \mathbb{R}^n$ ,  $\|P_C(y) - P_C(x)\| \leq \|y - x\|$ .

### Une première condition nécessaire d'optimalité

Soit  $K \subset \mathbb{R}^n$ . On dit qu'un vecteur  $h \in \mathbb{R}^n$  est  $K$ -admissible au point  $x \in K$  s'il existe

- une suite  $(t_k)_{k \in \mathbb{N}}$  de réels strictement positifs vérifiant  $\lim_{k \rightarrow +\infty} t_k = 0$ ,
- une suite  $(h_k)_{k \in \mathbb{N}}$  de vecteurs de  $\mathbb{R}^n$  vérifiant  $\lim_{k \rightarrow +\infty} h_k = h$ ,

telles que pour tout  $k \in \mathbb{N}$ ,

$$x + t_k h_k \in K$$

On appelle cône  $K$ -admissible au point  $x \in K$  l'ensemble

$$\mathcal{A}_K(x) := \{h \in \mathbb{R}^n, h \text{ est un vecteur } K \text{ admissible au point } x\}$$

7. Décrire  $\mathcal{A}_K(x)$  dans le cas où  $x$  est dans l'intérieur de  $K$ .

8. Montrer que si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est différentiable en  $x^* \in K$  et admet un minimum local sur  $K$  en  $x^*$ , alors

$$\forall h \in \mathcal{A}_K(x^*), \langle \nabla f(x^*), h \rangle \geq 0$$

Qu'exprime ce résultat dans le cas particulier où  $x^*$  est dans l'intérieur de  $K$  ?

## 2 Pénalisation

Dans le but d'approcher un minimum d'une fonction  $f$  sur  $K \subset \mathbb{R}^n$ , on cherche à se ramener à la minimisation d'une fonction sur  $\mathbb{R}^n$  tout entier. Pour ce faire, on propose d'ajouter à  $f$  un terme de "pénalisation", qui prend de grandes valeurs en dehors de  $K$ , et de minimiser la nouvelle fonction pénalisée sur  $\mathbb{R}^n$  tout entier. Cette partie a pour but de justifier cette approche dans un cas particulier.

Dans toute cette partie, on considère une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  différentiable sur  $\mathbb{R}^n$  et  $\alpha$ -convexe, où  $\alpha \in \mathbb{R}_+^*$ . On pose

$$K = \{x \in \mathbb{R}^n, g_1(x) \leq 0, \dots, g_p(x) \leq 0\}$$

où  $p \in \mathbb{N}^*$  et  $g_1, \dots, g_p$  sont des fonctions convexes de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , différentiables sur  $\mathbb{R}^n$ . On suppose de plus que l'ensemble  $K$  est non vide.

1. Montrer qu'il existe un unique élément  $x^* \in K$  tel que  $f(x^*) = \inf_{x \in K} f(x)$ .

Pour tout  $k \in \mathbb{N}$ , on introduit la fonction  $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$  définie par

$$\forall x \in \mathbb{R}^n, f_k(x) = f(x) + k\Psi(x)$$

où  $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$  est la fonction définie par

$$\forall x \in \mathbb{R}^n, \Psi(x) = \sum_{i=1}^p \max(0, g_i(x))^2$$

2. Pour tout  $x \in \mathbb{R}^n$ , calculer  $\lim_{k \rightarrow +\infty} f_k(x)$ .
3. Montrer que pour tout  $k \in \mathbb{N}$ , il existe un unique  $x_k \in \mathbb{R}^n$  tel que  $f_k(x_k) = \inf_{x \in \mathbb{R}^n} f_k(x)$ .  
*Indication : on pourra commencer par montrer que si  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction convexe, et  $h : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction convexe croissante, alors  $h \circ g$  est convexe.*
4. Montrer que pour tout  $k \in \mathbb{N}$ ,  $f(x_k) \leq f(x^*)$ .
5. On considère une sous-suite  $(x_{\varphi(k)})_{k \in \mathbb{N}}$  de  $(x_k)_{k \in \mathbb{N}}$  qui converge vers  $\bar{x} \in \mathbb{R}^n$ .
  - (a) Montrer que  $\bar{x} \in K$ .
  - (b) En déduire que  $\bar{x} = x^*$ .
6. En déduire que la suite  $(x_k)_{k \in \mathbb{N}}$  converge vers  $x^*$ .
7. Montrer que la suite  $(f_k(x_k))_{k \in \mathbb{N}}$  converge vers  $f(x^*)$ .

## 3 Théorème de Karush-Kuhn-Tucker

Le but de cette partie est d'établir une condition nécessaire d'optimalité dans le cas où le domaine  $K$  est décrit par des contraintes de type inégalité.

### Lemme de Farkas

Soient  $m \in \mathbb{N}^*$  et  $(u_1, \dots, u_m)$  une famille de vecteurs de  $\mathbb{R}^n$ . On note

$$C = \left\{ \sum_{i=1}^m \mu_i u_i, \forall i \in \llbracket 1, m \rrbracket \mu_i \geq 0 \right\}$$

On cherche à démontrer le résultat suivant.

**Lemme 1.** Si  $v \in \mathbb{R}^n$ , alors une et une seule des deux assertions suivantes est vérifiée :

- (i)  $v \in C$ ,

(ii) il existe  $w \in \mathbb{R}^n$  tel que  $\langle v, w \rangle < 0$  et  $\forall i \in \llbracket 1, m \rrbracket, \langle u_i, w \rangle \geq 0$ .

1. Le but de cette question est de montrer que  $C$  est un convexe fermé de  $\mathbb{R}^n$ .

(a) Montrer que  $C$  est convexe.

(b) Montrer que si  $(u_1, \dots, u_m)$  est une famille libre, alors  $C$  est fermé.

(c) Pour tout  $I \subset \llbracket 1, m \rrbracket$ , on pose  $C_I = \{\sum_{i \in I} \mu_i u_i, \forall i \in I \mu_i \geq 0\}$ . Montrer que

$$C = \bigcup_I C_I$$

où l'union est prise sur les ensembles  $I \subset \llbracket 1, m \rrbracket$  tels que  $(u_i)_{i \in I}$  est une famille libre. En déduire que  $C$  est fermé.

2. On considère un vecteur  $v \in \mathbb{R}^n \setminus C$ .

(a) Montrer que  $\langle P_C(v), P_C(v) - v \rangle = 0$ .

(b) On pose  $w = P_C(v) - v$ . Montrer que  $\langle v, w \rangle < 0$  et  $\langle u_i, w \rangle \geq 0$  pour tout  $i \in \llbracket 1, m \rrbracket$ .

3. Conclure la preuve du lemme 1.

### Condition nécessaire d'optimalité

Soit  $p \in \mathbb{N}^*$ . Dans toute la suite de cette partie, on suppose que  $f, g_1, \dots, g_p$  sont des fonctions de  $\mathbb{R}^n$  dans  $\mathbb{R}$  différentiables sur  $\mathbb{R}^n$ , et que

$$K = \{x \in \mathbb{R}^n, g_1(x) \leq 0, \dots, g_p(x) \leq 0\}$$

est non vide. Pour tout  $x \in K$ , on note

$$I_x = \{i \in \llbracket 1, p \rrbracket, g_i(x) = 0\}$$

4. Montrer que pour tout  $x \in K$ ,

$$\mathcal{A}_K(x) \subset \{h \in \mathbb{R}^n, \forall i \in I_x, \langle \nabla g_i(x), h \rangle \leq 0\}$$

5. On considère  $x^* \in K$  et on fait l'hypothèse suivante :

$$\exists v \in \mathbb{R}^n, \forall i \in I_{x^*}, \langle \nabla g_i(x^*), v \rangle < 0 \quad (H)$$

Montrer que  $\mathcal{A}_K(x^*) = \{h \in \mathbb{R}^n, \forall i \in I_{x^*}, \langle \nabla g_i(x^*), h \rangle \leq 0\}$ .

6. Montrer que si  $x^* \in K$  est tel que  $(\nabla g_i(x^*))_{i \in I_{x^*}}$  forme une famille libre, alors l'hypothèse (H) est vérifiée.

7. On suppose que  $f$  atteint en  $x^* \in K$  un minimum local sur  $K$ , et que l'hypothèse (H) est vérifiée. Montrer qu'il existe des réels positifs  $\mu_1^*, \dots, \mu_p^*$  tels que

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^p \mu_i^* \nabla g_i(x^*) = 0, \\ \forall i \in \llbracket 1, p \rrbracket, \mu_i^* g_i(x^*) = 0 \end{cases} \quad (1)$$

8. On suppose dans cette question que les fonctions  $f, g_1, \dots, g_p$  sont convexes. Soient  $x^* \in K$  et  $\mu_1^*, \dots, \mu_p^*$  tels que (1) soit vérifié. Montrer que  $f$  admet en  $x^*$  un minimum global sur  $K$ .

## 4 Etude du problème dual

Le but de cette partie est d'aborder la minimisation d'une fonction  $f$  sur un sous-domaine  $K$  de  $\mathbb{R}^n$  en considérant le problème "dual" associé. Dans un cas particulier, on propose une approche basée sur l'étude du problème dual pour obtenir une approximation du minimum de  $f$  sur  $K$ .

Soit  $p \in \llbracket 1, n \rrbracket$ . On suppose dans toute cette partie que  $f, g_1, \dots, g_p$  sont des fonctions de  $\mathbb{R}^n$  dans  $\mathbb{R}$  différentiables. On fait l'hypothèse supplémentaire que  $f$  est  $\alpha$ -convexe pour un certain  $\alpha \in \mathbb{R}_+^*$ , et que les fonctions  $g_1, \dots, g_p$  sont convexes. On suppose par ailleurs que

$$K = \{x \in \mathbb{R}^n, g_1(x) \leq 0, \dots, g_p(x) \leq 0\}$$

est non vide. Dans toute la suite, on note  $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{pmatrix}$  pour tout  $x \in \mathbb{R}^n$ .

On introduit une fonction  $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}_+^p$  définie par

$$\mathcal{L}(x, \mu) = f(x) + \sum_{i=1}^p \mu_i g_i(x)$$

pour tout  $x \in \mathbb{R}^n$  et tout  $(\mu_1, \dots, \mu_p) \in \mathbb{R}_+^p$ . On s'intéresse au problème : trouver  $x^* \in K$  tel que

$$f(x^*) = \inf_{x \in K} f(x) \quad (P)$$

1. Montrer que

$$\inf_{x \in K} f(x) = \inf_{x \in \mathbb{R}^n} \sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(x, \mu)$$

2. Montrer que pour tout  $\mu \in \mathbb{R}_+^p$ , il existe un unique  $x_\mu \in \mathbb{R}^n$  vérifiant  $\mathcal{L}(x_\mu, \mu) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu)$ .

Pour tout  $\mu \in \mathbb{R}_+^p$ , on note

$$G(\mu) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu) = \mathcal{L}(x_\mu, \mu)$$

On va s'intéresser au problème dit dual : trouver  $\mu^* \in \mathbb{R}_+^p$  tel que

$$G(\mu^*) = \sup_{\mu \in \mathbb{R}_+^p} G(\mu) = \sup_{\mu \in \mathbb{R}_+^p} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu) \quad (Q)$$

On dit que  $(\bar{x}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}_+^p$  est un **point selle de  $\mathcal{L}$**  si

$$\mathcal{L}(\bar{x}, \bar{\mu}) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \bar{\mu}) \quad \text{et} \quad \mathcal{L}(\bar{x}, \bar{\mu}) = \sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(\bar{x}, \mu)$$

3. On suppose dans cette question que  $(\bar{x}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}_+^p$  est un point selle de  $\mathcal{L}$ .

- Montrer que  $\bar{x}$  est solution de (P).
- Montrer que  $\bar{\mu}$  est solution de (Q).
- Montrer que

$$\inf_{x \in \mathbb{R}^n} \sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(x, \mu) = \sup_{\mu \in \mathbb{R}_+^p} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu)$$

4. On considère  $x^* \in K$  une solution de (P) satisfaisant l'hypothèse (H). Soit  $\mu^* = (\mu_1^*, \dots, \mu_p^*)$  comme dans la question 3.7. Montrer que  $\mu^*$  est solution de (Q).

5. On suppose dans toute cette question que la fonction  $\mu \in \mathbb{R}_+^p \mapsto x_\mu$  est continue. On considère une solution  $\bar{\mu} \in \mathbb{R}_+^p$  de (Q).

- (a) Soient  $\mu \in \mathbb{R}_+^p$  et  $\xi \in \mathbb{R}^p$  tels que  $\mu + \xi \in \mathbb{R}_+^p$ . Montrer que pour tout  $t \in [0, 1]$ ,  $\mu + t\xi \in \mathbb{R}_+^p$ , et

$$\lim_{t \rightarrow 0^+} \frac{G(\mu + t\xi) - G(\mu)}{t} = \langle g(x_\mu), \xi \rangle$$

En déduire que pour tout  $\mu \in \mathbb{R}_+^p$ ,  $\langle g(x_{\bar{\mu}}), \mu - \bar{\mu} \rangle \leq 0$ .

- (b) Montrer que  $x_{\bar{\mu}}$  est solution de (P).

6. (Théorème d'Uzawa). Soient  $A \in \mathcal{M}_{p,n}(\mathbb{R})$  une matrice de rang  $p$  et  $b \in \mathbb{R}^p$ . On suppose que la fonction  $g$  est de la forme

$$g : x \mapsto Ax + b$$

- (a) Montrer que pour tout  $\mu \in \mathbb{R}_+^p$ ,  $\nabla f(x_\mu) = -A^T \mu$ , et en déduire que la fonction  $\mu \mapsto x_\mu$  est continue sur  $\mathbb{R}_+^p$ .
- (b) Montrer que (P) admet une unique solution  $x^* \in K$  et que (Q) admet une unique solution  $\mu^* \in \mathbb{R}_+^p$ .

Soit  $\rho > 0$ . On définit la suite  $(\mu^k)_{k \in \mathbb{N}}$  par récurrence de la manière suivante :

- on fixe  $\mu^0 \in \mathbb{R}_+^p$
- pour tout  $k \in \mathbb{N}$ , on pose  $\mu^{k+1} = P_{\mathbb{R}_+^p}(\mu^k + \rho g(x_{\mu^k}))$

où  $P_{\mathbb{R}_+^p} : \mathbb{R}^p \rightarrow \mathbb{R}_+^p$  est la projection sur le convexe fermé  $\mathbb{R}_+^p$  de  $\mathbb{R}^p$ .

- (c) Montrer que  $\mu^* = P_{\mathbb{R}_+^p}(\mu^* + \rho g(x_{\mu^*}))$ .

On suppose désormais que  $\|Ax\| \leq \sqrt{\frac{\alpha}{\rho}} \|x\|$  pour tout  $x \in \mathbb{R}^n$ .

- (d) Montrer que la suite  $(x_{\mu^k})_{k \in \mathbb{N}}$  converge vers  $x^*$ .
- (e) Montrer que la suite  $(\mu^k)_{k \in \mathbb{N}}$  converge vers  $\mu^*$ .