

I - PRÉLIMINAIRES

Fonctions convexes

I.1. (a) \Rightarrow : On suppose que f est convexe. Soit $x, y \in \mathbb{R}^n$. Montrons $\varphi_{x,y}$ est convexe.

Soit $a, b \in \mathbb{R}$. Soit $\lambda \in [0, 1]$. On a :

$$\varphi_{x,y}(\lambda a + (1 - \lambda)b) = f(x + (\lambda a + (1 - \lambda)b)(y - x)) = f[\lambda(x + a(y - x)) + (1 - \lambda)(x + b(y - x))]$$

donc on obtient le résultat attendu

$$\varphi_{x,y}(\lambda a + (1 - \lambda)b) \leq \lambda f(x + a(y - x)) + (1 - \lambda)f(x + b(y - x)) \leq \lambda \varphi_{x,y}(a) + (1 - \lambda)\varphi_{x,y}(b)$$

\Leftarrow : On suppose que pour tous $x, y \in \mathbb{R}^n$, $\varphi_{x,y}$ est convexe. Soit $x, y \in \mathbb{R}^n$ et $\lambda \in [0, 1]$. On a

$$f((1 - \lambda)x + \lambda y) = \varphi_{x,y}(\lambda) = \varphi_{x,y}(\lambda \times 1 + (1 - \lambda) \times 0) \leq \lambda \varphi_{x,y}(1) + (1 - \lambda)\varphi_{x,y}(0) = \lambda f(y) + (1 - \lambda)f(x)$$

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

(b) La fonction $\gamma : t \mapsto x + t(y - x)$ est dérivable sur \mathbb{R} de dérivée $\gamma' : t \mapsto y - x$

or $\varphi_{x,y} = f \circ \gamma$ et f est différentiable. Ainsi d'après le cours, $\varphi_{x,y}$ est dérivable et pour $t \in \mathbb{R}$, on a

$$\varphi'_{x,y}(t) = df(\gamma(t)) \cdot \gamma'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$$

donc pour tous $x, y \in \mathbb{R}^n$, $\varphi_{x,y}$ est dérivable et $\forall t \in \mathbb{R}$, $\varphi'_{x,y}(t) = \langle \nabla f(x + t(y - x)), y - x \rangle$

(c) On suppose que f est différentiable sur \mathbb{R}^n .

\Rightarrow : On suppose que f est convexe. Soit $x, y \in \mathbb{R}^n$. La fonction $\varphi_{x,y}$ est convexe d'après (a) et dérivable d'après (b) : $\forall t \in \mathbb{R}$, $\varphi'_{x,y}(t) = \langle \nabla f(x + t(y - x)), y - x \rangle$

La courbe d'une fonction convexe étant au-dessus de ses tangentes, on a :

$$\forall t \in \mathbb{R}, \varphi'_{x,y}(0)t + \varphi_{x,y}(0) \leq \varphi_{x,y}(t)$$

or $\varphi_{x,y}(0) = f(x)$, $\varphi_{x,y}(1) = f(y)$ et $\varphi'_{x,y}(0) = \langle \nabla f(x), y - x \rangle$ et

Pour conclure il suffit d'appliquer l'inégalité à $t = 1$.

\Leftarrow : On suppose que pour tous $a, b \in \mathbb{R}^n$, $f(b) \geq f(a) + \langle \nabla f(a), b - a \rangle$.

Soit $x, y \in \mathbb{R}^n$. Montrons que $\varphi_{x,y}$ est convexe. Il suffit d'établir que $\varphi'_{x,y}$ est croissante.

Soit $u < v$ dans \mathbb{R} . On a en appliquant l'hypothèse à $a = x + u(y - x)$ et $b = x + v(y - x)$

$$\varphi_{x,y}(u) \geq \varphi_{x,y}(v) + \langle \nabla f(x + u(y - x)), (u - v)(y - x) \rangle$$

ainsi en utilisant (b), on a

$$\frac{\varphi_{x,y}(v) - \varphi_{x,y}(u)}{v - u} \leq \langle \nabla f(x + u(y - x)), y - x \rangle = \varphi'_{x,y}(u)$$

puis avec $a = x + u(y - x)$ et $b = x + v(y - x)$, on trouve

$$\varphi_{x,y}(v) \geq \varphi_{x,y}(u) + \langle \nabla f(x + v(y - x)), (v - u)(y - x) \rangle$$

Finalement, on obtient

$$\varphi'_{x,y}(v) \geq \frac{\varphi_{x,y}(v) - \varphi_{x,y}(u)}{v - u} \geq \varphi'_{x,y}(u)$$

Ce qui permet de conclure.

On a montré que 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) Si f est convexe, alors pour tous $x, y \in \mathbb{R}^n$, $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$ et $f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$

En sommant et par linéarité on obtient $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$

Réciproquement on suppose que pour tout $a, b \in \mathbb{R}^n$, $\langle \nabla f(b) - \nabla f(a), b - a \rangle \geq 0$.

Soit $x, y \in \mathbb{R}^n$. Soit $u < v$ dans \mathbb{R} . On a d'après (b) :

$$\varphi'_{x,y}(v) - \varphi'_{x,y}(u) = \langle \nabla f(x + v(y - x)) - \nabla f(x + u(y - x)), y - x \rangle$$

ainsi

$$\varphi'_{x,y}(v) - \varphi'_{x,y}(u) = \frac{\langle \nabla f(x + v(y - x)) - \nabla f(x + u(y - x)), (v - u)(y - x) \rangle}{v - u} \geq 0$$

On peut conclure à la convexité de $\varphi_{x,y}$ comme à la question précédente.

d'où f est convexe si et seulement si pour tous $x, y \in \mathbb{R}^n$, on a $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$

I.2. On suppose que $\nabla f(x^*) = 0$.

Comme f est convexe et différentiable, alors (b) donne : $\forall y \in \mathbb{R}^n$, $f(y) \geq f(x^*) + \langle 0, y - x^* \rangle = f(x^*)$

si $\nabla f(x^*) = 0$ alors f admet un minimum global en x^*

I.3. (a) On remarque que l'application $Q : x = (x_1, \dots, x_n) \mapsto \|x\| = \sum_{i=1}^n x_i^2$ est de classe \mathcal{C}^∞ donc différentiable

et que pour tout $i \in \llbracket 1, n \rrbracket$, on a $\frac{\partial Q}{\partial x_i} : x \mapsto 2x_i$.

Ainsi $\nabla Q = 2\text{Id}$.

On aurait pu aussi remarquer que $\langle \cdot, \cdot \rangle$ est bilinéaire en dimension finie,

on aurait trouvé que $\forall x \in \mathbb{R}^n$, $\forall h \in \mathbb{R}^n$, $dQ(x) \cdot h = 2\langle x, h \rangle$.

Ainsi g_α est différentiable et $\nabla g_\alpha(x) = \nabla f(x) - \alpha x$ pour tout $x \in \mathbb{R}^n$

Soit $x, y \in \mathbb{R}^n$. On a

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

or $\|y\|^2 - \|x\|^2 - 2\langle x, y - x \rangle = \|y\|^2 + \|x\|^2 - 2\langle x, y \rangle = \|y - x\|^2$. Ainsi

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

On conclut à l'aide de 1(c) appliqué à g_α que f est α -convexe si et seulement si g_α est convexe

(b) Soit $x, y \in \mathbb{R}^n$. On a :

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

Ainsi $\langle \nabla g_\alpha(y) - \nabla g_\alpha(x), y - x \rangle \geq 0 \iff \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \alpha \|y - x\|^2$

À l'aide de 1(d) appliqué à g_α et de la question précédente,

on déduit que f est α -convexe si et seulement si $\forall x, y \in \mathbb{R}^n$, $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \alpha \|y - x\|^2$

Fonctions coercives

I.4. On suppose K est un fermé non vide de \mathbb{R}^n . Prenons $a \in K$ et notons $M = |f(a)|$.

Comme f est coercive, cela nous fournit $R > 0$, tel que $\forall x \in \mathbb{R}^n$, $\|x\| > R \Rightarrow f(x) > M$.

Je note $C = K \cap \mathcal{B}(0, R)$ (boule fermée de centre 0 et de rayon R).

Ainsi $\forall x \in K \setminus C$, $f(x) > f(a)$ et donc $a \in C$.

C est fermé par intersection de fermés et C est bornée car $C \subset \mathcal{B}(0, R)$ ainsi C est compact car $\dim(\mathbb{R}^n) < \infty$. Le théorème des bornes atteinte avec la fonction continue sur le compact C , nous fournit $x^* \in C$ tel que $f(x^*) = \min_{x \in C} f(x)$.

En particulier $f(x^*) \leq f(a)$ et donc $\forall x \in K \setminus C, f(x) > f(x^*)$

On a établi $\text{l'existence de } x^* \in K \text{ tel que } f(x^*) = \inf_{x \in K} f(x)$

I.5. On suppose K est un convexe fermé non vide de \mathbb{R}^n . Prenons $a \in K$. Comme f est α -convexe, on a

$$\forall x \in \mathbb{R}^n, f(x) \geq f(a) + \langle \nabla f(a), x - a \rangle + \frac{\alpha}{2} \|x - a\|^2 = \frac{\alpha}{2} \|x\|^2 + \langle x, \alpha a + \nabla f(a) \rangle + f(a) + \frac{\alpha}{2} \|a\|^2 - \langle a, \nabla f(a) \rangle$$

En posant $b = -\|\alpha a + \nabla f(a)\|$ et $c = f(a) + \frac{\alpha}{2} \|a\|^2 - \langle a, \nabla f(a) \rangle$, on obtient avec l'inégalité de Cauchy-Schwarz :

$$f(x) \geq \frac{\alpha}{2} \|x\|^2 - b\|x\| + c$$

Comme $\lim_{t \rightarrow +\infty} \frac{\alpha}{2} t^2 - bt + c = +\infty$, on en déduit que f est coercive et continue car différentiable, d'après la question précédente, f admet un minimum sur K .

On considère $x, y \in K$ points en lesquels f admet un minimum sur K .

Comme $\forall t \in [0, 1], x + t(y - x) \in K$ car K convexe.

La fonction $\varphi_{x,y}$ (introduite en 1) admet sur $[0, 1]$ un minimum en 0 et 1.

Ainsi $\forall t \in]0, 1], \frac{\varphi_{x,y}(t) - \varphi_{x,y}(0)}{t - 0} \geq 0$ d'où $\varphi'_{x,y}(0) \geq 0$ et de même $\varphi'_{x,y}(1) \leq 0$

On rappelle que $\varphi'_{x,y} : t \mapsto \langle \nabla f(x + t(y - x)), y - x \rangle$ donc $\langle \nabla f(x), y - x \rangle \geq 0$ et $\langle \nabla f(y), y - x \rangle \leq 0$ ainsi $\langle \nabla f(y) - \nabla f(x), y - x \rangle \leq 0$

d'où d'après (b), on a $\alpha \|y - x\|^2 \leq 0$ donc $x = y$

d'où f admet un unique minimum sur K

Projection sur un convexe fermé

I.6. (a) On considère l'application $f : z \mapsto \|z - x\|^2$.

f est différentiable sur \mathbb{R}^n et $\forall z \in \mathbb{R}^n, \nabla f(z) = 2(z - x)$ (calcul analogues aux précédents)

Soit $u, v \in \mathbb{R}^n$. On a

$$f(v) - f(u) - \langle \nabla f(u), v - u \rangle = \|v - x\|^2 - \|u - x\|^2 - \langle 2(u - x), v - u \rangle = \|v\|^2 - \|u\|^2 - 2\langle u, v - u \rangle$$

donc $f(v) - f(u) - \langle \nabla f(u), v - u \rangle = \|v - u\|^2$. Ainsi

$$f(v) = f(u) + \langle \nabla f(u), v - u \rangle + \|v - u\|^2 \geq f(u) + \langle \nabla f(u), v - u \rangle + \frac{2}{2} \|v - u\|^2$$

donc f est 2-convexe. On peut donc appliquer 5.

Ainsi f admet un unique minimum sur C .

Ceci nous fournit $a \in C$, tel que $\forall z \in C \setminus \{a\}, \|z - x\|^2 > \|a - x\|^2$

donc $\forall z \in C \setminus \{a\}, \|z - x\| > \|a - x\|$

Ainsi $P_C(x) = a$ est l'unique point de C tel que $\|P_C(x) - x\| = \inf_{y \in C} \|y - x\|$

(b) Soit $y \in C$. On considère la fonction $\psi_y : t \in \mathbb{R} \mapsto \|x - (\bar{x} + t(y - \bar{x}))\|^2$. Soit $t \in \mathbb{R}$. On a :

$$\psi_y(t) = t^2 \|y - \bar{x}\|^2 - 2t \langle x - \bar{x}, y - \bar{x} \rangle + \|x - \bar{x}\|^2$$

Comme C est convexe, alors $\forall t \in [0, 1], \bar{x} + t(y - \bar{x}) \in C$.

Si $\bar{x} = P_C(x)$, alors $\forall t \in [0, 1], \psi_y(t) \geq \psi_y(0)$

donc $\psi'_y(0) \geq 0$ comme en 5 puis $-2 \langle x - \bar{x}, y - \bar{x} \rangle \geq 0$

et donc $\langle x - \bar{x}, y - \bar{x} \rangle \leq 0$.

Réciproquement on suppose que $\forall y \in C, \langle x - \bar{x}, y - \bar{x} \rangle \leq 0$.

Prenons $y = P_C(x)$, alors la fonction ψ_y admet un minimum en 1.

donc $\psi'_y(1) = 0$ d'où $\|y - \bar{x}\|^2 = \langle x - \bar{x}, y - \bar{x} \rangle \leq 0$

donc $y - \bar{x} = 0$ et ainsi $\bar{x} = y = P_C(x)$.

On a montré que $\boxed{\bar{x} = P_C(x) \text{ si et seulement si } \forall y \in C, \langle x - \bar{x}, y - \bar{x} \rangle \leq 0}$

(c) Soit $x, y \in \mathbb{R}^n$. On a $P_C(x)$ et $P_C(y) \in C$ donc d'après (b), on a

$$\langle x - P_C(x), P_C(y) - P_C(x) \rangle \leq 0 \quad \text{et} \quad \langle y - P_C(y), P_C(x) - P_C(y) \rangle \leq 0$$

ainsi $\langle (x - y) - (P_C(x) - P_C(y)), P_C(x) - P_C(y) \rangle \geq 0$ d'où en utilisant Cauchy-Schwarz, on a

$$\|P_C(y) - P_C(x)\|^2 = \langle P_C(x) - P_C(y), P_C(x) - P_C(y) \rangle \leq \langle (x - y), P_C(x) - P_C(y) \rangle \leq \|y - x\| \cdot \|P_C(y) - P_C(x)\|$$

Comme $\|P_C(y) - P_C(x)\| = 0$ ou $\|P_C(y) - P_C(x)\| > 0$, on a bien $\boxed{\|P_C(y) - P_C(x)\| \leq \|y - x\|}$

Une première condition nécessaire d'optimalité

I.7. On suppose que x est dans l'intérieur de K , ceci nous fournit $r > 0$ tel que $\mathcal{B}(x, r) \subset K$.

Soit $h \in \mathbb{R}^n$. Pour $k \in \mathbb{N}$, je pose $h_k = h$ et $t_k = \frac{r}{(k+1)\|h\|}$ si $h \neq 0$ et $t_k = \frac{1}{(k+1)}$ sinon.

On a bien $\lim_{k \rightarrow +\infty} t_k = 0$ et $\lim_{k \rightarrow +\infty} h_k = h$ et les $t_k > 0$.

De plus $\forall k \in \mathbb{N}, x + t_k h_k \in K$ car $\mathcal{B}(x, r) \subset K$

$\boxed{\text{D'où } \mathcal{A}_K(x) = \mathbb{R}^n \text{ dans le cas où } x \text{ est dans l'intérieur de } K}$

I.8. On suppose que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable en $x^* \in K$ et admet un minimum local sur K en x^* .

Ceci nous fournit un voisinage de x^* tel que $\forall x \in V \cap K, f(x) \geq f(x^*)$.

Soit $h \in \mathcal{A}_K(x^*)$. On considère alors (t_k) et (h_k) de la K -admissibilité de h .

Par théorème d'opération sur les limites, on a $\lim_{k \rightarrow +\infty} x^* + t_k h_k = x^* + 0h = x^*$

Ce qui nous fournit $p \in \mathbb{N}^*$ tel que $\forall k \geq p, x^* + t_k h_k \in V$.

Soit $k \geq p$. On a alors $f(x^* + t_k h_k) \geq f(x^*)$

Or quand $k \rightarrow +\infty$, on a $t_k h_k \rightarrow 0$ et donc

$$f(x^* + t_k h_k) = f(x^*) + \langle \nabla f(x^*), t_k h_k \rangle + o(\|t_k h_k\|)$$

donc

$$f(x^* + t_k h_k) - f(x^*) = t_k (\langle \nabla f(x^*), h_k \rangle + \|h_k\| \varepsilon_k) \quad \text{avec } \varepsilon_k \rightarrow 0$$

On remarque que $\langle \nabla f(x^*), h_k \rangle \rightarrow \langle \nabla f(x^*), h \rangle$ par continuité du produit scalaire

Si $\langle \nabla f(x^*), h \rangle = 0$, on a bien $\langle \nabla f(x^*), h \rangle \geq 0$.

Sinon on a $f(x^* + t_k h_k) - f(x^*) \underset{k \rightarrow +\infty}{\sim} t_k \langle \nabla f(x^*), h \rangle$ donc $\lim_{k \rightarrow +\infty} \frac{f(x^* + t_k h_k) - f(x^*)}{t_k} = \langle \nabla f(x^*), h \rangle$

Par prolongement des inégalités, on a $\forall h \in \mathcal{A}_K(x^*), \langle \nabla f(x^*), h \rangle \geq 0$

Dans le cas où x^* est dans l'intérieur de K alors $\forall h \in \mathbb{R}^n, \langle \nabla f(x^*), h \rangle \geq 0$ selon 7.

En particulier pour $h = -\nabla f(x^*)$, cela donne $\nabla f(x^*) = 0$ pour x^* est dans l'intérieur de K

On reconnaît la condition nécessaire d'extremum pour une fonction différentiable sur un ouvert.

II - PÉNALISATION

II.1. La fonction f est différentiable et α -convexe sur \mathbb{R}^n . K est non vide.

Pour conclure en utilisant **I.5.**, il suffit de montrer que K est fermé et convexe. On a

$$K = \bigcap_{i=1}^p g_i^{-1}]-\infty, 0]$$

Pour chaque i , g_i est continue sur \mathbb{R}^n (car différentiable), donc $g_i^{-1}]-\infty, 0]$ est fermé en tant que image réciproque d'un fermé par une application continue.

Ainsi K est une partie fermée de \mathbb{R}^n en tant qu'intersection de fermés.

Soit $x, y \in K$. Soit $\lambda \in [0, 1]$. On a

$$\forall i \in [1, p], g_i(\lambda x + (1 - \lambda)y) \leq \lambda g_i(x) + (1 - \lambda)g_i(y) \leq 0$$

donc $\lambda x + (1 - \lambda)y \in K$. Ainsi K est bien convexe.

Ainsi $\boxed{\text{il existe un unique élément } x^* \in K \text{ tel que } f(x^*) = \inf_{x \in K} f(x)}$

II.2. Pour tout $x \in \mathbb{R}^n$, on a $\psi(x) \geq 0$ et l'équivalence $x \in K \iff \psi(x) = 0$.

On en déduit que $\lim_{k \rightarrow +\infty} f_k(x) = \begin{cases} f(x) & \text{si } x \in K \\ +\infty & \text{sinon} \end{cases}$

II.3. Soit $k \in \mathbb{N}$. Comme \mathbb{R}^n est non vide, fermé et convexe, pour conclure en utilisant **I.5.**, il suffit de montrer que f_k est différentiable et α -convexe sur \mathbb{R}^n .

étape 1 : On va montrer que f_k est différentiable.

On note la fonction $\varphi : t \mapsto \max(0, t)^2$. de sorte que $\forall t \in \mathbb{R}, \varphi(t) = \left(\frac{t + |t|}{2}\right)^2 = \frac{t^2 + t|t|}{2}$

φ est dérivable sur \mathbb{R}^* : $\forall t > 0, \varphi'(t) = 2t$ et $\forall t < 0, \varphi'(t) = 0$ de plus $\lim_{t \rightarrow 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t \rightarrow 0} \frac{t + |t|}{2} = 0$

donc φ est dérivable sur \mathbb{R} donc y est différentiable.

Ainsi par composition, pour tout $i \in [1, n]$, $x \mapsto \max(0, g_i(x))^2$ est différentiable sur \mathbb{R}^n .

Par combinaison linéaire, f_k est alors différentiable sur \mathbb{R}^n .

étape 2 : On va montrer le résultat de l'indication.

On considère $g : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et $h : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction convexe croissante.

Soit $\lambda \in [0, 1]$ et $x, y \in \mathbb{R}^n$. On a : $g((1 - \lambda)x + \lambda y) \leq (1 - \lambda)g(x) + \lambda g(y)$.

Comme h est croissante puis convexe :

$$h \circ g((1 - \lambda)x + \lambda y) \leq h((1 - \lambda)g(x) + \lambda g(y)) \leq (1 - \lambda)h(g(x)) + \lambda h(g(y)) \leq (1 - \lambda)h \circ g(x) + \lambda h \circ g(y)$$

ainsi $h \circ g$ est bien convexe.

étape 3 : Soit $i \in \llbracket 1, n \rrbracket$. On va montrer que $\varphi \circ g_i : x \mapsto \max(0, g_i(x))^2$ est convexe.

Il suffit d'établir que φ est convexe et croissante pour utiliser l'étape 2.

Ceci est vrai car d'après l'étape 1 il est clair que φ' est positive et croissante sur \mathbb{R} .

étape 4 : On va montrer que f_k est convexe.

En utilisant I.3.(a), il suffit de montrer que $x \mapsto f_k(x) - \frac{\alpha}{2}\|x\|^2$ est convexe.

Toujours d'après I.3.(a), on a $x \mapsto f(x) - \frac{\alpha}{2}\|x\|^2$ est convexe.

Par définition des fonctions convexes, on voit qu'une combinaison linéaire à coefficients positifs de fonctions convexes est convexe.

On peut conclure que il existe un unique $x_k \in \mathbb{R}^n$ tel que $f_k(x_k) = \inf_{x \in \mathbb{R}^n} f_k(x)$

II.4. Soit $k \in \mathbb{N}$. Comme $x^* \in K$ et $f \leq f_k$ et par définition de x_k , on a $f(x_k) \leq f_k(x_k) \leq f_k(x^*) = f(x^*)$

II.5. (a) Par l'absurde si $\bar{x} \notin K$ alors $\psi(\bar{x}) > 0$.

Alors $\lim_{k \rightarrow +\infty} \psi(x_{\varphi(k)}) = \psi(\bar{x})$ car ψ est continue car différentiable sur \mathbb{R}^n .

or d'après la question précédente, on a

$$\forall k \in \mathbb{N}, f(x_{\varphi(k)}) + k\psi(x_{\varphi(k)}) = f_{\varphi(k)}(x_{\varphi(k)}) \leq f_k(x^*) = f(x^*)$$

Donc la suite $(f_{\varphi(k)}(x_{\varphi(k)}))_{k \in \mathbb{N}}$ diverge vers $+\infty$ et est majorée

Ceci est absurde. On a montré que $\bar{x} \in K$

(b) f est continue donc en passant à la limite de l'inégalité de 5, on a $f(\bar{x}) \leq f(x^*)$.

Or $\bar{x} \in K$ selon (a), donc d'après l'unicité de 1, on a bien $\bar{x} = x^*$

II.6. On suppose par l'absurde que la suite $(x_k)_{k \in \mathbb{N}}$ n'est pas bornée.

On construit $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ par $\varphi(0) = 0$ et la relation de récurrence :

$$\forall k \in \mathbb{N}, \varphi(k+1) = \min \{p > \varphi(k) \mid \|x_p\| \geq k+1\}$$

φ est bien définie car toute partie non vide de \mathbb{N} admet un minimum.

φ est une extractrice car strictement croissante.

De plus $\forall k \in \mathbb{N}, \|x_{\varphi(k)}\| \geq k$ d'où $\lim_{k \rightarrow +\infty} \|x_{\varphi(k)}\| = +\infty$.

or f est coercive car f est différentiable et α -convexe en utilisant I.5

donc $\lim_{k \rightarrow +\infty} f(x_{\varphi(k)}) = +\infty$ ce qui est absurde d'après 4.

Ainsi la suite $(x_k)_{k \in \mathbb{N}}$ est bornée dans l'espace \mathbb{R}^n qui est de dimension finie.

De plus, cette suite admet x^* comme unique valeur d'adhérence.

On en déduit que la suite $(x_k)_{k \in \mathbb{N}}$ converge vers x^*

II.7. Soit $k \in \mathbb{N}$. On a $f(x_k) + k\psi(x_k) = f_k(x_k) \leq f(x^*)$ (majoration vue en II.4) donc

$$0 \leq k\psi(x_k) \leq f(x^*) - f(x_k)$$

Or par continuité de f et selon la question précédente $f(x^*) - f(x_k) \xrightarrow[k \rightarrow +\infty]{} 0$

Par théorème des gendarmes, on a alors $k\psi(x_k) \xrightarrow[k \rightarrow +\infty]{} 0$

Ainsi la suite $(f_k(x_k))_{k \in \mathbb{N}}$ converge vers $f(x^*)$

III - THÉORÈME DE KARUSH-KUHN-TUCKER

III.1. (a) Soit $x, y \in C$ et $\lambda \in [0, 1]$.

On peut écrire $x = \sum_{i=1}^m \mu_i u_i$ et $y = \sum_{i=1}^m \mu'_i u_i$ où les $\mu_i \geq 0$ et les $\mu'_i \geq 0$.

Alors $\lambda x + (1 - \lambda)y \in C$ car les $\lambda\mu_i + (1 - \lambda)\mu'_i \geq 0$.

On a montré que C est convexe

(b) On suppose que (u_1, \dots, u_m) est une famille libre.

Je note $F = \text{Vect}(u_1, \dots, u_m)$ de sorte que (u_1, \dots, u_m) est une base de F .

Pour $i \in \llbracket 1, m \rrbracket$, je note γ_i l'application qui à chaque vecteur de F associe sa i -ème coordonnée dans cette base. Comme γ_i est linéaire et que $\dim(F) < +\infty$, alors γ_i est continue.

De plus $C \subset F$ et $C = \bigcap_{i=1}^m \gamma_i^{-1}(\mathbb{R}^+)$.

or les $\gamma_i^{-1}(\mathbb{R}^+)$ sont des fermés de F en tant qu'images réciproques de fermés par des applications continues

De plus F est un fermé de E en tant que sous-espace vectoriel de dimension finie

ainsi les $\gamma_i^{-1}(\mathbb{R}^+)$ sont des fermés de E

ainsi par intersection C est fermé

(c) On considère qu'une famille indexée par l'ensemble vide est libre.

On est obligé d'envisager ce cas, lorsque $\forall i \in I, u_i = 0$.

Je note $\Lambda = \{I \in \mathcal{P}(\llbracket 1, m \rrbracket) \mid (u_i)_{i \in I} \text{ est une famille libre}\}$.

On remarque que $\forall I \in \mathcal{P}(\llbracket 1, m \rrbracket), C_I \subset C$ (quitte à rajouter des 0).

Ainsi $\bigcup_{I \in \Lambda} C_I \subset C$.

Pour montrer l'inclusion réciproque, on va montrer par récurrence sur $k \in \mathbb{N}$, la propriété :

$$\mathcal{H}_k : \forall (\mu_1, \dots, \mu_m) \in (\mathbb{R}^+)^m, |\{i \in \llbracket 1, m \rrbracket \mid \mu_i > 0\}| = k \implies \sum_{i=1}^m \mu_i u_i \in \bigcup_{I \in \Lambda} C_I$$

Initialisation : La propriété \mathcal{H}_0 est vraie car $\sum_{i=1}^m 0u_i = 0 \in C_\emptyset = \{0\}$

Hérédité : Soit $k \in \mathbb{N}$ tel que pour tout $p \in \llbracket 0, k \rrbracket$, \mathcal{H}_p est vraie. Montrons \mathcal{H}_{k+1} .

Soit $(\mu_1, \dots, \mu_m) \in (\mathbb{R}^+)^m$ tel que $|\{i \in \llbracket 1, m \rrbracket \mid \mu_i > 0\}| = k + 1$.

On note $x = \sum_{i=1}^m \mu_i u_i$ et $J = \{i \in \llbracket 1, m \rrbracket \mid \mu_i > 0\}$ de sorte que $x = \sum_{i \in J} \mu_i u_i$.

Il s'agit d'établir que $x \in \bigcup_{I \in \Lambda} C_I$.

Premier cas : Si la famille $(u_i)_{i \in J}$ est libre, alors $x = \sum_{i=1}^m \mu_i u_i \in C_J \subset \bigcup_{I \in \Lambda} C_I$.

Deuxième cas : Si la famille $(u_i)_{i \in J}$ est liée, il existe alors une famille de réels non tous nuls $(\lambda_i)_{i \in J}$ tel que $\sum_{i \in J} \lambda_i u_i = 0$.

Je note $S^+ = \{i \in J \mid \lambda_i > 0\}$ et $S^- = \{i \in J \mid \lambda_i < 0\}$

de sorte que $\sum_{i \in S^+} \lambda_i u_i = \sum_{i \in S^-} (-\lambda_i) u_i$ et $S^+ \cup S^- \neq \emptyset$.

Si $S^+ \neq \emptyset$, je prends $j_0 \in S^+$ tel que $\frac{\mu_{j_0}}{\lambda_{j_0}} = \min \left\{ \frac{\mu_j}{\lambda_j} \mid j \in S^+ \right\}$.

Je pose $S' = S^+ \setminus \{j_0\}$. Alors

$$x = \sum_{i \in S'} \underbrace{\left(\mu_i - \frac{\mu_{j_0}}{\lambda_{j_0}} \lambda_i \right)}_{\geq 0} u_i + \sum_{i \in S^-} \underbrace{\left(\mu_i - \frac{\mu_{j_0}}{\lambda_{j_0}} \lambda_i \right)}_{\geq 0} u_i$$

On peut alors utiliser l'hypothèse de récurrence pour conclure que $x \in \bigcup_{I \in \Lambda} C_I$.

Si $S^+ = \emptyset$, alors on se ramène au cas précédent en remarquant $\sum_{i \in J} (-\lambda_i) u_i = 0$.

Conclusion : On a montré que pour tout $k \in \mathbb{N}$, \mathcal{H}_k est vraie.

En particulier pour tout $k \in \llbracket 0, m \rrbracket$, la propriété \mathcal{H}_k est vraie et ainsi $\bigcup_{I \in \Lambda} C_I \supset C$.

On a bien $C = \bigcup_{I \in \Lambda} C_I$ or d'après (b), pour tout $I \in \Lambda$, C_I est fermé.

On en déduit que C est fermé en tant que réunion finie de fermés en effet $|\Lambda| \leq 2^m < +\infty$.

III.2. (a) Comme C est convexe, fermé et non vide car $0 \in C$, on peut appliquer I.6.(b) :

$$\forall y \in C, \langle v - P_C(v), y - P_C(v) \rangle \leq 0$$

On remarque $\forall \lambda \geq 0, \forall x \in C, \lambda x \in C$

d'où $0 \in C$ et $2P_C(v) \in C$ ce qui permet de conclure $\langle P_C(v), P_C(v) - v \rangle = 0$

(b) À l'aide de I.6(b), et de l'égalité précédente, on a :

$$\forall y \in C, \langle w, y \rangle \geq 0$$

Or pour tout $i \in \llbracket 1, m \rrbracket$, on a $u_i \in C$ et donc $\langle u_i, w \rangle \geq 0$

De plus, à l'aide de (a) :

$$\langle v, w \rangle = \langle v, P_C(v) \rangle - \langle v, v \rangle = \langle P_C(v), P_C(v) \rangle - \langle v, v \rangle = \|P_C(v)\|^2 - \|v\|^2$$

Comme $0 \in C$, on a $P_C(0) = 0$ et ainsi d'après I.6.(c), $\|P_C(v)\| \leq \|v\|$.

Par l'absurde si on avait $\|P_C(v)\| = \|v\|$ alors d'après (a) :

$$\langle P_C(v), v \rangle = \langle P_C(v), P_C(v) \rangle = \langle v, v \rangle$$

donc $\langle P_C(v), v \rangle = \|P_C(v)\| \cdot \|v\|$ ainsi selon le cas d'égalité de l'inégalité de Cauchy-Schwarz, on a l'existence de $\lambda \in \mathbb{R}$ tel que $P_C(v) = \lambda v$ car $v \neq 0$.

On trouve ensuite $\lambda = 1$ donc $v = P_C(v) \in C$ ce qui est absurde.

Ainsi $\|P_C(v)\| < \|v\|$ et $\langle v, w \rangle < 0$

III.3. On vient de montrer que le (i) ou le (ii) du lemme 1 est toujours vrai.

Par l'absurde si il existait $v \in C$ et $w \in \mathbb{R}^n$ tels que $\langle v, w \rangle < 0$ et $\forall i \in \llbracket 1, m \rrbracket, \langle u_i, w \rangle \geq 0$.

On pourrait écrire $v = \sum_{i=1}^m \mu_i u_i$ où $\forall i \in \llbracket 1, m \rrbracket, \mu_i \geq 0$. Ainsi

$$0 > \langle v, w \rangle = \sum_{i=1}^m \mu_i \langle v, u_i \rangle \geq 0$$

Ce qui est absurde. Ceci prouve le lemme 1

Condition nécessaire d'optimalité

III.4. Soit $x \in K$, $h \in \mathcal{A}_K(x)$ et $i \in I_x$. On a

$$\forall y \in K, -g_i(x) = 0 \leq -g_i(y)$$

Donc $-g_i$ différentiable sur \mathbb{R}^n présente un minimum local sur K en x .

Ainsi $\langle \nabla(-g_i)(x), h \rangle \geq 0$ en utilisant I.8.

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

III.5. Soit $h \in \mathbb{R}^n$ tel que $\forall i \in I_{x^*}$, $\langle \nabla g_i(x^*), h \rangle \leq 0$. Selon la question précédente, il suffit d'établir que $h \in \mathcal{A}_K(x^*)$.

L'hypothèse (H) nous fournit $v \in \mathbb{R}^n$ tel que $\forall i \in I_{x^*}$, $\langle \nabla g_i(x^*), v \rangle < 0$.

Soit $\alpha > 0$. On a $\forall i \in I_{x^*}$, $\langle \nabla g_i(x^*), \alpha v \rangle < 0$

Soit $i \in \llbracket 1, p \rrbracket$. On a pose

$$G_{i,\alpha} : t \mapsto g_i(x^* + t\alpha v + th)$$

Si $i \notin I_{x^*}$, la continuité de $G_{i,\alpha}$ et $G_{i,\alpha}(0) = g_i(0) < 0$, nous fournit

$\rho_i(\alpha) > 0$ tel que $\forall t \in [0, \rho_i(\alpha)]$, $g_i(x^* + t\alpha v + th) < 0$

Si $i \in I_{x^*}$ alors $G_{i,\alpha}$ s'annule en 0 et a comme dérivée $G'_{i,\alpha} : t \mapsto \langle \nabla g_i(x^* + t\alpha v + th), \alpha v + h \rangle$.

donc $G'_{i,\alpha}(0) = \alpha \langle \nabla g_i(x^*), v \rangle + \langle \nabla g_i(x^*), h \rangle < 0$ et $G_{i,\alpha}(0) = g_i(0) = 0$

ce qui nous fournit $\rho_i > 0$ tel que $\forall t \in]0, \rho_i(\alpha)]$, $g_i(x^* + t(\alpha v + h)) < 0$

On pose $r(\alpha) = \min \{\rho_i(\alpha) \mid i \in \llbracket 1, p \rrbracket\}$ de sorte que

$$r(\alpha) > 0 \text{ et } \forall i \in \llbracket 1, p \rrbracket, \forall t \in]0, r(\alpha)], g_i(x^* + t\alpha v + th) < 0$$

Pour $k \in \mathbb{N}$, je pose $\alpha_k = \frac{1}{k+1}$ et $t_k = \min\left(\frac{1}{k+1}, r(\alpha_k)\right)$ et $h_k = h + \alpha_k v$.

Ainsi $\forall k \in \mathbb{N}$, $0 < t_k \leq r(\alpha_k)$ et $\lim_{k \rightarrow +\infty} t_k = 0$ et $\lim_{k \rightarrow +\infty} h_k = h$

De plus, $\forall k \in \mathbb{N}$, $x^* + t_k h_k \in K$ car $\forall i \in \llbracket 1, p \rrbracket$, $g_i(x^* + t_k h_k) = g_i(x^* + t_k h + \alpha_k t_k v) < 0$

Ainsi $h \in \mathcal{A}_K(x^*)$

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

III.6. Soit $p \in \mathbb{N}^*$. On considère une famille libre $(u_i)_{1 \leq i \leq p}$.

La méthode d'orthonormalisation de Gram-Schmidt appliquée à la famille libre $(u_i)_{1 \leq i \leq p}$ nous fournit une famille orthonormale $(\varepsilon_i)_{1 \leq i \leq p}$ tel que

$$\begin{cases} \langle u_i, \varepsilon_i \rangle > 0 \\ \forall k \in \llbracket 1, p \rrbracket, \text{Vect}((\varepsilon_i)_{1 \leq i \leq k}) = \text{Vect}((u_i)_{1 \leq i \leq k}) \end{cases}$$

On va montrer par récurrence sur $k \in \llbracket 1, p \rrbracket$ l'existence de $v \in \text{Vect}((u_i)_{1 \leq i \leq k})$ tel que $\forall i \in \llbracket 1, k \rrbracket$, $\langle u_i, v \rangle < 0$.

Pour l'initialisation, je prend $v_1 = -u_1$.

Pour l'hérédité : soit $k \in \llbracket 1, p-1 \rrbracket$ tel qu'il existe $v \in \text{Vect}((u_i)_{1 \leq i \leq k})$ tel que $\forall i \in \llbracket 1, k \rrbracket$, $\langle u_i, v \rangle < 0$.

Je pose $w(t) = v + t\varepsilon_{k+1}$.

On a $\forall i \in \llbracket 1, k \rrbracket$, $\langle u_i, w(t) \rangle = \langle u_i, v \rangle < 0$ et $\langle u_{k+1}, w(t) \rangle = \langle u_{k+1}, v \rangle + t \langle u_{k+1}, \varepsilon_{k+1} \rangle$

En prenant $t = -\frac{1 + \langle u_{k+1}, v \rangle}{\langle u_{k+1}, \varepsilon_{k+1} \rangle}$, on a bien $\forall i \in \llbracket 1, k+1 \rrbracket$, $\langle u_i, w(t) \rangle < 0$

On donc montré que la propriété était vraie pour tout $k \in \llbracket 1, p \rrbracket$.

L'hypothèse (H) est vérifiée si $x^* \in K$ est tel que $(\nabla g_i(x^*))_{i \in I_{x^*}}$ forme une famille libre et $I_{x^*} \neq \emptyset$.

Dans le cas où $I_{x^*} = \emptyset$, n'importe quel vecteur $v \in \mathbb{R}^n$ convient.

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

III.7. Soit $w \in \mathbb{R}^n$ tel que $\forall i \in I_{x^*}, \langle \nabla g_i(x^*), w \rangle \leq 0$.

On a $w \in \mathcal{A}_K(x^*)$ d'après 5.

Ainsi d'après I.8, on a $\langle \nabla f(x^*), w \rangle \geq 0$

On vient de montrer

$$\forall w \in \mathbb{R}^n, (\forall i \in I_{x^*}, \langle \nabla g_i(x^*), w \rangle \leq 0) \implies \langle \nabla f(x^*), w \rangle \geq 0$$

Ce qui est la négation de : $\exists w \in \mathbb{R}^n, \langle \nabla f(x^*), w \rangle < 0$ et $\forall i \in I_{x^*}, \langle -\nabla g_i(x^*), w \rangle \geq 0$

Le lemme 1, nous fournit une famille de réels positifs $(\mu_i^*)_{i \in I_{x^*}}$, telle que $\nabla f(x^*) = \sum_{i \in I_{x^*}} \mu_i^* (-\nabla g_i(x^*))$

En complétant la famille par des zéros,

on trouve des réels positifs μ_1^*, \dots, μ_p^* tels que :

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

III.8. On remarque que $\nabla \left(f + \sum_{i=1}^p \mu_i^* g_i \right) (x^*) = 0$ et que $f + \sum_{i=1}^p \mu_i^* g_i$ est convexe.

Alors selon I.2, $f + \sum_{i=1}^p \mu_i^* g_i$ admet en x^* un minimum global sur \mathbb{R}^n .

Soit $x \in K$. On a : $f(x^*) = f(x^*) + \sum_{i=1}^p \mu_i^* g_i(x^*) \leq f(x) + \sum_{i=1}^p \mu_i^* g_i(x) \leq f(x)$

Ainsi f admet en x^* un minimum global sur K

IV - ÉTUDE DU PROBLÈME DUAL

IV.1. Pour une partie non vide est non majorée (respectivement non minorée), la borne supérieure (respectivement inférieure) est $+\infty$ (respectivement $-\infty$).

Soit $x \in \mathbb{R}^n \setminus K$. Il existe $i_0 \in \llbracket 1, p \rrbracket$ tel que $g_{i_0}(x) > 0$.

Dans ce cas $\sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(x, \mu) = +\infty$.

Par conséquent

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

Soit $x \in K$ et $\mu = (\mu_1, \dots, \mu_p) \in \mathbb{R}_+^p$.

On a : $\mathcal{L}(x, \mu) = f(x) + \sum_{i=1}^p \mu_i g_i(x) \leq f(x) = f(x) + \sum_{i=1}^p 0 g_i(x)$ donc

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

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

IV.2. Soit $\mu \in \mathbb{R}_+^p$. En faisant comme en **II.3.**, on montre que la fonction $x \mapsto \mathcal{L}(x, \mu)$ est α -convexe.

Ainsi en appliquant **II.1.** à cette fonction différentiable et α -convexe sur \mathbb{R}^n (défini avec 0 contrainte),

$$\text{il existe un unique } x_\mu \in \mathbb{R}^n \text{ vérifiant } \mathcal{L}(x_\mu, \mu) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu)$$

IV.3. (a) Il s'agit d'établir que : (i) $\bar{x} \in K$ et (ii) $f(\bar{x}) = \inf_{x \in K} f(x)$.

Pour (i) : on procède par l'absurde en supposant que $\bar{x} \notin K$.

Alors $\sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(\bar{x}, \mu) = +\infty$ comme en **IV.1.**) donc

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

Absurde car une borne inférieure généralisée est dans $[-\infty, +\infty[$.

Soit $y \in K$ pour établir (ii), il suffit de montrer que $f(\bar{x}) \leq f(y)$.

Comme les $g_i(\bar{x}) \leq 0$ car $\bar{x} \in K$, on a

$$\inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \bar{\mu}) = \mathcal{L}(\bar{x}, \bar{\mu}) = \sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(\bar{x}, \mu) = \sup_{\mu \in \mathbb{R}_+^p} \left(f(\bar{x}) + \sum_{i=1}^p \mu_i g_i(\bar{x}) \right) = f(\bar{x})$$

Ainsi $f(\bar{x}) \leq \mathcal{L}(y, \bar{\mu}) = f(y) + \sum_{i=1}^p \bar{\mu}_i g_i(y) \leq f(y)$ car $y \in K \subset \mathbb{R}^n$

donc \bar{x} est bien solution de (P)

(b) Soit $\mu \in \mathbb{R}_+^p$. Il suffit de montrer que $G(\mu) \leq G(\bar{\mu})$. On a

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

donc

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

donc $\bar{\mu}$ est bien solution de (Q)

(c) D'après (a) puis **IV.1.**, on a $\bar{x} \in K$ et

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

et d'après (b), on a : $\mathcal{L}(\bar{x}, \bar{\mu}) = G(\bar{\mu}) = \sup_{\mu \in \mathbb{R}_+^p} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu)$

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

IV.4. On remarque que $\nabla \left(f + \sum_{i=1}^p \mu_i^* g_i \right) (x^*) = 0$ et que $f + \sum_{i=1}^p \mu_i^* g_i$ est α -convexe (analogue à l'étape 4 de II.3).

Par définition des fonctions α -convexes, on a donc

$$\forall y \in \mathbb{R}^n, \left(f + \sum_{i=1}^p \mu_i^* g_i \right) (y) \geq \left(f + \sum_{i=1}^p \mu_i^* g_i \right) (x^*) + \langle 0, y - x^* \rangle + \frac{\alpha}{2} \|y - x^*\|^2 \geq \left(f + \sum_{i=1}^p \mu_i^* g_i \right) (x^*)$$

avec égalité si et seulement si $y = x^*$.

Donc $f + \sum_{i=1}^p \mu_i^* g_i$ admet un unique minimum global en x^* d'où $\mathcal{L}(x^*, \mu^*) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu^*)$ (1)

Soit $\mu = (\mu_1, \dots, \mu_p) \in \mathbb{R}_+^p$. On a selon la deuxième condition de III.8 :

$$\mathcal{L}(x^*, \mu^*) - \mathcal{L}(x^*, \mu) = \sum_{i=1}^p \mu_i^* g_i(x^*) - \sum_{i=1}^p \mu_i g_i(x^*) = - \sum_{i=1}^p \mu_i g_i(x^*) \geq 0$$

car les $g_i(x^*) \leq 0$ car $x^* \in K$.

D'où $\mathcal{L}(x^*, \mu^*) = \sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(x^*, \mu)$ (2)

Avec (1) et (2), on déduit que (x^*, μ^*) est un point selle de \mathcal{L} d'où μ^* est solution de (Q) d'après IV.3.(b)

IV.5. Pour soulager l'écriture je note l'application $X : \mu \mapsto x_\mu = X(\mu)$ qui est bien définie sur \mathbb{R}_+^p et continue par hypothèse.

(a) Soit $t \in [0, 1]$.

Pour $i \in \llbracket 1, p \rrbracket$, les formes linéaires de $\mathbb{R}^p : (u_1, \dots, u_p) \mapsto -u_i$ sont convexes.

Ainsi comme en II.1, \mathbb{R}_+^p est une partie convexe de \mathbb{R}^p .

De plus $\mu + t\xi = t(\mu + \xi) + (1-t)\mu$ or $\mu + \xi \in \mathbb{R}_+^p$ et $\mu \in \mathbb{R}_+^p$ donc $\mu + t\xi \in \mathbb{R}_+^p$

On suppose désormais que $t \in]0, 1]$. On a alors l'existence des membres et l'égalité

$$\frac{G(\mu + t\xi) - G(\mu)}{t} = \frac{\mathcal{L}(X(\mu + t\xi), \mu + t\xi) - \mathcal{L}(X(\mu + t\xi), \mu) + \mathcal{L}(X(\mu + t\xi), \mu) - \mathcal{L}(X(\mu), \mu)}{t}$$

or par continuités de g , X et du produit scalaire :

$$\frac{\mathcal{L}(X(\mu + t\xi), \mu + t\xi) - \mathcal{L}(X(\mu + t\xi), \mu)}{t} = \langle g(X(\mu + t\xi)), \xi \rangle \xrightarrow{t \rightarrow 0} \langle g(X(\mu)), \xi \rangle = \langle g(x_\mu), \xi \rangle$$

et $\frac{\mathcal{L}(X(\mu + t\xi), \mu) - \mathcal{L}(X(\mu), \mu)}{t} \geq 0$ par définition de $X(\mu)$

D'un autre côté :

$$\frac{G(\mu + t\xi) - G(\mu)}{t} = \frac{\mathcal{L}(X(\mu + t\xi), \mu + t\xi) - \mathcal{L}(X(\mu), \mu + t\xi)}{t} + \langle g(x_\mu), \xi \rangle$$

d'où par définition de $X(\mu + t\xi)$, on a

$$\langle g(x_\mu), \xi \rangle \leq \frac{G(\mu + t\xi) - G(\mu)}{t} \leq \langle g(X(\mu + t\xi)), \xi \rangle$$

par théorème des gendarmes, on obtient bien : $\lim_{t \rightarrow 0^+} \frac{G(\mu + t\xi) - G(\mu)}{t} = \langle g(x_\mu), \xi \rangle$

Soit $\mu \in \mathbb{R}_+^p$. Je pose maintenant $\xi = \mu - \bar{\mu}$ de sorte que $\bar{\mu} \in \mathbb{R}_+^p$ et $\bar{\mu} + \xi = \mu \in \mathbb{R}_+^p$.

D'après ce que l'on vient de voir la fonction $H : t \mapsto G(\bar{\mu} + t\xi)$ est dérivable en 0

et $H'(0) = \langle g(x_{\bar{\mu}}), \xi \rangle = \langle g(x_{\bar{\mu}}), \mu - \bar{\mu} \rangle$

or cette fonction admet un maximum en 0 car $\bar{\mu} \in \mathbb{R}_+^p$ est solution de (Q)

d'où $H'(0) \leq 0$ c'est à dire $\langle g(x_{\bar{\mu}}), \mu - \bar{\mu} \rangle \leq 0$

(b) Par définition de $x_{\bar{\mu}}$, on a $\mathcal{L}(x_{\bar{\mu}}, \bar{\mu}) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \bar{\mu})$ (i)

Soit $\mu \in \mathbb{R}_+^p$. D'après (a), on a

$$\mathcal{L}(x_{\bar{\mu}}, \bar{\mu}) - \mathcal{L}(x_{\bar{\mu}}, \mu) = \langle g(x_{\bar{\mu}}), \bar{\mu} - \mu \rangle \geq 0$$

Ainsi $\mathcal{L}(x_{\bar{\mu}}, \bar{\mu}) = \sup_{\mu \in \mathbb{R}_+^p} \mathcal{L}(x_{\bar{\mu}}, \mu)$ (ii)

Avec (i) et (ii), $(x_{\bar{\mu}}, \bar{\mu})$ est un point selle pour \mathcal{L} .

Alors selon IV.3.(a) $x_{\bar{\mu}}$ est solution de (P)

IV.6. (a) On remarque au préalable que comme $g : x \mapsto Ax + b$ alors pour $i \in \llbracket 1, p \rrbracket$, on a

$$g_i : x \mapsto \langle L_i^\top, x \rangle + b_i$$

où L_i est la i -ème ligne de A et b_i est le i -ème coefficient de b .

Ainsi les g_i sont bien convexes et différentiables sur \mathbb{R}^n et $\forall x \in \mathbb{R}^n$, $\nabla g_i(x) = L_i^\top$.

De plus, comme A est de rang p , il existe $x \in \mathbb{R}^n$ tel que $g(x) = 0$ et donc K est non vide on peut donc utiliser les résultats de la partie IV.

Soit $\mu \in \mathbb{R}_+^p$. On pose alors $H : x \in \mathbb{R}^n \mapsto \mathcal{L}(x, \mu)$. Ainsi

$$\forall x \in \mathbb{R}^n, H(x) = f(x) + \langle g(x), \mu \rangle = f(x) + \langle Ax, \mu \rangle + \langle b, \mu \rangle = f(x) + \langle x, A^\top \mu \rangle + \langle b, \mu \rangle$$

Ainsi H est différentiable et $\forall x \in \mathbb{R}^n$, $\nabla H(x) = \nabla f(x) + A^\top \mu$

Or la fonction H admet un minimum sur l'ouvert \mathbb{R}^n en x_μ donc $\nabla H(x_\mu) = 0$

ainsi $\boxed{\text{pour tout } \mu \in \mathbb{R}_+^p, \nabla f(x_\mu) = -A^\top \mu}$

Soit $\lambda \in \mathbb{R}_+^p$. On a alors selon I.3.(b) :

$$\langle A^\top \lambda - A^\top \mu, x_\lambda - x_\mu \rangle = \langle \nabla f(x_\lambda) - \nabla f(x_\mu), x_\lambda - x_\mu \rangle \geq \alpha \|x_\lambda - x_\mu\|^2$$

À l'aide de Cauchy-Schwarz et comme c'est valable si $x_\lambda = x_\mu$, on a alors

$$\|x_\lambda - x_\mu\| \leq \frac{\|A^\top (\lambda - \mu)\|}{\alpha}$$

Comme l'application $x \mapsto Ax$ est continue, on en conclut que $\lim_{\lambda \rightarrow \mu} x_\lambda = x_\mu$

Ainsi $\boxed{\text{la fonction } \mu \mapsto x_\mu \text{ est continue sur } \mathbb{R}_+^p}$

(b) Les hypothèses de II.1 étant vérifiées : $\boxed{(P) \text{ admet une unique solution } x^* \in K}$ De plus

$$\text{rg} \left((\nabla g_i(x^*))_{1 \leq i \leq p} \right) = \text{rg} \left(\left(L_i^\top \right)_{1 \leq i \leq p} \right) = \text{rg}(A) = p$$

donc la famille $(\nabla g_i(x^*))_{1 \leq i \leq p}$ est libre et il en est donc de même pour $(\nabla g_i(x^*))_{i \in I_{x^*}}$

Ainsi l'hypothèse (H) est vérifiée selon III.6

On peut donc choisir $\mu^* \in \mathbb{R}_+^p$ vérifiant les conditions de III.7 selon III.7

Selon IV.4, μ^* est alors solution de (Q) et on a aussi $x_{\mu^*} = x^*$.

Alors selon (a), on a : $-\nabla f(x^*) = A^\top \mu^*$

L'application de $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ $\tau : \mu \mapsto A^\top \mu$ est de rang $\text{rg}(A^\top) = \text{rg}(A) = p$

Selon le théorème du rang, on a donc $\dim(\text{Ker } \tau) = p - p = 0$

τ est donc injective ainsi μ^* est l'unique antécédent de $-\nabla f(x^*)$ par τ (l'unicité de x^* a été établie)

d'où $\boxed{(Q) \text{ admet une unique solution } \mu^* \in \mathbb{R}_+^p}$

(c) En utilisant I.6.(b), il suffit d'établir que

$$\forall y \in \mathbb{R}_+^p, \langle \mu^* + \rho g(x_{\mu^*}) - \mu^*, y - \mu^* \rangle \leq 0$$

comme $\rho > 0$ ceci équivaut à

$$\forall \mu \in \mathbb{R}_+^p, \langle g(x_{\mu^*}), \mu - \mu^* \rangle \leq 0$$

or cette propriété est vraie selon IV.5.(a) ainsi $\mu^* = P_{\mathbb{R}_+^p}(\mu^* + \rho g(x_{\mu^*}))$

Remarques : La norme de Ax est à prendre dans \mathbb{R}^p pour $x \in \mathbb{R}^n$.

Par ailleurs comme en dimension finie une application linéaire est continue, il existe $K > 0$ tel que

$$\forall x \in \mathbb{R}^n, \|Ax\| \leq K\|x\|$$

En prenant $\rho = \frac{\alpha}{K^2}$, on a bien $\|Ax\| \leq \sqrt{\frac{\alpha}{\rho}}\|x\|$ pour tout $x \in \mathbb{R}^n$.

(d) On a (avec la question précédente)

$$\|\mu^{k+1} - \mu^*\| = \|P_{\mathbb{R}_+^p}(\mu^k + \rho g(x_{\mu^k})) - P_{\mathbb{R}_+^p}(\mu^* + \rho g(x_{\mu^*}))\|$$

Comme $P_{\mathbb{R}_+^p}$ est 1 lipschitzienne,

$$\|\mu^{k+1} - \mu^*\| \leq \|\mu^k + \rho g(x_{\mu^k}) - \mu^* - \rho g(x_{\mu^*})\| = \|\mu^k - \mu^* + \rho A(x_{\mu^k} - x_{\mu^*})\|$$

On élève au carré et on développe la carré scalaire du membre de droite

$$\|\mu^{k+1} - \mu^*\|^2 \leq \|\mu^k - \mu^*\|^2 + \rho^2 \|A(x_{\mu^k} - x_{\mu^*})\|^2 + 2\rho \langle \mu^k - \mu^*, A(x_{\mu^k} - x_{\mu^*}) \rangle$$

L'hypothèse sur A donne enfin

$$\|\mu^{k+1} - \mu^*\|^2 \leq \|\mu^k - \mu^*\|^2 + \alpha\rho \|x_{\mu^k} - x_{\mu^*}\|^2 + 2\rho \langle \mu^k - \mu^*, A(x_{\mu^k} - x_{\mu^*}) \rangle$$

Comme $\langle Ax, y \rangle = \langle x, A^\top y \rangle$, on a aussi, en utilisant (b) :

$$2\rho \langle \mu^k - \mu^*, A(x_{\mu^k} - x_{\mu^*}) \rangle = 2\rho \langle A^\top(\mu^k - \mu^*), x_{\mu^k} - x_{\mu^*} \rangle = -2\rho \langle \nabla f(x_{\mu^k}) - \nabla f(x_{\mu^*}), x_{\mu^k} - x_{\mu^*} \rangle$$

et comme f est α -convexe en utilisant I.3.(b),

$$2\rho \langle \mu^k - \mu^*, A(x_{\mu^k} - x_{\mu^*}) \rangle \leq -2\rho\alpha \|x_{\mu^k} - x_{\mu^*}\|^2$$

On peut ainsi conclure que

$$\|\mu^{k+1} - \mu^*\|^2 \leq \|\mu^k - \mu^*\|^2 - \alpha\rho \|x_{\mu^k} - x_{\mu^*}\|^2$$

La suite $(\|\mu^k - \mu^*\|^2)_{k \geq 0}$ est décroissante et minorée par 0 et converge donc vers une limite $\ell \geq 0$. De plus,

$$\|x_{\mu^k} - x_{\mu^*}\|^2 \leq \frac{1}{\alpha\rho} (\|\mu^k - \mu^*\|^2 - \|\mu^{k+1} - \mu^*\|^2)$$

et $(x_{\mu^k})_{k \in \mathbb{N}}$ converge vers $x_{\mu^*} = x^*$ d'après le théorème des gendarmes.

(e) D'après (d), on remarque que la suite $(\mu^k)_k$ est bornée.

Comme \mathbb{R}^p est de dimension finie, on considère $\bar{\mu}$ une valeur d'adhérence et il suffit d'établir que $\bar{\mu} = \mu^*$.

On considère alors une extractrice φ telle que $(\mu^{\varphi(k)})_k$ converge vers $\bar{\mu}$.

Comme $\mu \mapsto x_\mu$ est continue, la suite $(x_{\mu^{\varphi(k)}})_k$ converge vers $x_{\bar{\mu}}$.

Or d'après (d) la suite extraite $(x_{\mu^{\varphi(k)}})_k$ converge vers x^* .

Donc par unicité de la limite $x_{\mu^*} = x^* = x_{\bar{\mu}}$

Ainsi à l'aide de (a),

$$\tau(\mu^*) = A^\top \mu^* = -\nabla f(x^*) = A^\top \bar{\mu} = \tau(\bar{\mu})$$

τ a été introduit en (b) et est injective d'où $\bar{\mu} = \mu^*$ donc

$$\mu^k \xrightarrow{k \rightarrow +\infty} \mu^*$$