3.5.3 Exercices (algorithmes pour l'optimisation avec contraintes)

Exercice 133 (Méthode de pénalisation)**.**
Soit f une fonction continue et strictement convexe de $\mathbb{R}^n$ dans $\mathbb{R}$, satisfaisant de plus :

$$\lim_{|x|\to+\infty} f(x) = +\infty.$$

Soit K un sous ensemble non vide, convexe (c'est–à–dire tel que $\forall (x,y) \in K^2$, $tx + (1-t)y \in K$, $\forall t \in]0,1[$), et fermé de $\mathbb{R}^n$. Soit ψ une fonction continue de $\mathbb{R}^n$ dans $[0,+\infty[$ telle que $\psi(x) = 0$ si et seulement si $x \in K$. Pour $n \in \mathbb{N}$, on définit la fonction f_k par $f_k(x) = f(x) + n\psi(x)$.

1. Montrer qu'il existe au moins un élément $\bar{x}_k \in \mathbb{R}^n$ tel que $f_k(\bar{x}_k) = \inf_{x\in\mathbb{R}^n} f_k(x)$, et qu'il existe un unique élément $\bar{x}_K \in K$ tel que $f(\bar{x}_K) = \inf_{x\in K} f(x)$.
2. Montrer que pour tout $n \in \mathbb{N}$,
$$f(\bar{x}_n) \le f_k(\bar{x}_n) \le f(\bar{x}_K).$$
3. En déduire qu'il existe une sous-suite $(\bar{x}_{n_k})_{k\in\mathbb{N}}$ et $y \in K$ tels que $\bar{x}_{n_k} \to y$ lorsque $k \to +\infty$.
4. Montrer que $y = \bar{x}_K$. En déduire que toute la suite $(\bar{x}_k)_{n\in\mathbb{N}}$ converge vers $\bar{x}_K$.
5. Déduire de ces questions un algorithme (dit "de pénalisation") de résolution du problème de minimisation suivant :
$$\begin{cases} \text{Trouver } \bar{x}_K \in K; \\ f(\bar{x}_K) \le f(x), \forall x \in K, \end{cases}$$
en donnant un exemple de fonction ψ.

Exercice 134 (Convergence de l'algorithme d'Uzawa)**.** *Corrigé en page 279*
Soient $n \ge 1$ $p \in \mathbb{N}^\star$. Soit $f \in C^1(\mathbb{R}^n, \mathbb{R})$ une fonction telle que

$$\exists \alpha > 0,\ (\nabla f(x) - \nabla f(y)) \cdot (x-y) \ge \alpha |x-y|^2,\ \forall x, y \in \mathbb{R}^n.$$

Soit $C \in M_{p,n}(\mathbb{R})$ (C est donc une matrice, à éléments réels, ayant p lignes et n colonnes) et $d \in \mathbb{R}^p$. On note $D = \{x \in \mathbb{R}^n, Cx \le d\}$ et $\mathcal{C}^+ = \{u \in \mathbb{R}^p, u \ge 0\}$.
On suppose $D \ne \emptyset$ et on s'intéresse au problème suivant :

$$x \in D,\quad f(x) \le f(y),\ \forall y \in D. \tag{3.68}$$

1. Montrer que $f(y) \ge f(x) + \nabla f(x) \cdot (y-x) + \frac{\alpha}{2}|x-y|^2$ pour tout $x, y \in \mathbb{R}^n$.
2. Montrer que f est strictement convexe et que $f(x) \to \infty$ quand $|x| \to \infty$. En déduire qu'il existe une et une seule solution au problème (3.68).

Dans la suite, on note $\overline{x}$ cette solution.
Pour $u \in \mathbb{R}^p$ et $x \in \mathbb{R}^n$, on pose $L(x,u) = f(x) + u \cdot (Cx - d)$.

3. Soit $u \in \mathbb{R}^p$ (dans cette question, u est fixé). Montrer que l'application $x \to L(x,u)$ est strictement convexe (de $\mathbb{R}^n$ dans $\mathbb{R}$) et que $L(x,u) \to \infty$ quand $|x| \to \infty$ [Utiliser la question 1]. En déduire qu'il existe une et une seule solution au problème suivant :
$$x \in \mathbb{R}^n,\quad L(x,u) \le L(y,u),\ \forall y \in \mathbb{R}^n. \tag{3.69}$$
Dans la suite, on note x_u cette solution. Montrer que x_u est aussi l'unique élément de $\mathbb{R}^n$ t.q. $\nabla f(x_u) + C^t u = 0$.
4. On admet que le théorème de Kuhn-Tucker s'applique ici (cf. cours). Il existe donc $\overline{u} \in \mathcal{C}^+$ t.q. $\nabla f(\overline{x}) + C^t \overline{u} = 0$ et $\overline{u} \cdot (C\overline{x} - d) = 0$. Montrer que $(\overline{x}, \overline{u})$ est un point selle de L sur $\mathbb{R}^n \times \mathcal{C}^+$, c'est-à-dire :
$$L(\overline{x}, v) \le L(\overline{x}, \overline{u}) \le L(y, \overline{u}),\ \forall (y,v) \in \mathbb{R}^n \times \mathcal{C}^+. \tag{3.70}$$

Pour $u \in \mathbb{R}^p$, on pose $M(u) = L(x_u, u)$ (de sorte que $M(u) = \inf\{L(x,u),\, x \in \mathbb{R}^n\}$). On considère alors le problème suivant :

$$u \in \mathcal{C}^+, \;\; M(u) \geq M(v), \; \forall v \in \mathcal{C}^+. \tag{3.71}$$

5. Soit $(x,u) \in \mathbb{R}^n \times \mathcal{C}^+$ un point selle de L sur $\mathbb{R}^n \times \mathcal{C}^+$ (c'est-à-dire $L(x,v) \leq L(x,u) \leq L(y,u)$, pour tout $(y,v) \in \mathbb{R}^n \times \mathcal{C}^+$). Montrer que $x = \overline{x} = x_u$ (on rappelle que $\overline{x}$ est l'unique solution de (3.68) et x_u est l'unique solution de (3.69)) et que u est solution de (3.71). [On pourra commencer par montrer, en utilisant la première inégalité, que $x \in D$ et $u \cdot (Cx - d) = 0$.]

   Montrer que $\nabla f(\overline{x}) + C^t u = 0$ et que $u = P_{\mathcal{C}^+}(u + \rho(C\overline{x} - d))$, pour tout $\rho > 0$, où $P_{\mathcal{C}^+}$ désigne l'opérateur de projection orthogonale sur $\mathcal{C}^+$. [on rappelle que si $v \in \mathbb{R}^p$ et $w \in \mathcal{C}^+$, on a $w = P_{\mathcal{C}^+} v \iff ((v-w)\cdot(w-z) \geq 0, \forall z \in \mathcal{C}^+)$.]
6. Déduire des questions 2, 4 et 5 que le problème (3.71) admet au moins une solution.
7. On admet que l'application $u \mapsto x_u$ est dérivable. Montrer que l'algorithme du gradient à pas fixe avec projection pour trouver la solution de (3.71) s'écrit (on désigne par $\rho > 0$ le pas de l'algorithme) :

   Initialisation. $u_0 \in \mathcal{C}^+$.

   Itérations. Pour $u_k \in \mathcal{C}^+$ connu ($k \geq 0$). On calcule $x_k \in \mathbb{R}^n$ t.q. $\nabla f(x_k) + C^t u_k = 0$ (montrer qu'un tel x_k existe et est unique) et on pose $u_{k+1} = P_{\mathcal{C}^+}(u_k + \rho(Cx_k - d))$.

Dans la suite, on s'intéresse à la convergence de la suite $(x_k, u_k)_{k\in\mathbb{N}}$ donnée par cet algorithme.

8. Soit ρ t.q. $0 < \rho < 2\alpha/\|C\|^2$ avec $\|C\| = \sup\{|Cx|,\, x \in \mathbb{R}^n \text{ t.q. } |x| = 1\}$. Soit $(\overline{x}, \overline{u}) \in \mathbb{R}^n \times \mathcal{C}^+$ un point selle de L sur $\mathbb{R}^n \times \mathcal{C}^+$ (c'est-à-dire vérifiant (3.70)) et $(x_k, u_k)_{k\in\mathbb{N}}$ la suite donnée par l'algorithme de la question précédente. Montrer que

   $$|u_{k+1} - \overline{u}|^2 \leq |u_k - \overline{u}|^2 - \rho(2\alpha - \rho\|C\|^2)|x_k - \overline{x}|^2, \; \forall k \in \mathbb{R}^n.$$

   En déduire que $x_k \to \overline{x}$ quand $k \to \infty$.

   Montrer que la suite $(u_k)_{k\in\mathbb{N}}$ est bornée et que, si $\tilde{u}$ est une valeur d'adhérence de la suite $(u_k)_{k\in\mathbb{N}}$, on a $\nabla f(\overline{x}) + C^t\tilde{u} = 0$. En déduire que, si rang(C)=p, on a $u_k \to \overline{u}$ quand $k \to \infty$ et que $\overline{u}$ est *l'unique* élément de $\mathcal{C}^+$ t.q. $\nabla f(\overline{x}) + C^t\overline{u} = 0$.