Exercice 1 : Gradient projeté

Soit V un espace de Hilbert. Soit $K \subset V$ un convexe fermé non vide. On souhaite écrire un algorithme pour résoudre le problème de minimisation :

$$J(u) = \inf_{v \in K} J(v),$$

où J est une fonctionnelle α-convexe différentiable telle que ∇J soit Lipschitzienne c-à-d. il existe $C > 0$ tel que :

$$\|\nabla J(v) - \nabla J(w)\| \leq C\|v - w\| \; \forall v, w \in V.$$

a) Montrer que l'application $v \to v - \mu\nabla J(v)$ est contractante si $0 < \mu < 2\alpha/C^2$.

b) On note P_K l'opérateur de projection sur K. Montrer que $\|P_K x - P_K y\|^2 \leq (P_K x - P_K y, x - y)$ pour tout $x, y \in V$ et en déduire que

$$\|P_K x - P_K y\| \leq \|x - y\| \text{ pour tout } x, y \in V.$$

c) Montrer que la solution du problème de minimisation vérifie $u = P_K(u - \mu\nabla J(u))$ pour tout $\mu > 0$.

d) Expliquer la méthode de gradient projeté :

$$u^{n+1} = P_K(u^n - \mu\nabla J(u^n)).$$

e) Montrer que sous la condition $0 < \mu < 2\alpha/C^2$ la méthode de gradient projeté converge.

Exercice 2 : Gradient projeté : application

Soit J la fonctionnelle de l'exercice précédent. On cherche à résoudre

$$J(u) = \inf_{v \in K} J(v),$$

avec $K = \{v \in \mathbb{R}^N \text{ tel que} \sum_{i=1}^{N} v_i = 1\}$.

a) Montrer que K est convexe.

b) Calculer n le vecteur normal à K et expliciter P_K, l'opérateur de projection orthogonale de $\mathbb{R}^N$ sur K.

c) Ecrire l'itération k de la méthode de gradient projecté.

d) Vérifier que si cette suite converge, alors sa limite vérifie le théorème des extrema liés.

Exercice 3 : Méthode de pénalisation

On considère le problème :

$$\min_{h(x)=0} f(x)$$

où f et h sont des fonctions de $\mathbb{R}^N$ dans $\mathbb{R}$. Au lieu de résoudre ce problème on se propose de résoudre le problème sans contrainte :

$$\min_{x\in\mathbb{R}^N} \left(f(x) + \frac{\mu}{2}|h(x)|^2 \right),$$

avec μ “grand”.

On considère le cas où $f(x) = \frac{1}{2}x^T Q x$ avec Q diagonale, avec des éléments diagonaux strictement positifs et $h(x) = C^t x$ avec $C^t = (1, 0, \cdots, 0)$. On utilise une méthode de gradient à pas optimal pour résoudre ce nouveau problème : comment varie le taux de convergence en fonction de μ ? Commenter.