Devoir - calcul différentiel

Problème - *Optimisation - Gradient à pas optimal*
Soit J une fonction de classe $\mathcal{C}^1$, de $\mathbb{R}^n$ dans $\mathbb{R}$. On suppose qu'il existe un $\alpha > 0$ réel pour lequel l'application J vérifie la condition,

$$\forall (u, v) \in \mathbb{R}^n \times \mathbb{R}^n < \nabla J(v) - \nabla J(u), v - u >\geqslant \alpha ||v - u||^2 \tag{1}$$

où $\nabla J(u)$ est le gradient de J en u.

L'objectif de cet exercice est de prouver l'existence puis de trouver un élément $x \in \mathbb{R}^n$ vérifiant la condition,

$$J(x) = \inf_{u \in \mathbb{R}^n} J(u).$$

On construit pour cela une suite $(u_n)_{n \in \mathbb{N}}$ dite *minimisante* pour J, *i.e.* telle que si la suite $(J(u_n))_{n \in \mathbb{N}}$ converge vers $J(x)$ alors la suite $(u_n)_{n \in \mathbb{N}}$ converge vers l'élement x cherché.

Partie A - Étude théorique Justifions dans un premier temps l'existence et l'unicité du minimum de J.

i. Vérifier que la condition (1) est équivalente à la condition suivante,

$$\forall (u, v) \in \mathbb{R}^n \times \mathbb{R}^n J(v) \geqslant J(u) + < \nabla J(u), v - u > + \frac{\alpha}{2} ||v - u||^2 \tag{2}$$

Indication : on pensera à Taylor avec reste intégral.

ii. En déduire que la fonction J est coercive sur $\mathbb{R}^n$, *i.e.* $\lim_{u \to \infty} J(u) = \infty$
Indication : on cherchera une minoration en utilisant Cauchy-Schwarz.

iii. En déduire par compacité l'existence pour J d'un minimum global x sur $\mathbb{R}^n$.
Vérifier en utilisant (2) qu'il est unique.

Partie B - Une fonction intermédiaire Considérant $u \in \mathbb{R}^n - \{x\}$, on définit la fonction,

$$\begin{array}{rccl} \phi_u : & \mathbb{R} & \to & \mathbb{R} \\ & t & \mapsto & J(u + t\nabla J(u)) \end{array}$$

i. Calculer la dérivée de ϕ_u et montrer que ϕ_u est une fonction convexe.

ii. Montrer que $\lim_{t \to \infty} \phi_u(t) = \infty$.
En déduire par un raisonnement similaire à la partie précédente, que ϕ_u admet un unique minimum global sur $\mathbb{R}$.

Partie C - Minimisation On construit ici la suite minimisante annoncée.
Le résultat précédent permet de définir la suite minimisante $(u_n)_{n \in \mathbb{N}}$ de $\mathbb{R}^n$ de la façon suivante,

$$\left\{ \begin{array}{l} u_0 \in \mathbb{R}^n, \\ u_{n+1} = u_n + t_n \nabla J(u_n) \end{array} \right.$$

où t_n est l'unique réel obtenu précédement qui minimise ϕ_{u_n}, *i.e.* vérifiant la condition

$$\phi_{u_n}(t_n) = \min(\phi_{u_n}).$$

i. Considérant $\phi'_{u_k}(t_k)$, montrer que quelque soit $n \in \mathbb{N}$, $\nabla J(u_n) \perp \nabla J(u_{n+1})$.
En déduire la relation,

$$J(u_k) - J(u_{k+1}) > \frac{\alpha}{2}||u_k - u_{k+1}||^2$$

ii. Étudiant succintement la monotonie de la suite $(J(u_n))_{n\in\mathbb{N}}$, en déduire que la suite $(J(u_k) - J(u_{k+1}))_{n\in\mathbb{N}}$ tend vers 0, puis que la suite $(u_{n+1} - u_n)_{n\in\mathbb{N}}$ tend vers 0,

iii. Montrer que la suite $(u_k)_{n\in\mathbb{N}}$ est bornée. Par un argument de continuité uniforme, déduire de la question précédente que la suite $(\nabla J(u_{n+1}) - \nabla J(u_n))_{n\in\mathbb{N}}$ tend vers 0,

iv. Utilisant la relation d'orthogonalité, montrer enfin que la suite $(\nabla J(u_n))_{n\in\mathbb{N}}$ tend vers 0.
Indication : on évaluera la quantité $||\nabla J(u_k)||^2$.

v. En déduire que la suite $(u_n)_{n\in\mathbb{N}}$ converge vers le point x précédement défini.
Indication : on partira de la condition (1).

Quelques remarques :

i. la méthode du gradient à pas optimal est un exemple *d'algorithme de descente*, ainsi nommés parce que chaque terme de la suite est obtenu à partir du précédent en suivant la direction de plus grande pente,

ii. il existe une version moins efficace de cet algorithme, à pas fixe (*i.e.* le pas t_n est constant dans la suite $(u_n)_{n\in\mathbb{N}}$) ; ainsi qu'une méthode plus efficace, la méthode du gradient conjugué. Dans cette dernière, la direction choisie est orthogonale à toutes les directions précédament obtenues (à mettre en relation avec la question C.i),

iii. ces méthodes s'appliquent notamment à la résolution numérique de systèmes linéaires de la forme $A.X = B$.