Université de Paris
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

Master $1^{\text{ère}}$ année, MMA, Optimisation

Seconde session du 18/06/2021 - Correction

Durée 2h. Aucun document n'est autorisé. Les exercices sont indépendants. Le barème est indicatif. On prendra soin de bien justifier les réponses.

Exercice 1. Question de cours (3 points)

Énoncer et démontrer la condition nécessaire et suffisante de convexité pour une fonction $f : \Omega \to \mathbb{R}$ différentiable sur un ouvert $\Omega \subset \mathbb{R}^N$.

Correction. Voir le cours.

Exercice 2. (5 points)

Soient $N, M, P \in \mathbb{N}^*$. On définit pour tout $x \in \mathbb{R}^N$,

$$f(x) = \frac{1}{2} \|Hx - y\|_2^2 + \lambda \|Dx\|_2^2 + \mu \|x\|_2^2 ,$$

avec $H : \mathbb{R}^N \to \mathbb{R}^M$ et $D : \mathbb{R}^N \to \mathbb{R}^P$ deux opérateurs linéaires, $y \in \mathbb{R}^M$ un vecteur fixé, $\lambda, \mu > 0$ deux constantes et $\|\cdot\|_2$ la norme euclidienne, où par abus de langage on utilise la même notation peu importe l'espace sous jacent ($\mathbb{R}^N, \mathbb{R}^M$ ou $\mathbb{R}^P$).

1. (a) Rappeler la définition d'une fonctionnelle quadratique.
   (b) Montrer que f est une fonctionnelle quadratique.
   (c) Que vaut $\nabla f(x)$ pour tout $x \in \mathbb{R}^N$? *On pourra utiliser directement les résultats sur les fonctionnelles quadratiques.*
   (d) Que vaut $\nabla^2 f(x)$ pour tout $x \in \mathbb{R}^N$? *On pourra utiliser directement les résultats sur les fonctionnelles quadratiques.*
2. Démontrer que f est strictement convexe.
3. Démontrer que f admet un unique minimum global.
4. Les hypothèses du théorème de convergence de la méthode de gradient à pas optimal sont-elles satisfaites ?
5. Écrire en Python l'algorithme de descente du gradient à pas optimal pour la fonction f.
6. On note $(x_n)_{n\in\mathbb{N}}$ la suite produite par l'algorithme de descente du gradient à pas optimal appliqué à f. Que dire des vecteurs $\Delta x_n = x_{n+1} - x_n$ et $\Delta x_{n+1} = x_{n+2} - x_{n+1}$?

Correction.

1. (a) C'est une fonction $f : \mathbb{R}^N \to \mathbb{R}$ qui s'écrit pour tout $x \in \mathbb{R}^N$, $f(x) = \frac{1}{2}\langle Ax,\, x\rangle - \langle b,\, x\rangle + c$ avec A une matrice symétrique de taille $N \times N$, $b \in \mathbb{R}^N$ et $c \in \mathbb{R}$.
   (b) En écrivant les normes euclidiennes au carré sous forme de produits scalaires, en développant et passant à la transposée, on montre que f est une fonctionnelle quadratique avec $A = H^T H + 2\lambda D^T D + 2\mu I_N$, qui est bien une matrice symétrique, $b = H^T y$ et $c = \frac{1}{2}\|y\|_2^2$.
   (c) D'après le cours, on sait que pour tout $x \in \mathbb{R}^N$, $\nabla f(x) = Ax - b$.
   (d) D'après le cours, on sait que pour tout $x \in \mathbb{R}^N$, $\nabla^2 f(x) = A$.
2. Comme pour tout $x \in \mathbb{R}^N$, $\nabla^2 f(x) = A$, avec A qui est une matrice symétrique définie positive, puisque $A \succeq 2\mu I_N$, on en déduit que f est strictement convexe.
3. La fonction f admet un point critique puisque A est inversible, donc un minimum globale comme f est convexe, et il est unique puisque f est strictement convexe.
4. On a pour tout $x \in \mathbb{R}^N$, $2\mu I_N \preceq \nabla^2 f(x) \preceq (\lambda_{\max}(H^T H)) + 2\lambda\lambda_{\max}(D^T D) + 2\mu)I_N$, donc les hypothèses sont bien satisfaites.
5. Voir le cours.
6. On sait pour la méthode de la descente de gradient à pas optimal que les gradients successifs sont orthogonaux, donc pour tout $n \in \mathbb{N}$, $\Delta x_n \perp \Delta x_{n+1}$.

Exercice 3. (13 points)

Soit $N \in \mathbb{N}^*$. Dans la suite, on considère $\mathbb{R}^N$ munit de son produit scalaire canonique noté $\langle \cdot,\, \cdot \rangle$ et on note $\|\cdot\|_2$ la norme euclidienne (issue du produit scalaire canonique).

Les parties I et II sont indépendantes. La partie III utilise certains résultats des deux premières.

Partie I : questions préliminaires. Dans toute cette partie uniquement on suppose que $N = 2$. On note (e_1, e_2) la base canonique de $\mathbb{R}^2$.

1. Décrire l'ensemble $\mathcal{H} = \{x \in \mathbb{R}^2 : \langle e_1,\, x\rangle = 1\}$. Le représenter graphiquement.
2. (a) Donner l'ensemble de définition, noté V, de $g : x \mapsto -\ln(1 - \langle e_1,\, x\rangle)$. Le représenter graphiquement.
   (b) Est-ce un ensemble ouvert ou fermé ? Est-ce un ensemble convexe ? Justifier brièvement.
3. (a) En déduire l'ensemble de définition, noté U, de la fonction
   $$f : x \mapsto -\ln(1 - \langle e_1,\, x\rangle) - \ln(1 - \langle e_2,\, x\rangle) - \ln(2 - \langle -(e_1 + e_2),\, x\rangle).$$
   Le représenter graphiquement.
   (b) Est-ce un ensemble ouvert ou fermé ? Est-ce un ensemble convexe ? Justifier brièvement.
   (c) U est-il borné ? *On pourra également observer, en faisant un dessin, que pour tout $v \in \mathbb{R}^2 \setminus \{0\}$, il est toujours possible d'avoir $\langle a,\, v\rangle > 0$ avec un $a \in \mathbb{R}^2$ choisi dans $\{e_1, e_2, -(e_1 + e_2)\}$.*

4. On considère $\tilde{f} : x \mapsto -\ln(-\langle e_1, x\rangle) - \ln(-\langle e_2, x\rangle)$. Représenter graphiquement l'ensemble de définition de $\tilde{f}$. Est-il borné ? Donner un $v \in \mathbb{R}^2 \setminus \{0\}$ tel que $\langle e_1, v\rangle \leq 0$ <u>et</u> $\langle e_2, v\rangle \leq 0$. Que peut-on dire alors de la demi-droite $\mathbb{R}_+^* v$?

Partie II : existence d'un minimum. On considère la fonction $f : U \to \mathbb{R}$ où U est un ouvert, non vide, borné de $\mathbb{R}^N$. On suppose que f est continue, bornée inférieurement et tel que pour tout $\alpha \in \mathbb{R}$, $f^{-1}(]-\infty, \alpha])$ est fermé <u>dans $\mathbb{R}^N$</u>.

1. Justifier que $p = \inf_{x\in U} f(x)$ est fini ($p > -\infty$).
2. Démontrer l'existence d'une suite minimisante, c'est-à-dire d'une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de U, telle que $f(x_n) \underset{n\to+\infty}{\longrightarrow} p$. *On pourra utiliser la caractérisation de la borne inférieure.*
3. Démontrer qu'il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, on a $x_n \in f^{-1}(]-\infty, p+1])$.
4. En déduire qu'il existe $x^* \in U$ tel que $f(x^*) = p$, c'est-à-dire que le problème $\inf_{x\in U} f(x)$ admet au moins une solution.

Partie III : application. On considère $f : U \to \mathbb{R}$ avec $U \subset \mathbb{R}^N$ tel que

$$\forall x \in U, \quad f(x) = -\sum_{i=1}^{m} \ln(b_i - \langle a_i, x\rangle),$$

avec $m \in \mathbb{N}^*$, et pout tout $i \in \{1, \ldots, m\}$, $b_i \in \mathbb{R}$, $a_i \in \mathbb{R}^N$. On fait de plus l'hypothèse suivante

$$\forall v \in \mathbb{R}^N \setminus \{0\}, \ \exists i_0 \in \{1, \ldots, m\}, \quad \langle a_{i_0}, v\rangle > 0.$$

On suppose de plus que U est non vide.

1. Montrer que $\mathrm{Vect}(a_1, \ldots, a_m) = \mathbb{R}^N$.
2. (a) Donner le plus grand ouvert $U \subset \mathbb{R}^N$ sur lequel f est défini.
   (b) Justifier que U est convexe.
3. Justifier brièvement que f est $\mathcal{C}^2$ sur U.
4. (a) Démontrer que pour tout $x \in U$, et tout $h \in \mathbb{R}^N$, la différentielle de f en x appliquée à h est donnée par

$$\mathrm{d}f(x)(h) = \sum_{i=1}^{m} \frac{\langle a_i, h\rangle}{b_i - \langle a_i, x\rangle}.$$

   (b) En déduire que pour tout $x \in U$, $\nabla f(x) = \sum_{i=1}^{m} \frac{1}{b_i - \langle a_i, x\rangle} a_i$.
5. (a) Démontrer que pour tout $x \in U$, et tout $(h, k) \in \mathbb{R}^N \times \mathbb{R}^N$, la différentielle seconde de f en x appliquée à (h, k) est donnée par

$$\mathrm{d}^2 f(x)(h, k) = \sum_{i=1}^{m} \frac{\langle a_i, h\rangle \langle a_i, k\rangle}{(b_i - \langle a_i, x\rangle)^2}.$$

(b) Démontrer que f est strictement convexe sur U.

6. Montrer que U est borné. *On pourra raisonner par l'absurde et utiliser la compacité de la sphère unité de* $\mathbb{R}^N$.

7. (a) Soient $x, x_0 \in U$. Justifier que

$$f(x) \geq f(x_0) + \left\langle \sum_{i=1}^{m} \frac{1}{b_i - \langle a_i,\, x_0 \rangle} a_i,\, x - x_0 \right\rangle.$$

Puis montrer que

$$f(x_0) + \left\langle \sum_{i=1}^{m} \frac{1}{b_i - \langle a_i,\, x_0 \rangle} a_i,\, x - x_0 \right\rangle = f(x_0) + m - \sum_{i=1}^{m} \frac{b_i - \langle a_i,\, x \rangle}{b_i - \langle a_i,\, x_0 \rangle}.$$

(b) Déduire des deux précédentes questions que f est bornée inférieurement.

8. Déduire de l'ensemble des questions que f admet un unique minimum sur U.

Correction.

Partie I.

1. C'est un hyperplan affine de vecteur normal e_1 et passant par le point $(1, 0)$.

2. (a) g est définie pour les $x \in \mathbb{R}^2$ satisfaisant $1 - \langle e_1,\, x \rangle > 0$. L'ensemble de définition de g est donc l'ensemble V représentant le demi plan ouvert contenant le point $(0, 0)$ et dont la frontière est $\mathcal{H}$ (il est situé "à gauche" de $\mathcal{H}$).

   (b) V est ouvert, vu la condition ouverte sur $x \in \mathbb{R}^2$, $1 - \langle e_1,\, x \rangle > 0$, le définissant. C'est un ensemble convexe comme c'est un demi plan.

3. (a) L'ensemble de définition U de f est l'intersection des demi plans ouverts $\{x \in \mathbb{R}^2 : \langle e_1,\, x \rangle < 1\}$, $\{x \in \mathbb{R}^2 : \langle e_2,\, x \rangle < 1\}$, $\{x \in \mathbb{R}^2 : \langle -(e_1 + e_2),\, x \rangle < 2\}$ car il faut que chacun des termes dans les logarithmes soient strictement positifs. La représentation graphique de U donne l'intérieur d'un triangle rectangle contenant $(0, 0)$.

   (b) U est ouvert comme intersection finie d'ouverts et est convexe comme intersection finie de convexes.

   (c) On remarque graphiquement que U est borné (puisque c'est l'intérieur d'un triangle rectangle).

4. Cette fois l'ensemble de définition $\tilde{U}$ de $\tilde{f}$ est l'orthant ouvert négatif de $\mathbb{R}^2$, c'est-à-dire les $x \in \mathbb{R}^2$ tels que leurs deux coordonnées soient strictement négatives. Cet ensemble n'est donc pas borné. $\tilde{U}$ est défini comme l'intersection des demi plans ouverts situés respectivement à gauche et en dessous de respectivement la droite de vecteur normal e_1 et la droite de vecteur normal e_2. On remarque que le vecteur $v = -(e_1 + e_2)$ satisfait $\langle e_1,\, v \rangle < 0$ et $\langle e_2,\, v \rangle < 0$. Alors qu'il n'était pas possible de trouver un tel v dans le cas du triangle rectangle. On remarque de plus que ce v défini une direction suivant laquelle $\tilde{U}$ n'est pas borné ; plus précisemment la demi droite (non borné donc) $\mathbb{R}_+^* v$ est inclue dans $\tilde{U}$.

Partie II.

1. L'ensemble $\{f(x) : x \in U\}$ est un sous ensemble non vide (car U est non vide) de $\mathbb{R}$, minorée car f est borné inférieurement, donc admet une borne inférieure finie i.e. $p > -\infty$.

2. D'après la caractérisation de la borne inférieure : pour tout $\varepsilon > 0$, il existe $x \in U$ tel que $0 \leq f(x) - p < \varepsilon$. Il suffit donc de prendre par exemple $\varepsilon = \frac{1}{n+1}$ pour tout $n \in \mathbb{N}$, pour produire une suite $(x_n)_{n\in\mathbb{N}}$ de l'ensemble U qui satisfait $f(x_n) \underset{n\to+\infty}{\longrightarrow} p$.

3. Puisque $f(x_n) \underset{n\to+\infty}{\longrightarrow} p$, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, on a $0 \leq f(x_n) - p \leq 1$, c'est-à-dire $f(x_n) \leq p + 1$, ou encore $x_n \in f^{-1}(]-\infty, p+1])$.

4. D'après la question précédente, la suite $(x_n)_{n\in\mathbb{N}}$ est contenue à partir d'un certain rang dans $f^{-1}(]-\infty, p+1])$, qui est fermé dans $\mathbb{R}^N$ par hypothèse. De plus comme U est borné et $f^{-1}(]-\infty, p+1]) \subset U$, c'est aussi le cas de $f^{-1}(]-\infty, p+1])$, qui est donc un compact de $\mathbb{R}^N$. Par conséquent on peut extraire une sous suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ de $(x_n)_{n\in\mathbb{N}}$ qui converge vers un $x^* \in f^{-1}(]-\infty, p+1]) \subset U$. Par continuité de f, on a alors $f(x^*) = \lim_{n\to+\infty} f(x_{\varphi(n)}) = p$. D'où f admet x^* comme minimum global sur U.

Partie III.

1. Soit $v \in \mathrm{Vect}(a_1, \ldots, a_m)^{\perp}$, alors par définition pour tout i, $\langle a_i,\, v\rangle = 0$. Nécessairement $v = 0$ car sinon contradiction avec l'hypothèse de l'énoncé. D'où $\mathrm{Vect}(a_1, \ldots, a_m)^{\perp} = \{0\}$ i.e. $\mathrm{Vect}(a_1, \ldots, a_m) = \mathbb{R}^N$.

2. (a) On a $U = \{x \in \mathbb{R}^N : \forall i \in \{1, \ldots, m\},\ \langle a_i,\, x\rangle < b_i\}$. C'est bien un ouvert, car les conditions le définissant sont ouvertes. Ou alors c'est l'intersection de m demi espaces affines ouverts.

   (b) U est l'intersection de m demi espaces affines ouverts qui sont convexes donc U est convexe.

3. Chacune des fonctions $x \in U \mapsto -\ln(b_i - \langle a_i,\, x\rangle)$ est $\mathcal{C}^2$ sur l'ouvert U comme composée de fonctions $\mathcal{C}^2$. Donc f est $\mathcal{C}^2$ sur U comme somme de fonctions $\mathcal{C}^2$ sur U.

4. (a) On pourrait appliquer directement la formule de la différentielle d'une composée. On peut aussi faire le calcul en partant de la définition. Soit $x \in U$ et $h \in \mathbb{R}^N$ (tel que $x + h \in U$). Alors $f(x+h) = -\sum_{i=1}^m \ln(b_i - \langle a_i,\, x+h\rangle) = -\sum_{i=1}^m g_i(\langle a_i,\, x\rangle + \langle a_i,\, h\rangle)$ où pour tout i, $g_i : t \mapsto \ln(b_i - t)$. Les fonctions g_i sont dérivables sur $]-\infty, b_i[$ et $g_i'(t) = -\frac{1}{b_i - t}$. D'où $g_i(\langle a_i,\, x\rangle + \langle a_i,\, h\rangle) = g_i(\langle a_i,\, x\rangle) + g_i'(\langle a_i,\, x\rangle)\,\langle a_i,\, h\rangle + o(\langle a_i,\, h\rangle) = g_i(\langle a_i,\, x\rangle) - \frac{\langle a_i,\, h\rangle}{b_i - \langle a_i,\, x\rangle} + o(\|h\|_2)$. Ainsi $f(x+h) = f(x) + \sum_{i=1}^m \frac{\langle a_i,\, h\rangle}{b_i - \langle a_i,\, x\rangle} + o(\|h\|_2)$. D'où par définition de la différentielle de f en x appliquée à h, $\mathrm{d}f(x)(h) = \sum_{i=1}^m \frac{\langle a_i,\, h\rangle}{b_i - \langle a_i,\, x\rangle}$ (cette expression est vraie ici pour tout $h \in \mathbb{R}^N$).

   (b) Pour une fonction $f : U \to \mathbb{R}$ différentiable sur l'ouvert U de $\mathbb{R}^N$, le gradient de f en tout $x \in U$ est par définition l'unique vecteur tel que pour tout $h \in \mathbb{R}^N$, $\mathrm{d}f(x)(h) = \langle \nabla f(x),\, h\rangle$. Donc par identification, on a pour tout $x \in U$, $\nabla f(x) = \sum_{i=1}^m \frac{1}{b_i - \langle a_i,\, x\rangle} a_i$.

5. (a) On pourrait appliquer directement la formule de la différentielle d'une composée. On peut aussi faire le calcul en partant de la définition. Soit $x \in U$ et $k \in \mathbb{R}^N$ (tel que $x+k \in U$). Alors la différentielle seconde de f en x appliquée à k peut-être obtenue en linéarisant $\mathrm{d}f(x+k)$ (puisque la différentielle seconde de f est la différentielle de $x \mapsto \mathrm{d}f(x)$). On a $\mathrm{d}f(x+k) = \sum_{i=1}^m \frac{\langle a_i, \cdot\rangle}{b_i - \langle a_i, x+k\rangle} = \sum_{i=1}^m \langle a_i, \cdot\rangle\, \tilde{g}_i(\langle a_i, x\rangle + \langle a_i, k\rangle)$, où pour tout i, $\tilde{g}_i : t \mapsto \frac{1}{b_i - t}$. Les fonctions $\tilde{g}_i$ sont dérivables sur $\mathbb{R} \setminus \{b_i\}$ et $\tilde{g}_i'(t) = \frac{1}{(b_i-t)^2}$. D'où $\tilde{g}_i(\langle a_i, x\rangle + \langle a_i, k\rangle) = \tilde{g}_i(\langle a_i, x\rangle) + \tilde{g}_i'(\langle a_i, x\rangle)\,\langle a_i, k\rangle + o(\langle a_i, k\rangle) = g_i(\langle a_i, x\rangle) + \frac{\langle a_i, k\rangle}{(b_i - \langle a_i, x\rangle)^2} + o(\|k\|_2)$. Ainsi, $\mathrm{d}f(x+k) = \mathrm{d}f(x) + \sum_{i=1}^m \langle a_i, \cdot\rangle \frac{\langle a_i, k\rangle}{(b_i-\langle a_i, x\rangle)^2} + o(\|k\|_2)$. D'où par définition, pour tout $h, k \in \mathbb{R}^N$, $\mathrm{d}^2 f(x)(h,k) = \sum_{i=1}^m \langle a_i, h\rangle \frac{\langle a_i, k\rangle}{(b_i - \langle a_i, x\rangle)^2}$.

   (b) On a pour tout $x \in U$ et tout $h \in \mathbb{R}^N \setminus \{0\}$, $\mathrm{d}^2 f(x)(h,h) = \sum_{i=1}^m \frac{\langle a_i, h\rangle^2}{(b_i - \langle a_i, x\rangle)^2} > 0$ (quantité positive puis forcément non nulle car au moins un des $\langle a_i, h\rangle$ est strictement positif par hypothèse sur les a_i). Donc f est strictement convexe sur U.

6. Supposons par l'absurde que U n'est pas borné. Alors il existe une suite $(x_n)_{n\in\mathbb{N}}$ de U telle que $\|x_n\|_2 \to +\infty$. On peut alors supposer que $x_n \neq 0$ pour tout $n \in \mathbb{N}$ (quitte à tronquer la suite de ses premiers termes). Alors pour tout $n \in \mathbb{N}$, soit $v_n = \frac{x_n}{\|x_n\|_2}$. Comme pour tout $n \in \mathbb{N}$, $\|v_n\|_2 = 1$ et que la sphère unité de $\mathbb{R}^N$ est compact, il existe une suite extraite $(v_{\varphi(n)})_{n\in\mathbb{N}}$ qui converge vers v tel que $\|v\|_2 = 1$. Or pour tout $i \in \{1, \dots, m\}$, tout $n \in \mathbb{N}$, on a : $\left\langle a_i, x_{\varphi(n)}\right\rangle < b_i$ d'où $\left\langle a_i, v_{\varphi(n)}\right\rangle < \frac{b_i}{\left\|x_{\varphi(n)}\right\|}$. Donc par passage à la limite, comme $x_{\varphi(n)} \to +\infty$, on obtient que pour tout $i \in \{1, \dots, m\}$, $\langle a_i, v\rangle \leq 0$. Ceci contredit l'hypothèse de l'énoncé sur les a_i (car $v \neq 0$). Donc U est borné.

7. (a) f est une fonction convexe différentiable sur l'ouvert convexe U donc pour tout $x, x_0 \in U$, on a $f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0\rangle$ (voir la question de cours). On obtient l'inégalité souhaitée en utilisant l'expression de $\nabla f(x_0)$ trouvé à la question 4.(b).

   Pour la deuxième partie, on a

$$
\begin{aligned}
f(x_0) + \left\langle \sum_{i=1}^m \frac{1}{b_i - \langle a_i, x_0\rangle} a_i,\, x - x_0 \right\rangle &= f(x_0) + \sum_{i=1}^m \frac{\langle a_i, x - x_0\rangle}{b_i - \langle a_i, x_0\rangle}, \\
&= f(x_0) - \sum_{i=1}^m \frac{-b_i + b_i + \langle a_i, x_0 - x\rangle}{b_i - \langle a_i, x_0\rangle}, \\
&= f(x_0) + m - \sum_{i=1}^m \frac{b_i - \langle a_i, x\rangle}{b_i - \langle a_i, x_0\rangle}.
\end{aligned}
$$

   (b) Comme U est borné, il existe $C > 0$ telle que pour tout $x \in U$, $\|x\|_2 \leq C$. Soit $x \in U$ alors pour tout i, $\langle a_i, x\rangle \geq -C\,\|a_i\|_2$ (Inégalité de Cauchy-Schwarz). D'où $\sum_{i=1}^m \frac{b_i - \langle a_i, x\rangle}{b_i - \langle a_i, x_0\rangle} \leq \sum_{i=1}^m \frac{b_i + \|a_i\|_2 C}{b_i - \langle a_i, x_0\rangle}$ et donc

$$
f(x) \geq f(x_0) + m - \sum_{i=1}^m \frac{b_i + \|a_i\|_2\, C}{b_i - \langle a_i, x_0\rangle}.
$$

Ce terme de droite étant une constante, on en déduit que f est borné inférieurement sur U. *En fait il était possible de se passer de l'égalité montrée à la question précédente et directement borner inférieurement, comme au dessus, le terme de droite dans l'inégalité à la question précédente.*

8. On va appliquer le résultat obtenu à la partie II. On sait que U est un ouvert non vide et que f est continue. De plus, on a montré que U est borné dans $\mathbb{R}^N$ et que f est bornée inférieurement. Il ne nous reste plus qu'à vérifier que pour tout $\alpha \in \mathbb{R}$, $f^{-1}(]-\infty, \alpha])$ est fermé dans $\mathbb{R}^N$. Utilisons la caractérisation séquentielle des fermés. Soit donc $\alpha \in \mathbb{R}$ et soit $(x_n)_{n\in\mathbb{N}}$ une suite de $V_\alpha = f^{-1}(]-\infty, \alpha])$ qui converge vers $x \in \mathbb{R}^N$. Montrons que $x \in V_\alpha$. Comme pour tout $n \in \mathbb{N}$, $f(x_n) \leq \alpha$, par continuité de f et passage à la limite dans l'inégalité, on en déduit que $f(x) \leq \alpha$. D'où $x \in V_\alpha$ comme souhaité. Ainsi d'après Partie II. 4., f admet un minimum global sur U. Comme f est strictement convexe, ce minimum global est unique.