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

Université Paris Cité

Master 1ère année, MMA, 2022-2023
Optimisation

Seconde Session du 22/06/2023

Durée 1h30. Aucun document n'est autorisé. Les exercices sont indépendants.

Total : 28.25 = 5.5 + 8.75 + 14, bonus : 5.5.

> Pour l'ensemble du sujet, on considère $M, N, P \in \mathbb{N}^*$. On munit $\mathbb{R}^M$, $\mathbb{R}^N$ et $\mathbb{R}^P$ de leurs produits scalaires canoniques respectifs que l'on note $\langle \cdot, \cdot \rangle$ indépendamment de la dimension (par abus de langage). Associés à ces produits scalaires, on utilise de même la notation unique $\|\cdot\|_2$ pour les normes euclidiennes.

Exercice 1 (Cours) (5.5pt)

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction différentiable.

1. a) Montrer que f est convexe sur $\mathbb{R}^N$ si et seulement si

$$\forall x, y \in \mathbb{R}^N, \quad f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle .$$

b) On suppose que $x \in \mathbb{R}^N$ est un minimum local de f et que f est convexe. Montrer que x est un point critique de f. En déduire alors que x est un minimum global de f.

2. Soit $m > 0$. Montrer que f est m-fortement convexe sur $\mathbb{R}^N$ si et seulement si

$$\forall x, y \in \mathbb{R}^N, \quad f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle + \frac{m}{2} \|y - x\|_2^2 .$$

Correction.

5.5 = (2.25 + 2) + 1.25

1. a) Supposons f est convexe sur $\mathbb{R}^N$. Soit $x, y \in \mathbb{R}^N$. Alors pour tout $t \in [0, 1]$, on a $f(x + t(y - x)) \leqslant f(x) + t(f(y) - f(x))$, i.e.

$$\frac{f(x + t(y - x)) - f(x)}{t} \leqslant f(y) - f(x).$$

Par passage à la limite quand $t \to 0$, on trouve donc $\langle \nabla f(x), y - x \rangle \leqslant f(y) - f(x)$ (1pt).

Réciproquement, supposons que pour tout $x, y \in \mathbb{R}^N$

$$\forall x, y \in \mathbb{R}^N, \quad f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle . \tag{1}$$

Soit $x, y \in \mathbb{R}^N$ et $t \in [0, 1]$. Alors appliquons l'inégalité (1) en $((1-t)x + ty, x)$ et $((1-t)x + ty, y)$, on obtient

$$f(x) \geqslant f((1-t)x + ty) + \left\langle \nabla f((1-t)x + ty), \underbrace{x - (1-t)x - ty}_{=t(x-y)} \right\rangle ,$$

$$f(y) \geqslant f((1-t)x + ty) + \left\langle \nabla f((1-t)x + ty), \underbrace{y - (1-t)x - ty}_{=-(1-t)(x-y)} \right\rangle , \text{(0.75pt)}$$

d'où en multipliant la première inégalité par $(1-t)$, la seconde par t et en les sommant, on obtient

$$(1-t)f(x) + tf(y) \geqslant f((1-t)x + ty),$$

c'est-à-dire f est convexe sur $\mathbb{R}^N$ (0.5pt).

b) Comme x est un minimum local de f, il existe $r > 0$ tel que pour tout $y \in B(x, r)$, $f(y) \geqslant f(x)$ (0.25pt). Soit $h \in \mathbb{R}^N$, il existe donc $t_0 > 0$ tel que pour tout $t \in]0, t_0[$, on a $f(x + th) - f(x) \geqslant 0$. En divisant cette inégalité par t et en faisant $t \to 0$, on trouve donc $\langle \nabla f(x), h \rangle \geqslant 0$ (0.5pt). Cette inégalité est également vraie en remplaçant h par $-h$, d'où $\langle \nabla f(x), h \rangle = 0$ (0.25pt). Comme cette égalité est vraie pour tout $h \in \mathbb{R}^N$, on a $\nabla f(x) \in (\mathbb{R}^N)^\perp$, i.e. $\nabla f(x) = 0$. x est donc bien un point critique de f (0.5pt).

Comme f est convexe on sait d'après la précédente question que pour tout $y \in \mathbb{R}^N$, $f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle$. Comme $\nabla f(x) = 0$, on a donc pour tout $y \in \mathbb{R}^N$, $f(y) \geqslant f(x)$. D'où x est un minimum global de f (0.5pt).

> Notez les passages surlignés avec le passage du local au global.

2. f est m-fortement convexe sur $\mathbb{R}^N$ si et seulement si $g = f - \frac{m}{2}\|\cdot\|_2^2$ est convexe sur $\mathbb{R}^N$ (0.25pt). Or g est différentiable, comme somme de fonctions différentiables, donc c'est équivalent à avoir
$$\forall x, y \in \mathbb{R}^N, \quad g(y) \geqslant g(x) + \langle \nabla g(x),\, y - x\rangle \,. \,(0.25pt)$$
On a pour tout $x \in \mathbb{R}^N$, $\nabla g(x) = \nabla f(x) - mx$ (0.25pt) et après quelques opérations arithmétiques, on obtient bien en raisonnant toujours par équivalence l'inégalité souhaitée (0.5pt).

> Je ne détaille pas ici ces opérations (tout de même attendues !), vous les retrouverez dans le cours.

Exercice 2 (8.75pt)

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 $H \in \mathcal{M}_{M,N}(\mathbb{R})$ et $D \in \mathcal{M}_{P,N}(\mathbb{R})$, $y \in \mathbb{R}^M$ un vecteur fixé et $\lambda, \mu > 0$ deux constantes.

1. a) Rappeler la définition d'une fonctionnelle quadratique. *On précisera bien la dimension des matrices et vecteurs impliquées.*

   b) Montrer que f est une fonctionnelle quadratique. *On vérifiera la cohérence des objets mathématiques entre eux à travers les dimensions des matrices et vecteurs impliquées.*

   c) Exprimer $\nabla f(x)$, en fonction des données du problème, pour tout $x \in \mathbb{R}^N$?

   d) Exprimer $\nabla^2 f(x)$, en fonction des données du problème, pour tout $x \in \mathbb{R}^N$?
2. Démontrer que f est strictement convexe.
3. Démontrer que f admet un unique minimum global.
4. Montrer que les hypothèses du théorème de convergence de la méthode de descente de gradient à pas optimal sont satisfaites.
5. On note $(x_n)_{n\in\mathbb{N}}$ la suite produite par l'algorithme de descente de gradient à pas optimal appliqué à f.

   a) Que pouvez-vous dire de la convergence de $(x_n)_{n\in\mathbb{N}}$?

   b) Que dire, pour tout $n \in \mathbb{N}$, des vecteurs $\Delta x_n = x_{n+1} - x_n$ et $\Delta x_{n+1} = x_{n+2} - x_{n+1}$?
6. Écrire en Python une fonction codant l'algorithme de descente du gradient à pas optimal pour la fonction f. Elle prendra en argument les différents paramètres du problème : H, D, y, λ, μ, ainsi que d'autres éventuelles paramètres que vous jugerez pertinent et renverra la suite des itérés $(x_n)_{n\in\mathbb{N}}$.

Correction.
8.75 = (0.75 + 1.5 + 0.5 + 0.25) + 1 + 0.75 + 1 + (0.5 + 0.5) + 2

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 $c \in \mathbb{R}$ (0.75pt).

   b) En écrivant les normes euclidiennes au carré sous forme de produits scalaires, en développant et passant à la transposée (0.5pt), on montre que f est une fonctionnelle quadratique avec $A = H^TH + 2\lambda D^TD + 2\mu I_N$, qui est bien une matrice symétrique (0.25pt), $b = H^Ty$ et $c = \frac{1}{2}\|y\|_2^2$ (0.25pt). On remarque que H^TH, D^TD sont des matrices de tailles $N \times N$, donc A est bien comme attendue une matrice de taille $N \times N$. Puis H^Ty est un vecteur de taille N (0.5pt).

   c) D'après le cours, on sait que pour tout $x \in \mathbb{R}^N$, $\nabla f(x) = Ax - b$, avec ici $A = H^TH + 2\lambda D^TD + 2\mu I_N$ et $b = H^Ty$ (0.5pt).

   d) D'après le cours, on sait que pour tout $x \in \mathbb{R}^N$, $\nabla^2 f(x) = A$, avec ici $A = H^TH + 2\lambda D^TD + 2\mu I_N$ (0.25pt).
2. Soit $x \in \mathbb{R}^N$, on a pour tout $h \in \mathbb{R}^N$, $h^T\nabla^2 f(x)h = \underbrace{\|Hh\|_2^2 + 2\lambda\|Dh\|_2^2}_{\geqslant 0} + 2\mu\|h\|_2^2 \geqslant 2\mu\|h\|_2^2$. D'où $h^T\nabla^2 f(x)h > 0$ dès que $h \neq 0$, i.e. $\nabla^2 f(x)$ est définie positive. Comme c'est vrai pour tout $x \in \mathbb{R}^N$, f est strictement convexe sur $\mathbb{R}^N$ (1pt).
3. La fonction f admet un point critique (c'est $A^{-1}b$) puisque A est inversible, et ce point critique est un minimum globale comme f est convexe. Enfin ce minimum global est unique puisque f est strictement convexe (0.75pt).

> Il y a plusieurs manière de justifier cette question. On pouvait aussi dire que comme f est convexe, alors l'ensemble de ses minimums globaux est inclu dans l'ensemble de ses points critiques et comme celui-ci est réduit à un singleton contenant l'élément $A^{-1}b$, f admet donc un unique minimum global. Sinon on pouvait aussi dire que f admet un minimum global car est continue et coercive (à montrer !). Puis conclure qu'il est unique par strict convexité de nouveau.

4. La fonction f est $\mathcal{C}^2$ sur $\mathbb{R}^N$ et on a pour tout $x \in \mathbb{R}^N$, $2\mu I_N \preceq \nabla^2 f(x) \preceq (\lambda_{\max}(H^T H)) + 2\lambda\lambda_{\max}(D^T D) + 2\mu) I_N$, où $\lambda_{\max}(A)$ correspond à la plus grande valeur propre de la matrice symétrique A. Ainsi on sait d'après le cours que cela implique que les hypothèses du théorème de convergence de l'algorithme de la descente de gradient à pas optimal sont satisfaites (Corollaire 3.20) (1pt).

> Pour rappel les hypothèses du résultat de convergence demande que la fonction soit $\mathcal{C}^2$ et convexe sur $\mathbb{R}^N$, et que sur l'ensemble $C = \{x \in \mathbb{R}^N : f(x) \leqslant f(x^{(0)})\}$, qui doit être compact, f vérifie $mI_N \preceq \nabla^2 f(x) \preceq MI_N$ avec $0 < m \leqslant M$. Toutes ces conditions sont bien vérifiées ici. Celle qui n'est pas triviale est C compact, mais on sait que f fortement convexe implique f coercive donc l'ensemble C est forcément bornée (et fermée par continuité de f).

5. a) Cette question n'était pas très bien posée voire induisait en erreur, car elle sous entendait de parler de la vitesse de convergence de la suite $(x_n)_{n\in\mathbb{N}}$, ce qui n'a pas été étudiée. Nous avons plutôt montrer la convergence linéaire de la suite $(f(x_n))_{n\in\mathbb{N}}$.

   La suite converge vers l'unique minimum global de f (0.25pt) et la suite $(f(x_n))_{n\in\mathbb{N}}$ converge linéairement vers la valeur minimale de f sur $\mathbb{R}^N$. (0.25pt).

   b) On sait pour la méthode de la descente de gradient à pas optimal que les gradients successifs sont orthogonaux, donc pour tout $n \in \mathbb{N}$, $\Delta x_n \perp \Delta x_{n+1}$. Si on relit les itérés successifs formant la suite $(x_n)_{n\in\mathbb{N}}$ par des segments, alors on se retrouve avec une trajectoire linéaire par morceaux avec des angles droits à chaque changement de direction (0.5pt).

6. 0.5 : mise en forme globale du code Python,
   0.25 : définition variables externes,
   0.25 : initialisation de l'algorithme,
   0.25 : condition d'arrêt,
   0.25 : écriture nouvel itéré,
   0.5 : cohérence logique de l'algorithme (ordre...).

Exercice 3 (19.5pt)

Partie I : un résultat général d'existence de 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 dans $\mathbb{R}^N$.

1. Justifier que $p = \inf_{x\in U} f(x)$ existe ($p > -\infty$).
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 $f(x_n) \underset{n\to+\infty}{\longrightarrow} p$.
3. Démontrer qu'il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, on a $x_n \in f^{-1}(]-\infty, p+1])$.
4. En déduire que le problème $\inf_{x\in U} f(x)$ admet au moins une solution.

Partie II : barrière logarithmique.

On considère la fonction

$$f : x \mapsto -\sum_{i=1}^{m} \ln(b_i - \langle a_i, x\rangle),$$

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

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

1. Questions préliminaires.

   a) On suppose dans cette question et les trois suivantes que $N = 2$. Représenter sur votre feuille en trois figures successives, les ensembles $E_1 = \{x \in \mathbb{R}^2 : \langle e_1, x\rangle = 0\}$, $F_1 = \{x \in \mathbb{R}^2 : \langle e_1, x\rangle = 1\}$ et $C_1 = \{x \in \mathbb{R}^2 : 1 - \langle e_1, x\rangle > 0\}$, où $e_1 = (1, 0)$ est le premier vecteur de la base canonique de $\mathbb{R}^2$.

   b) Représenter l'ensemble $C_2 = \{x \in \mathbb{R}^2 : 1 - \langle e_2, x\rangle > 0\}$.

   c) Finalement représenter l'ensemble $U = C_1 \cap C_2$. Trouver un vecteur $v \in \mathbb{R}^2 \setminus \{0\}$ tel que $\langle e_1, v\rangle \leqslant 0$ et $\cancel{\langle e_1, v\rangle \leqslant 0}$ $\langle e_2, v\rangle \leqslant 0$. Représenter ce vecteur sur votre figure ainsi que l'ensemble $\mathbb{R}_+^* v$. Montrer que $\mathbb{R}_+^* v \subset U$. L'ensemble U est-il borné ?

   d) Soit $C_1, C_2, \dots, C_m$ des sous ensembles convexes de $\mathbb{R}^N$. Montrer alors que $U = \cap_{i=1}^m C_i$ est également un convexe de $\mathbb{R}^N$.

2. a) Donner l'ensemble de définition de f, noté U. Montrer que U est un ouvert de $\mathbb{R}^n$.

   b) Montrer 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 $h \in \mathbb{R}^N$
$$df(x)(h) = \sum_{i=1}^{m} \frac{\langle a_i,\, h\rangle}{b_i - \langle a_i,\, x\rangle}.$$
 b) En déduire $\nabla f(x)$ pour tout $x \in \mathbb{R}^N$.

5. a) Démontrer que pour tout $x \in U$, et tout $(h,k) \in \mathbb{R}^N \times \mathbb{R}^N$
$$d^2 f(x)(h,k) = \sum_{i=1}^{m} \frac{\langle a_i,\, h\rangle \langle a_i,\, k\rangle}{(b_i - \langle a_i,\, x\rangle)^2}.$$
 b) En déduire $\nabla^2 f(x)$ pour tout $x \in \mathbb{R}^N$.
 c) Démontrer que f est strictement convexe sur U.

6. a) Soit $x, x_0 \in U$. Justifier que
$$f(x) \geqslant f(x_0) + \left\langle \sum_{i=1}^{m} \frac{1}{b_i - \langle a_i,\, x_0\rangle} a_i,\, x - x_0 \right\rangle.$$
 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.

7. Bonus : On admet que U est borné. Déduire alors de l'ensemble des questions que f admet un unique minimum sur U.

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

Correction.

Partie I. 2.5 = 0.5 + 0.5 + 0.5 + 1

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. $p > -\infty$ par la propriété de la borne inférieure (0.5pt).

2. D'après la caractérisation de la borne inférieure : pour tout $\varepsilon > 0$, il existe $x \in U$ tel que $0 \leqslant f(x) - p < \varepsilon$. 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 $f(x_n) \underset{n\to+\infty}{\longrightarrow} p$ (0.5pt).

3. Puisque $f(x_n) \underset{n\to+\infty}{\longrightarrow} p$, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, on a $0 \leqslant f(x_n) - p \leqslant 1$, c'est-à-dire $f(x_n) \leqslant p+1$, ou encore $x_n \in f^{-1}(]-\infty, p+1])$ (0.5pt).

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 $\mathbb{R}^N$ (0.5pt). 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 (0.5pt).

Partie II. 17 = (0.75 + 0.25 + 1.25 + 0.5) + (0.5 + 0.5) + 0.5 + (1.5 + 0.75) + (1.5 + 1 + 0.75) + (1.25 + 0.5) + 3.5 + 2

1. a) (0.75pt)
 b) (0.25pt)
 c) (0.5pt) pour les représentations.
 Le vecteur $v = (-1,-1)$ satisfait $\langle e_1,\, v\rangle = -1 < 0$ et $\langle e_2,\, v\rangle = -1 < 0$ (0.25pt). Soit $t > 0$, on a $1 - \langle e_1,\, tv\rangle = 1 + t > 0$, donc $tv \in C_1$. De même $1 - \langle e_2,\, tv\rangle = 1 + t > 0$, i.e. $tv \in C_2$. On a donc bien $\mathbb{R}_+^* v \subset U$ (0.5pt).
 d) Soit $x, y \in U$ et $t \in [0,1]$. Alors pour tout $i \in \{1,\dots,m\}$, $x, y \in C_i$ et donc par convexité de C_i, on a $(1-t)x + ty \in C_i$. D'où $(1-t)x + ty \in U$, i.e. U est convexe (0.5pt).

2. a) On a $U = \{x \in \mathbb{R}^N : \forall i \in \{1,\dots,m\},\ \langle a_i,\, x\rangle < b_i\} = \cap_{i=1}^{m} \underbrace{\{x \in \mathbb{R}^N :\ b_i - \langle a_i,\, x\rangle > 0\}}_{=C_i}$. Pour tout i, C_i est ouvert (car la condition le définissant est ouverte) donc U est ouvert par intersection finie d'ouverts (0.5pt).
 b) Pour tout i, C_i est convexe, donc d'après une précédente question, U est convexe (0.5pt).

3. Chacunes 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 (0.5pt)

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, $df(x)(h) = \sum_{i=1}^m \frac{\langle a_i, h\rangle}{b_i - \langle a_i, x\rangle}$ (cette expression est vraie ici pour tout $h \in \mathbb{R}^N$) (1.5pt).

   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 $h \in \mathbb{R}^N$, $df(x)(h) = \langle \nabla f(x), h\rangle$. Donc par identification, on a pour tout $x \in U$, $\nabla f(x) = \sum_{i=1}^m \frac{1}{b_i - \langle a_i, x\rangle} a_i$ (0.75pt).

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 $df(x+k)$ (puisque la différentielle seconde de f est la différentielle de $x \mapsto df(x)$). On a $df(x+k) = \sum_{i=1}^m \frac{\langle a_i, \cdot\rangle}{b_i - \langle a_i, x+k\rangle} = \sum_{i=1}^m \langle a_i, \cdot\rangle \tilde{g}_i(\langle a_i, x\rangle + \langle a_i, k\rangle)$, 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, $df(x+k) = df(x) + \sum_{i=1}^m \langle a_i, \cdot\rangle \frac{\langle a_i, k\rangle}{(b_i - \langle a_i, x\rangle)^2} + o(\|k\|_2)$. D'où par définition, pour tout $h, k \in \mathbb{R}^N$, $d^2 f(x)(h,k) = \sum_{i=1}^m \langle a_i, h\rangle \frac{\langle a_i, k\rangle}{(b_i - \langle a_i, x\rangle)^2}$ (1.5pt).

   b) D'après la précédente question, on a pour tout $x \in \mathbb{R}^N$ et tout $h, k \in \mathbb{R}^N$

$$
\begin{aligned}
d^2 f(x)(h,k) &= \sum_{i=1}^m \langle a_i, h\rangle \frac{\langle a_i, k\rangle}{(b_i - \langle a_i, x\rangle)^2} = \sum_{i=1}^m \frac{1}{(b_i - \langle a_i, x\rangle)^2} \langle a_i, h\rangle \langle a_i, k\rangle, \\
&= \sum_{i=1}^m \frac{1}{(b_i - \langle a_i, x\rangle)^2} \underbrace{a_i^T h}_{=h^T a_i} \; a_i^T h = \sum_{i=1}^m \frac{1}{(b_i - \langle a_i, x\rangle)^2} h^T \underbrace{a_i a_i^T}_{\in \mathcal{M}_N(\mathbb{R})} h, \\
&= h^T \left(\underbrace{\sum_{i=1}^m \frac{1}{(b_i - \langle a_i, x\rangle)^2} a_i a_i^T}_{\in \mathcal{M}_N(\mathbb{R})} \right) h.
\end{aligned}
$$

   Donc par définition, $\nabla^2 f(x) = \sum_{i=1}^m \frac{1}{(b_i - \langle a_i, x\rangle)^2} a_i a_i^T$ (1pt).

   c) On a pour tout $x \in U$ et tout $h \in \mathbb{R}^N \setminus \{0\}$, $d^2 f(x)(h,h) = \sum_{i=1}^m \frac{\langle a_i, h\rangle^2}{(b_i - \langle a_i, x\rangle)^2} > 0$ (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 (0.75pt).

6. a) f est une fonction convexe différentiable sur l'ouvert convexe U donc pour tout $x, x_0 \in U$, on a $f(x) \geqslant f(x_0) + \langle \nabla f(x_0), x - x_0\rangle$ (0.5pt). On obtient l'inégalité souhaitée en utilisant l'expression de $\nabla f(x_0)$ trouvée précédemment (0.25pt).

   Pour la deuxième partie de la question, on a

$$
\begin{aligned}
f(x_0) + \left\langle \sum_{i=1}^m \frac{1}{b_i - \langle a_i, x_0\rangle} a_i, x - x_0 \right\rangle &= f(x_0) + \sum_{i=1}^m \frac{\langle a_i, x - x_0\rangle}{b_i - \langle a_i, x_0\rangle}, \\
&= f(x_0) - \sum_{i=1}^m \frac{-b_i + b_i + \langle a_i, x_0 - x\rangle}{b_i - \langle a_i, x_0\rangle}, \\
&= f(x_0) + m - \sum_{i=1}^m \frac{b_i - \langle a_i, x\rangle}{b_i - \langle a_i, x_0\rangle} \text{ (0.5pt)}.
\end{aligned}
$$

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

$$
f(x) \geqslant f(x_0) + m - \sum_{i=1}^m \frac{b_i + \|a_i\|_2 C}{b_i - \langle a_i, x_0\rangle} \text{ (0.5pt)}.
$$

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.*

7. 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 $x \in \mathbb{R}^N$, il suffit de montrer que $x \in V_\alpha$.

> La subtilité et difficulté principale consiste à penser (et à réussir) à montrer que la limite x de la suite est dans U. Si ce n'était pas le cas, on ne pourrait pas écrire $f(x)$ et appliquer la caractérisation séquentielle de la continuité comme fait ci-dessous. Montrons donc d'abord ce point. L'idée est de remarquer que U est un ouvert donc si x est soit dans U soit sur sa frontière, qui est incluse dans l'union des ensemble $\{x \in \mathbb{R}^N : b_i - \langle a_i,\, x\rangle = 0\}$.

Montrons d'abord que $x \in U$. On a par défaut $x \in \overline{U} = \overset{\circ}{U} \cup \mathrm{Fr}(U)$, où $\mathrm{Fr}(U)$ dénote la frontière de U et l'union est disjointe. On a donc soit $x \in \overset{\circ}{U}$, soit $x \in \mathrm{Fr}(U)$. Supposons par l'absurde que $x \in \mathrm{Fr}(U)$, alors il existe $i \in \{1, \dots, m\}$, tel que $b_i - \langle a_i,\, x\rangle = 0$, et donc nécessairement $f(x_n) \underset{n\to+\infty}{\longrightarrow} +\infty$, ce qui entre en contradiction avec le fait que pour tout $n \in \mathbb{N}$, $f(x_n) \leqslant \alpha$. D'où $x \in \overset{\circ}{U}$. Mais comme U est ouvert, on a $\overset{\circ}{U} = U$, ainsi $x \in U$. Par conséquent, par la caractérisation séquentielle de la continuité, on déduit que $f(x_n) \underset{n\to+\infty}{\longrightarrow} f(x)$. Mais comme pour tout $n \in \mathbb{N}$, $f(x_n) \leqslant \alpha$, par passage à la limite dans l'inégalité, on a $f(x) \leqslant \alpha$, i.e. $x \in V_\alpha$.

Ainsi d'après Partie I, f admet un minimum global sur U. Comme f est strictement convexe, ce minimum global est unique (3.5pt).

8. 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 suppqoser 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 $\left\|x_{\varphi(n)}\right\|_2 \to +\infty$, on obtient que pour tout $i \in \{1, \dots, m\}$, $\langle a_i,\, v\rangle \leqslant 0$. Ceci contredit l'hypothèse de l'énoncé sur les a_i (car $v \neq 0$). Donc U est borné (2pt).