TD2 : Optimisation

Exercice 1.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = 4(x^2+y^2) - (x^2+y^2)^2$.

1. Calculer le gradient $\nabla f(x,y)$ et la matrice hessienne $H_f(x,y)$, pour tout $(x,y) \in \mathbb{R}^2$.
2. Montrer que l'ensemble des éléments de $\mathbb{R}^2$ tels que $\nabla f(x,y) = 0$ est l'union d'un point P et d'un cercle $\mathcal{C}$.
3. Pour P et les points de $\mathcal{C}$, dire s'il s'agit de minimiseurs locaux, maximiseurs locaux ou si on ne peut pas trancher.

Exercice 2.
Soit la fonction
$$f(x,y,z) = x^2 - xy + y^2 - yz + z^2.$$
On cherche à déterminer de façon approchée le minimum de f sur $\mathbb{R}^3$.

1. Calculer le gradient $\nabla f(X)$ et la matrice hessienne $H_f(X)$, pour tout $X = (x,y,z) \in \mathbb{R}^3$.
2. Montrer que f est strictement convexe.
3. Montrer que $\lim_{\|X\|\to+\infty} f(X) = +\infty$.
4. Donner l'algorithme de la méthode du gradient à pas fixe pour la minimisation de la fonction f.
5. Donner une condition sur le pas α pour que la méthode de gradient à pas constant converge.
6. Donner l'algorithme de la méthode du gradient à pas optimal. De quelle équation, le pas α_n calculé à chaque itération est-il solution ?

Exercice 3.
Soit $X \in \mathbb{R}^3$, et soit $\{A_i\}_{1\leq i\leq n}$ un ensemble de points de $\mathbb{R}^3$. On définit
$$F_i(X) = \|X - A_i\|^2 \quad \text{et } \Lambda(X) = \sum_{i=1}^{n} F_i(X).$$

1. (a) Montrer que, pour tout $1 \leq i \leq n$, F_i est différentiable sur $\mathbb{R}^3$ et calculer son gradient ∇F_i.
   (b) En déduire le gradient de Λ.
2. (a) Montrer que F_i est une fonction strictement convexe et que $\displaystyle\lim_{\|X\|\to+\infty} F_i(X) = +\infty$.
   (b) En déduire que Λ est strictement convexe et que $\displaystyle\lim_{\|X\|\to+\infty} \Lambda(X) = +\infty$.
   (c) Montrer que Λ admet un unique minimiseur $X^\star \in \mathbb{R}^3$. Caractériser $X^\star$ en fonction des A_i.
3. On cherche à déterminer le minimum de Λ par une méthode de descente de gradient.
   (a) Expliciter la méthode de descente de gradient à pas fixe.
   (b) Montrer que, pour tout X, $Y \in \mathbb{R}^3$
   $$\|\nabla\Lambda(X) - \nabla\Lambda(Y)\| \leq 2n\|X - Y\|.$$
   (c) Montrer que, pour tout X, $Y \in \mathbb{R}^3$
   $$(\nabla\Lambda(X) - \nabla\Lambda(Y)) \cdot (X - Y) \geq 2n\|X - Y\|^2.$$

(d) En déduire un critère sur le pas pour que la méthode de descente de gradient à pas fixe converge.

4. Expliciter la méthode de descente de gradient à pas optimal.

Exercice 4.

Soit $f \in \mathcal{C}^1(\mathbb{R}^N; \mathbb{R})$. On considère une méthode de descente de gradient à pas constant $\alpha > 0$ donnée par :

$$\begin{cases} x_0 \in \mathbb{R}^N, \\ x_{n+1} = x_n - \alpha \nabla f(x_n). \end{cases} \tag{1}$$

On suppose que le gradient de f est globalement lipschitzien de constante $\gamma > 0$.

1. Soit $0 < \beta < 1$. Montrer que, si $\alpha < \dfrac{2(1-\beta)}{\gamma}$, alors pour tout $x \in \mathbb{R}^N$
$$f(x - \alpha \nabla f(x)) < f(x) - \beta\alpha \|\nabla f(x)\|^2.$$
2. En déduire que, si $\alpha < \dfrac{2(1-\beta)}{\gamma}$, la suite (x_n) définie par (1) vérifie l'une des deux propriétés suivantes : soit
$$\lim_{n\to+\infty} f(x_n) = -\infty$$
soit
$$\lim_{n\to+\infty} \nabla f(x_n) = 0.$$

Exercice 5.

Soient a et b deux réels. On considère la fonctionnelle $f : \mathbb{R}^2 \mapsto \mathbb{R}$ donnée par

$$f(x, y) = ax^2 + by^2.$$

1. Sous quelles hypothèses sur a et b la fonction f est-elle strictement convexe ?
2. En déduire, sous ces hypothèses, que le fonction f admet une unique minimiseur sur $\mathbb{R}$.
3. On note $X_n = (x_n, y_n)^t$ le vecteur donné par la méthode du gradient à pas optimal au rang n. La suite (X_n) est ainsi donnée par :
$$X_{n+1} = X_n - \alpha_n \nabla f(X_n), \ \text{ pour tout } n \in \mathbb{N}$$
où α_n minimise $\alpha \mapsto f(X_n - \alpha \nabla f(X_n))$. Exprimer x_{n+1} et y_{n+1} en fonction de x_n et y_n. En déduire que si $x_0 \neq 0$ et $y_0 \neq 0$, alors, pour tout $n \in \mathbb{N}$, $x_n \neq 0$ et $y_n \neq 0$.
4. En déduire qu'en général la méthode du gradient à pas optimal ne peut pas converger en un nombre fini d'itérations.
5. Calculer les deux premières itérations de la méthode du gradient conjugué. Que peut-on conclure ?

Exercice 6.

Soit A une matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ avec $n > m$ telle que A soit injective. Pour un vecteur $b \in \mathbb{R}^n$, on considère la fonction $f : \mathbb{R}^m \to \mathbb{R}$ définie par $f(x) = \|Ax - b\|_2^2$.

1. Montrer que f est coercive.
2. Calculer le gradient $\nabla f(x)$ et la matrice hessienne $H_f(x)$, pour tout $x \in \mathbb{R}^m$.
3. Montrer que, tout point $\overline{x} \in \mathbb{R}^m$ tel que
$$f(\overline{x}) = \min_{x \in \mathbb{R}^m} f(x) \tag{2}$$
satisfait
$$A^t A \overline{x} = A^t b.$$
4. Déduire des questions précédentes l'existence et l'unicité du point $\overline{x}$ qui satisfait (2).

Remarque : cet exercice reprend la méthode des moindres carrés.