Examen du cours d'optimisation différentiable

Durée : 1 heure 30
Les documents ainsi que les calculatrices ne sont pas autorisés.

****

Exercice 0.1 (Différentielle et gradient)
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application différentiable. Soit $u \in \mathbb{R}^n$.

a) Donner la définition de la différentielle de f en u et montrer qu'elle est unique.

b) Donner la définition du gradient de f en u à partir de sa différentielle.

****

Correction de l'exercice 0.1 **a)** `[1pt]` La différentielle de f en u est une forme linéaire continue sur $\mathbb{R}^n$ telle que quand $h \in \mathbb{R}^n$ et $h \to 0$, on a

$$f(u+h) = f(u) + df_u(h) + o(\|h\|_2).$$

(On prend ici la norme euclidienne mais n'importe quelle autre norme convient aussi vu qu'en dimension finie toutes les normes sont équivalentes). Si df_u^* est une autre forme linéaire vérifiant cette propriété alors pour tout $x \in \mathbb{R}^n$ quand $\lambda \in \mathbb{R}$ et $\lambda \to 0$, on a

$$df_u(\lambda x) = df_u^*(\lambda x) + o(\lambda)$$

et donc par linéarité, quand $\lambda \to 0$

$$df_u(x) = df_u^*(x) + o(1).$$

Alors en faisant tendre λ vers 0 dans la dernière égalité, on obtient $df_u(x) = df_u^*(x)$ et ceci pour tout $x \in \mathbb{R}^n$, il y a donc bien unicité de la différentielle quand elle existe.

b) `[1pt]` Comme df_u est une forme linéaire continue sur l'espace de Hilbert $\mathbb{R}^n$, on en déduit par le théorème de Riesz qu'il existe un vecteur noté $\nabla f(u)$ tel que pour tout $h \in \mathbb{R}^n$, $df_u(h) = \langle \nabla f(u), h \rangle$. Ce vecteur est unique et c'est ainsi qu'on définit le gradient de la fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ en u.

****

Exercice 0.2
Soit $f : (x,y) \in \mathbb{R}^2 \to x^4 + y^3/3 - 4y - 2$.

a) Déterminer les points critiques de f

b) Parmi ces points critiques, déterminer les minima locaux de f.

c) Montrer que ces minima locaux ne sont pas globaux

Mêmes questions avec la fonction $g : (x, y) \in \mathbb{R}^2 \to 3x^3 + xy^2 - 2xy$**.**

****

Correction de l'exercice 0.2 **a)** `[1pt]` En tant que polynôme, f est de classe $\mathcal{C}^\infty$. Les points critiques de f sont les éléments $(x, y) \in \mathbb{R}^2$ tels que $\nabla f(x, y) = 0$ càd tels que $4x^3 = 0$ et $y^2 - 4 = 0$. Il y a donc deux points critiques $(0, 2)$ et $(0, -2)$.

b) `[2pts]` Pour savoir si un point critique est un minimum local de f, on peut utiliser la condition du second ordre. La Hessienne de f en (x, y) est donnée par

$$\nabla^2 f(x, y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & 2y \end{pmatrix}.$$

En particulier la Hessienne de f en $(0, -2)$ admet 0 et -4 comme valeurs propres. Donc $(0, -2)$ ne peut pas être un mimum local de f. Pour $(0, 2)$, la Hessienne de f en $(0, 2)$ admet pour valeurs propres 0 et 4, elle est donc positive mais pas définie positive. On ne peut pas conclure sur la minimalité locale de f en $(0, 2)$ par la condition du second ordre. Cependant, on voit que f est la somme de deux fonctions $f(x, y) = g(x) + h(y)$ avec $g(x) = x^4$ et $h(y) = y^3/3 - 4y - 2$. Pour chercher les minima locaux de f il suffit de chercher ceux de g et h. Pour g, 0 est le seul minimum local et pour h (qui est un polynôme de degré 3), seul 2 est un minimum locale (on a $h''(2) = 4 > 0$). Donc $(0, 2)$ est bien un minimum local de f.

c) `[1pt]` Ce n'est pas un minimum global car $f(0, y) \to -\infty$ quand $y \to -\infty$.

d) `[1pt]` Pour tout $(x, y) \in \mathbb{R}^2$, on a

$$\nabla g(x, y) = \begin{pmatrix} 9x^2 + y^2 - 2y \\ 2xy - 2x \end{pmatrix} \text{ et } \nabla^2 g(x, y) = \begin{pmatrix} 18x^2 & 2y - 2 \\ 2y - 2 & 2x \end{pmatrix}.$$

Les points critiques (x, y) de f satisfont $\nabla f(x, y) = 0$. On obtient 4 solutions

$$(0, 0), (0, 2), (1/3, 1) \text{ et } (-1/3, 1).$$

e) `[2pts]` La Hessienne de f en ces points critiques est

$$\begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 2/3 \end{pmatrix} \text{ et } \begin{pmatrix} 2 & 0 \\ 0 & -2/3 \end{pmatrix}.$$

La première matrice a pour valeurs propres 2 et -2 donc $(0, 0)$ n'est pas un minimum local de g. La deuxième matrice a pour valeurs propres 2 et -2 donc $(0, 2)$ n'est pas un minimum local. La troisième matrice a deux valeurs propres strictement positives (car leur produit vaut $4/3$ et leur somme $2 + 2/3$) donc $(1/3, 1)$ est un minimum local. La quatrième matrice a deux valeurs propres non nulles de signes opposés (leur produit vaut $-4/3$) donc $(-1/3, 1)$ n'est pas un minimum local de g.

La fonction g a donc un seul minimum local qui est atteint en $(1/3, 1)$.

f) `[1pt]`Par ailleurs, $f(x, 0) \to -\infty$ quand $x \to -\infty$ donc g n'a pas de minimum global. En particulier, $(1/3, 1)$ est seulement un minimum local sans être global.

****

Exercice 0.3

Soient $p_1, \cdots, p_n$ des points de $\mathbb{R}^d$. On définit la fonction

$$f : x \in \mathbb{R}^d \to \sum_{i=1}^{n} \|x - p_i\|_2^2$$

où $\|\cdot\|_2$ est la norme euclidienne de $\mathbb{R}^d$.

a) Montrer que f admet un unique minimum sur $\mathbb{R}^d$,

b) Déterminer ce minimum et en donner une interprétation géométrique.

****

Correction de l'exercice 0.3 **a)** `[2pt]` La fonction f est un polynôme du second ordre, elle est donc $\mathcal{C}^\infty$. On a pour tout $x \in \mathbb{R}^d$,

$$\nabla f(x) = \sum_{i=1}^{n} 2(x - p_i) \text{ et } \nabla^2 f(x) = 2I_d.$$

Comme $\nabla f(x) \succeq 2I_d$, f est 2-convexe, en particulier, elle est coercive sur $\mathbb{R}^d$ donc admet un minimum sur $\mathbb{R}^d$ et strictement convexe donc ce minimum est unique.

b) `[2pt]` Il est de plus un point critique de f, càd tel que $\nabla f(x) = 0$ càd $x = (1/n)\sum_{i=1}^n p_i$ est l'unique minimum de f sur $\mathbb{R}^d$. C'est la moyenne/barycentre des $p_1, \cdots, p_n$.

****

Exercice 0.4

On considère trois points M_1, M_2 et M_3 du plan de coordonnées (x_1, y_1), (x_2, y_2) et (x_3, y_3). On considère aussi la fonction

$$f : (a, b) \in \mathbb{R}^2 \to \sum_{i=1}^{3} (y_i - ax_i - b)^2.$$

On suppose que les $x_i, i = 1, 2, 3$ ne sont pas tous égaux.

a) Montrer que f est coercive et strictement convexe. En déduire que f admet un unique minimum sur $\mathbb{R}^2$.

b) Montrer que f n'admet qu'un seul point critique.

c) Résoudre le problème de minimisation de f sur $\mathbb{R}^2$ et donner une interprétation géométrique de cette solution en fonction des points M_1, M_2 et M_3.

****

Correction de l'exercice 0.4 **a)** `[4pts]` On note

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \text{ et } \mathbb{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix}.$$

On a $f : z \in \mathbb{R}^2 \to \|Y - \mathbb{X}z\|_2^2$ où $\|\cdot\|_2$ est la norme euclidienne de $\mathbb{R}^3$.

On a pour tout $z \in \mathbb{R}^2$,

$$\nabla f(z) = 2\mathbb{X}^\top(\mathbb{X}z - Y) \text{ et } \nabla^2 f(z) = 2\mathbb{X}^\top\mathbb{X}$$

Comme les x_i ne sont pas tous égaux, $\mathbb{X}^\top$ est de rang 2. Par ailleurs, $\mathrm{Im}(\mathbb{X}^\top\mathbb{X}) = \mathrm{Im}(\mathbb{X}^\top)$ donc $\mathbb{X}^\top\mathbb{X}$ est aussi de rang 2 et comme c'est une matrice 2×2, elle est inversible. Par ailleurs, pour tout $z \in \mathbb{R}^2$, $\langle z, \mathbb{X}^\top\mathbb{X}z\rangle = \|\mathbb{X}z\|_2^2 \geq 0$. Alors $\mathbb{X}^\top\mathbb{X}$ est inversible et positive donc elle est définie positive. Si on note $\sigma_1 \geq \sigma_2$ les valeurs propres de $\mathbb{X}^\top\mathbb{X}$, on a $\sigma_2 > 0$ et $\nabla^2 f(z) \succeq \sigma_2 I_2$. Donc f est σ_2-convexe. Elle est donc coercive sur $\mathbb{R}^2$ et strictement convexe. Donc f admet un unique minimum sur $\mathbb{R}^2$.

b) `[1pt]` $z \in \mathbb{R}^2$ est un point critique de f ssi $\nabla f(z) = 0$ càd $\mathbb{X}^\top\mathbb{X}z = \mathbb{X}^\top Y$. Comme $\mathbb{X}^\top\mathbb{X}$ est inversible, on a $z = (\mathbb{X}^\top\mathbb{X})^{-1}\mathbb{X}^\top Y$. Il y a donc bien unicité du point critique de f.

c) `[1p]` $z = (\mathbb{X}^\top\mathbb{X})^{-1}\mathbb{X}^\top Y$ est l'unqiue solution au problème de minimisation de f sur $\mathbb{R}^2$ au vu des deux questions précédentes. C'est la solution des moindres carrés de la droite qui minimise les résidus aux points M_1, M_2 et M_3.