23

Extrema des fonctions réelles de *n* variables

ECG Maths Approfondies
Semestre 4

« L'objectif est présenter la démarche de recherche d'extrema et d'en acquérir une maîtrise raisonnable à partir d'outils théoriques. L'espace $\mathbb{R}^n$ sera muni [de son produit scalaire canonique et] de la norme euclidienne usuelle. La détermination de la nature topologique d'un ensemble n'est pas un objectif du programme ; elle devra toujours être précisée. Néanmoins, il est nécessaire de sensibiliser les étudiants aux notions d'ouverts et de fermés » (extrait du programme officiel).
Dans tout le cours, n désigne un entier naturel non nul.

A. Notions de topologie

Définition 23.1

Soit $\mathcal{O}$ une partie de $\mathbb{R}^n$. On dit que $\mathcal{O}$ est une **partie ouverte** ou plus simplement un **ouvert** de $\mathbb{R}^n$ si $\Omega = \varnothing$ ou si :

$$\forall a \in \mathcal{O},\ \exists \alpha \in \mathbb{R}_+^* \ / \ \forall x \in \mathbb{R}^n,\ ||x - a|| < \alpha \Longrightarrow x \in \mathcal{O}$$

Définition 23.2

Soit F une partie de $\mathbb{R}^n$. On dit que F est une **partie fermée** ou plus simplement un **fermé** de $\mathbb{R}^n$ si son complémentaire $\mathcal{O} = \mathbb{R}^n \setminus F$ est ouvert.

Remarques

a. $\varnothing$ et $\mathbb{R}^2$ sont à la fois ouverts et fermés.

b. On pourra être rassuré par la lecture du programme, qui indique clairement : « *la détermination de la nature topologique d'un ensemble [(ouvert, fermé ou aucun des deux)] n'est pas un objectif du programme et devra toujours être indiquée* ». L'introduction des définitions précédentes n'a donc pas d'autre but que de définir le cadre dans lequel la suite du cours va se situer.

c. Pour faire simple (et fort peu rigoureux), on pourra retenir qu'une partie de $\mathbb{R}^2$ est ouverte si elle ne contient aucun des points se situant à sa frontière et qu'une partie est fermée si elle contient tous les points se situant à sa frontière.

Proposition 23.3

Si φ est une fonction continue sur $\mathbb{R}^n$, alors :

i. pour tout réel a, les ensembles

$$\{x \in \mathbb{R}^n \ / \ \varphi(x) < a\} \quad \text{et} \quad \{x \in \mathbb{R}^n \ / \ \varphi(x) > a\}$$

sont ouverts,

ii. pour tout réel a, les ensembles

$$\{x \in \mathbb{R}^n \ / \ \varphi(x) \leqslant a\} \quad \text{et} \quad \{x \in \mathbb{R}^n \ / \ \varphi(x) \geqslant a\}$$

sont fermés.

Exemples 23.1

a. $\{(x,y)\in\mathbb{R}^2 \ / \ x^2+y-xy<2\}$ est une partie ouverte de $\mathbb{R}^2$.

b. Pour tout réel r strictement positif, l'ensemble $\{x\in\mathbb{R}^n \ / \ ||x-a||<r\}$ est ouvert.

c. Pour tout réel r positif ou nul, l'ensemble $\{x\in\mathbb{R}^n \ / \ ||x-a||\leqslant r\}$ est fermé.

Définition 23.4

On dit qu'une partie E de $\mathbb{R}^n$ est **bornée** s'il existe un réel M positif ou nul tel que :

$$\forall x\in E,\ ||x||\leqslant M$$

Exercice 23.1 Dire si les ensembles suivants sont bornés :

$$[0,1]^2\,,\quad E=\left\{(x,y)\in(\mathbb{R}^+)^2 \ / \ x+y\leqslant 1\right\},\quad F=\left\{(x,y)\in\mathbb{R}^2 \ / \ 0\leqslant x-y\leqslant 2\right\}$$

B. Fonctions réelles de *n* variables réelles

B.1. Continuité d'une fonction définie sur une partie de $\mathbb{R}^n$

Définition 23.5

Soit f une fonction définie sur une partie P de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}$. On dit que f est continue sur P si f est continue en tout point de P.

Proposition 23.6

Les fonctions polynômes définies sur $\mathbb{R}^n$ sont continues sur $\mathbb{R}^n$.

Proposition 23.7

Soit f et g deux fonctions continues sur une partie P de $\mathbb{R}^n$.

i. $f+g : x\mapsto f(x)+g(x)$ et $fg : x\mapsto f(x)\,g(x)$ sont continues sur P.

ii. Si g ne s'annule pas sur P, alors $\dfrac{f}{g} : x\mapsto \dfrac{f(x)}{g(x)}$ est continue sur P.

Proposition 23.8

Si f est une fonction continue sur une partie P de $\mathbb{R}^n$, prenant ses valeurs dans un intervalle I de $\mathbb{R}$ et si g est une fonction continue sur I et à valeurs dans $\mathbb{R}$, alors $g\circ f$ est continue sur P.

B.2. Dérivées partielles, gradient

Définition 23.9

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et $x=(x_1,\ldots,x_n)$ un élément de $\mathcal{O}$. Pour tout $i\in[\![1,n]\!]$, la fonction

$$f_{x,i} : t\mapsto f(x_1,\ldots,x_{i-1},t,x_{i+1},\ldots,x_n)$$

est appelée $\boldsymbol{i^{\grave{e}me}}$ **fonction partielle** de f en x ; c'est une fonction d'une variable définie sur

$$\{t\in\mathbb{R} \ / \ (x_1,\ldots,x_{i-1},t,x_{i+1},\ldots,x_n)\in\mathcal{O}\}$$

Définition 23.10

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et $x = (x_1, \ldots, x_n)$ un élément de $\mathcal{O}$. Pour tout $i \in [\![1, n]\!]$, on dit que f admet une $i^{ème}$ **dérivée partielle d'ordre 1** en x si la $i^{ème}$ fonction partielle $f_{a,i}$ est dérivable en x_i et, dans ce cas, on note :

$$\partial_i f(x) = f'_{x,i}(x_i)$$

Pour tout $i \in [\![1, n]\!]$, on dit que f admet une $i^{ème}$ dérivée partielle d'ordre 1 sur $\mathcal{O}$ si f admet une $i^{ème}$ dérivée partielle d'ordre 1 en tout point de $\mathcal{O}$.

Définition 23.11

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et x un élément de $\mathcal{O}$. Si $\partial_i f(x)$ existe pour tout $i \in [\![1, n]\!]$, alors le vecteur de $\mathbb{R}^n$

$$\nabla f(x) = (\partial_1 f(x), \ldots, \partial_n f(x))$$

est appelé **gradient** de f en x.

Définition 23.12

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$. On dit que f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$ si les fonctions $\partial_1 f, \ldots, \partial_n f$ sont définies et continues sur $\mathcal{O}$.

Proposition 23.13

Les fonctions polynômes définies sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ sont de classe $\mathcal{C}^1$ sur $\mathcal{O}$.

Proposition 23.14

Soit f et g deux fonctions de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$.

i. $f + g : x \mapsto f(x) + g(x)$ et $fg : x \mapsto f(x)\, g(x)$ sont de classe $\mathcal{C}^1$ sur $\mathcal{O}$.

ii. Si g ne s'annule pas sur $\mathbb{R}^n$, alors $\dfrac{f}{g} : x \mapsto \dfrac{f(x)}{g(x)}$ est de classe $\mathcal{C}^1$ sur $\mathcal{O}$.

Proposition 23.15

Si f est une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, prenant ses valeurs dans un intervalle I de $\mathbb{R}$ et si g est une fonction de classe $\mathcal{C}^1$ sur I et à valeurs dans $\mathbb{R}$, alors $g \circ f$ est de classe $\mathcal{C}^1$ sur $\mathcal{O}$.

B.3. Développement limité à l'ordre 1

Théorème 23.16

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}$ et $x \in \mathcal{O}$. Il existe une fonction ε définie sur $\mathbb{R}^n$, continue et nulle en 0 telle que :

$$\forall h \in \mathbb{R}^n \ / \ x + h \in \mathcal{O},\ \ f(x+h) = f(x) + \langle \nabla f(x), h \rangle + ||h||\, \varepsilon(h)$$

Cette égalité est appelée **développement limité** de f à l'ordre 1 en x.

B.4. Dérivée directionnelle

Proposition 23.17

Si f est de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors, pour tout $x \in \mathcal{O}$ et tout $h \in \mathbb{R}^n$, la fonction $g : t \mapsto f(x+th)$ est dérivable sur $I = \{t \in \mathbb{R} \ / \ x+th \in \mathcal{O}\}$ et :

$$\forall t \in I, \ g'(t) = \langle \nabla f(x+th), h \rangle$$

Définition 23.18

Si f est de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors, pour tout $h \in \mathbb{R}^n$, le nombre dérivé en 0 de la fonction $g : t \mapsto f(x+th)$ est appelé **dérivée directionnelle** de f en x dans la direction h et noté $\partial_h f(x)$:

$$\partial_h f(x) = \langle \nabla f(x), h \rangle$$

C. Calcul différentiel : ordre 2

C.1. Dérivées partielles d'ordre 2

Définition 23.19

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et $x = (x_1, \ldots, x_n)$ un élément de $\mathcal{O}$. Pour tout $(i,j) \in [\![1,n]\!]^2$, si f admet une $j^{\text{ème}}$ dérivée partielle sur $\mathcal{O}$ et si $\partial_j f$ admet une $i^{\text{ème}}$ dérivée partielle en x, alors on note :

$$\partial^2_{i,j} f(x) = \partial_i(\partial_j f)(x)$$

Les fonctions $\partial^2_{i,j} f$, lorsqu'elles existent, sont appelée dérivées partielles d'ordre 2 de f.

Exercice 23.2 Calculer les dérivées partielles d'ordre 1 et 2 de la fonction $f : (x,y) \mapsto \ln(1+x^2+y^2)+2x^3-4x^2y$.

Définition 23.20

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$. On dit que f est de classe $\mathcal{C}^2$ sur $\mathcal{O}$ si, pour tout $(i,j) \in [\![1,n]\!]^2$, la fonction $\partial^2_{i,j}$ est définie et continue sur $\mathcal{O}$.

Proposition 23.21

Les fonctions polynômes définies sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ sont de classe $\mathcal{C}^2$ sur $\mathcal{O}$.

Proposition 23.22

Soit f et g deux fonctions de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$.

i. $f+g : x \mapsto f(x)+g(x)$ et $fg : x \mapsto f(x)\,g(x)$ sont de classe $\mathcal{C}^2$ sur $\mathcal{O}$.

ii. Si g ne s'annule pas sur $\mathbb{R}^n$, alors $\dfrac{f}{g} : x \mapsto \dfrac{f(x)}{g(x)}$ est de classe $\mathcal{C}^2$ sur $\mathcal{O}$.

Proposition 23.23

Si f est une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, prenant ses valeurs dans un intervalle I de $\mathbb{R}$ et si g est une fonction de classe $\mathcal{C}^2$ sur I et à valeurs dans $\mathbb{R}$, alors $g \circ f$ est de classe $\mathcal{C}^2$ sur $\mathcal{O}$.

Théorème 23.24 ► Théorème de Schwarz

Si f est une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors :

$$\forall (i,j) \in [\![1,n]\!]^2 \,,\ \partial^2_{i,j} f = \partial^2_{j,i} f$$

C.2. Matrice hessienne

Définition 23.25

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et x un élément de $\mathcal{O}$. Si, pour tout $(i,j) \in [\![1,n]\!]^2$, $\partial^2_{i,j} f(x)$ existe, alors la matrice de $\mathcal{M}_n(\mathbb{R})$

$$\nabla^2 f(x) = \left(\partial^2_{i,j} f(x)\right)_{1 \leqslant i,j \leqslant n}$$

est appelé **matrice hessienne** de f en x.

Exemple 23.2 La fonction $f : (x,y) \mapsto 2x^2y + 3xy - y^4$ est une fonction polynôme, donc elle est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ et on a :

$$\forall (x,y) \in \mathbb{R}^2,\ \partial_1 f(x,y) = 4xy + 3y \quad \text{et} \quad \partial_2 f(x,y) = 2x^2 + 3x - 4y^3$$

d'où :

$$\forall (x,y) \in \mathbb{R}^2,\ \begin{cases} \partial^2_{1,1} f(x,y) = 4y \\ \partial^2_{1,2} f(x,y) = 4x + 3 \\ \partial^2_{2,2} f(x,y) = -12y \end{cases}$$

On en déduit que la matrice hessienne de f en $(0,1)$ est :

$$\nabla^2 f(0,1) = \begin{pmatrix} 4 & 3 \\ 3 & -12 \end{pmatrix}$$

Théorème 23.26

Si f est une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors sa matrice hessienne de f est symétrique réelle en tout point de $\mathcal{O}$.

Définition 23.27

Soit f est une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$. Pour tout vecteur h de $\mathbb{R}^n$, on note H la colonne de ses coordonnées dans la base canonique de $\mathbb{R}^n$.
Pour tout $x \in \mathcal{O}$, l'application $q_x : h \mapsto {}^t H \, \nabla^2 f(x) \, H = \left\langle \nabla^2 f(x) \, H, H \right\rangle$ est appelée **forme quadratique** associée à la matrice hessienne de f en x.

C.3. Développement limité à l'ordre 2

Théorème 23.28

Soit f est une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}$ et $x \in \mathcal{O}$. Il existe une fonction ε définie sur $\mathbb{R}^n$, continue et nulle en 0 telle que :

$$\forall h \in \mathbb{R}^n \ / \ x + h \in \mathcal{O},\ f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \, q_x(h) + ||h||^2 \, \varepsilon(h)$$

Cette égalité est appelée **développement limité** de f à l'ordre 2 en x.

Proposition 23.29 ► Unicité du développement limité

Soit f est une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}$ et $x \in \mathcal{O}$. Le développement limité de f à l'ordre 2 en x est unique ; autrement dit, s'il existe un réel α, un vecteur v de $\mathbb{R}^n$, une matrice M de $\mathcal{M}_n(\mathbb{R})$ et une fonction η définie sur $\mathbb{R}^n$, continue et nulle en 0 telle que, en notant H la colonne des coordonnées de h dans la base canonique :

$$\forall h \in \mathbb{R}^n,\ f(x+h) = \alpha + \langle v, h\rangle + \frac{1}{2}\langle MH, H\rangle + ||h||^2\,\eta(h)$$

alors : $\alpha = f(x)$, $v = \nabla f(x)$ et $M = \nabla^2 f(x)$.

C.4. Dérivée seconde directionnelle

Proposition 23.30

Si f est de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors, pour tout $x \in \mathcal{O}$ et tout $h \in \mathbb{R}^n$, la fonction $g : t \mapsto f(x+th)$ est deux fois dérivable sur $I = \{t \in \mathbb{R} \ /\ x+th \in \mathcal{O}\}$ et :

$$\forall t \in I,\ g''(t) = q_{x+th}(h)$$

Définition 23.31

Si f est de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors, pour tout $h \in \mathbb{R}^n$, le nombre dérivé en 0 de la fonction $g : t \mapsto f(x+th)$ est appelé **dérivée seconde directionnelle** de f en x dans la direction h et noté $\partial_h^2 f(x)$:

$$\partial_h^2 f(x) = q_x(h)$$

D. Recherche d'extrema

D.1. Définitions

Définition 23.32

Soit f une fonction définie sur une partie P de $\mathbb{R}^n$ et x_0 un élément de P.

1. On dit que $f(x_0)$ est un **minimum global** de f si :

$$\forall x \in P,\ f(x) \geqslant f(x_0)$$

2. On dit que $f(x_0)$ est un **maximum global** de f si :

$$\forall x \in P,\ f(x) \leqslant f(x_0)$$

3. On dit que $f(x_0)$ est un **extremum global** de f si c'est un minimum global ou un maximum global.

Définition 23.33

Soit f une fonction définie sur une partie P de $\mathbb{R}^n$ et x_0 un élément de P.

1. On dit que $f(x_0)$ est un **minimum local** de f s'il existe un réel α strictement positif tel que :

$$\forall x \in P \ /\ ||x - x_0|| < \alpha,\ f(x) \geqslant f(x_0)$$

2. On dit que $f(x_0)$ est un **maximum local** de f s'il existe un réel α strictement positif tel que :

$$\forall x \in P \ /\ ||x - x_0|| < \alpha,\ f(x) \leqslant f(x_0)$$

3. On dit que $f(x_0)$ est un **extremum local** de f si c'est un minimum local ou un maximum local.

D.2. Extremum sur un ensemble fermé borné

Théorème 23.34

Si f est une fonction continue sur une partie K de $\mathbb{R}^n$ et si K est une partie fermée et bornée, alors f admet un minimum global et un maximum global.

D.3. Condition d'ordre 1

Théorème 23.35

Si f est de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et si f admet un extremum local en x, alors : $\nabla f(x) = 0$.

Preuve Soit $x_0 \in \mathcal{O}$. On suppose que f admet un minimum local en x_0 (les autres cas se traitant de manière analogue).
On peut déjà remarquer que, comme $\mathcal{O}$ est ouvert, il existe un réel ε strictement positif tel que :

$$\forall x \in \mathbb{R}^n,\ ||x - x_0|| < \varepsilon \Longrightarrow x \in \mathcal{O}.$$

De plus, comme $f(x_0)$ est un minimum local de f sur $\mathcal{O}$, il existe un réel α strictement positif tel que :

$$\forall x \in \mathcal{O}\ /\ ||x - x_0|| < \alpha,\ f(x) \geqslant f(x_0).$$

En notant $\eta = \min(\varepsilon, \alpha)$ et en fixant $h \in \mathbb{R}^n$, on a alors, en particulier :

$$\forall t \in \mathbb{R}\ /\ ||(x_0 + th) - x_0|| < \eta,\ f(x_0 + th) \geqslant f(x_0)$$

soit encore, en notant $g : t \mapsto f(x_0 + th)$:

$$\forall t \in \left]-\eta, \eta\right[,\ g(t) \geqslant g(0).$$

g admet donc un minimum en x_0 sur $]-\eta, \eta[$. Comme l'intervalle $]-\eta, \eta[$ est ouvert et comme g est de classe $\mathcal{C}^1$ sur $]-\eta, \eta[$ d'après 23.17, on a donc :

$$\langle \nabla f(x_0), h\rangle = g'(0) = 0$$

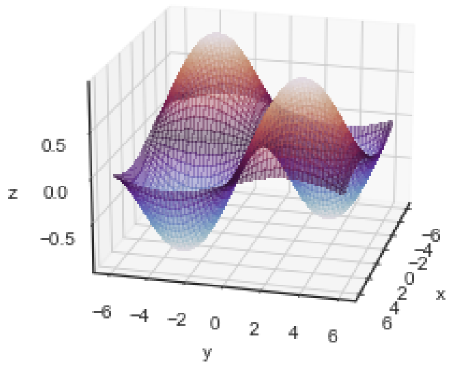
et donc, ce résultat étant valable pour tout $h \in \mathbb{R}^n$

$$\nabla f(x_0) = 0$$

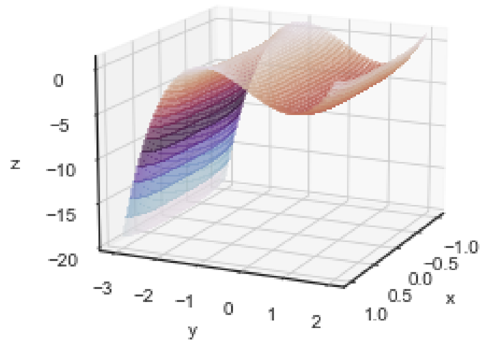
□

Définition 23.36

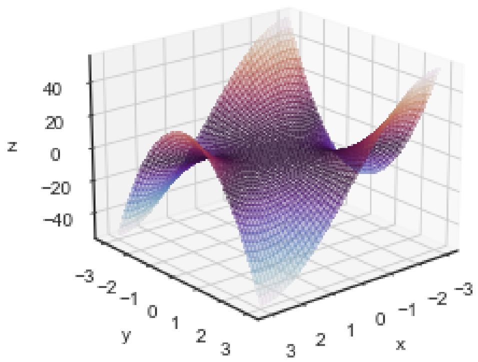
Si f est de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, on appelle **point critique** de f tout élément x de $\mathcal{O}$ tel que $\nabla f(x) = 0$.



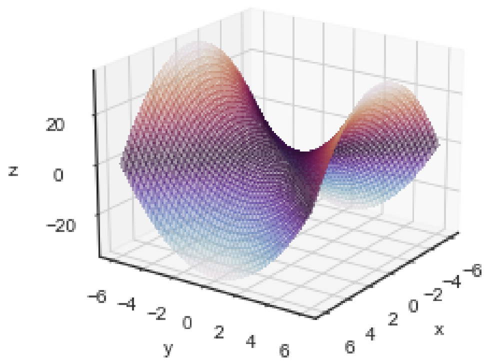
$a=(0,0)$ et $b=(\pi,\pi)$ sont des points critiques (entre autres) de $f:(x,y)\mapsto \sin\left(\frac{x}{2}\right)\sin\left(\frac{y}{2}\right)$; f a un maximum global en b, mais pas d'extremum en a.



$g:(x,y)\mapsto x^4+y^3-3y-2$ admet un minimum local en $(0,1)$ mais celui-ci n'est pas global.



$(0,0)$ est un point critique de $(x,y)\mapsto x^3-3xy^2$ mais cette fonction n'a pas d'extremum en ce point : c'est un point col (ou selle).



$(0,0)$ est un point critique de $h:(x,y)\mapsto y^2-x^2$ mais i n'a pas d'extremum en ce point : c'est un point selle (noter l'allure du graphe).

Remarques

a. Bien retenir qu'il ne s'agit que d'une condition nécessaire : le gradient d'une fonction peut s'annuler en un point a sans qu'elle admette un extremum en a (voir graphiques).

b. Attention aux hypothèses : il est important que f soit de classe $\mathcal{C}^1$ sur un ouvert pour que la recherche des points critiques ait un sens. Par exemple, la fonction $(x,y)\mapsto x+y$ admet un maximum global sur $[0,1]^2$, atteint en $(1,1)$ (et qui vaut 2), alors que son gradient ne s'annule pas.

c. On peut remarquer que, si x est un point critique de f, toutes les dérivées directionnelles en x sont nulles.

d. Retenir la méthode utilisée dans l'exemple suivant pour l'étude en $(0,0)$. Lorsque a est un point critique et qu'une étude générale du signe de $f(x)-f(a)$ n'est pas évidente, on peut commencer par étudier le signe lorsque x appartient à des droites particulières passant par a. En pratique, on commence par étudier le signe de $f(a+th)-f(a)$ lorsque t est proche de 0, pour des vecteurs h simples, du type $(1,0),(0,1),(1,1)$ ou $(1,-1)$. Pour démontrer qu'il n'y a pas d'extremum en a.Dans l'exemple, on avait $a=(0,0)$ et on a choisi les directions $(1,1)$ et $(1,0)$.

Exemple 23.3 Déterminons les extrema de la fonction $f:(x,y)\mapsto x^4+y^4-2(x-y)^2$ sur l'ouvert $\mathbb{R}^2$.

◇ f est une fonction polynôme donc elle est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et, comme $\mathbb{R}^2$ est un ouvert, elle ne peut ainsi admettre d'extremum qu'en un point critique (*i.e.* annulant son gradient). De plus, on a :

$$\forall (x,y)\in\mathbb{R}^2,\ \nabla f(x,y)=\big(4x^3-4(x-y),4y^3+4(x-y)\big)$$

et donc, pour tout $(x,y)\in\mathbb{R}^2$:

$$\begin{aligned}\nabla f(x,y)=0 &\iff \begin{cases}4x^3=4(x-y)\\4y^3=-4(x-y)\end{cases}\\&\iff\begin{cases}x^3=x-y\\y^3=-x^3\end{cases}\\&\iff\begin{cases}x(x^2-2)=0\\y=-x\end{cases}\\&\iff (x,y)\in\left\{(0,0),\left(\sqrt{2},-\sqrt{2}\right),\left(-\sqrt{2},\sqrt{2}\right)\right\}\end{aligned}$$

Ainsi, f ne peut admettre d'extremum qu'en l'un des points

$$x_1=(0,0),\quad x_2=\left(\sqrt{2},-\sqrt{2}\right)\quad\text{et}\quad x_3=\left(-\sqrt{2},\sqrt{2}\right)$$

- De plus on a :

$$\forall(x,y)\in\mathbb{R}^2,\ f(x,y)-f(0,0)=x^4+y^4-2(x-y)^2$$

On peut remarquer que :

$$\forall t\in\mathbb{R},\ \begin{cases}f(t,t)-f(0,0)=2t^4\\f(t,0)-f(0,0)=t^2(t^2-2)\end{cases}$$

et donc :

$$\forall t\in\left]0,\sqrt{2}\right[,\ \begin{cases}f(t,t)-f(0,0)>0\\f(t,0)-f(0,0)<0\end{cases}$$

ce qui suffit pour affirmer que $f(0,0)$ n'est pas un extremum local de f puisque, pour tout réel $\alpha>0$, on peut choisir deux points $a=(t,t)$ et $b=(t,0)$ vérifiant :

$$\begin{cases}||a-x_1||<\alpha\\||b-x_1||<\alpha\end{cases}\qquad\text{et}\qquad\begin{cases}f(a)>f(x_1)\\f(b)<f(x_1)\end{cases}$$

- Par ailleurs on a :

$$\begin{aligned}\forall(x,y)\in\mathbb{R}^2,\ f(x,y)-f(\left(\sqrt{2},-\sqrt{2}\right)&=x^4+y^4-2(x-y)^2+8\\&=x^4+y^4-2x^2+4xy-2y^2+8\\&=x^4-4x^2+y^4-4y^2+2(x^2+2xy+y^2)+8\\&=(x^2-2)^2+(y^2-2)^2+2(x+y)^2\\&\geqslant 0\end{aligned}$$

donc $f(x_2)=-8$ est le minimum global de f sur $\mathbb{R}^2$.

- De même, on prouve que $f(x_3)=-8$ est le minimum global de f sur $\mathbb{R}^2$.

Définition 23.37

Si f est de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$, on dit que x_0 est un **point selle** (ou que f admet un **col** en x_0 selon les cas) si x_0 est un point critique de f et si f n'admet pas d'extremum local en x_0.

D.4. Conditions suffisantes d'ordre 2

Théorème 23.38 ▶ Condition suffisante d'extremum local

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et x_0 un point critique de f.

i. Si les valeurs propres de $\nabla^2 f(x_0)$ sont toutes strictement positives, alors f admet un minimum local en x_0.

ii. Si les valeurs propres de $\nabla^2 f(x_0)$ sont toutes strictement négatives, alors f admet un maximum local en x_0.

iii. Si $\nabla^2 f(x_0)$ admet au moins une valeur propre strictement positive et une valeur propre strictement négative, alors f n'admet pas d'extremum en x_0.

Remarque Une preuve de ce théorème est proposée dans la section « Pour aller plus loin ».

Exercice 23.3 Déterminer les éventuels extremums locaux de la fonction $f : (x, y) \mapsto x^2 - y^2$ sur $\mathbb{R}^2$.

Remarque Les conditions 23.38 ne sont que des conditions suffisantes, et il arrivera qu'elles ne permettent pas de conclure, dans le cas où toutes les valeurs propres de la matrice hessienne en un point critique x_0 sont de même signe mais où 0 est valeur propre. Dans ce cas, f peut admettre, ou non, un extremum en x_0 et rien ne permet de conclure. En revanche, il y a un cas où l'on pourra conclure, qui est précisé dans le théorème suivant.

Théorème 23.39 ▶ Condition suffisante d'extremum global

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et x_0 un point critique de f.
Si, pour tout $x \in \mathbb{R}^n$, les valeurs propres de $\nabla^2 f(x)$ sont positives ou nulles (respectivement toutes négatives ou nulles), alors f admet en x_0 un minimum global (respectivement un maximum global).

Remarque Une preuve de ce théorème est proposée dans la section « Pour aller plus loin ».

Exercice 23.4 Montrer que la fonction $f : (x, y, z) \mapsto x^2 + 2y^2 + z^2$ admet un extremum global et préciser sa nature et sa valeur.

D.5. Recherche d'extrema sous contraintes d'égalités linéaires

Dans tout ce paragraphe, $\mathcal{O}$ désigne un ouvert de $\mathbb{R}^n$ et f est une fonction définie sur $\mathcal{O}$. Pour tout $x \in \mathbb{R}^n$, on convient de noter X la colonne des coordonnées de x dans la base canonique de $\mathbb{R}^n$.
De plus, p est un entier naturel non nul, $g_1, \dots, g_p$ sont des applications linéaires (*i.e.* des fonctions affines s'annulant en 0) définies sur $\mathbb{R}^n$, $b_1, \dots, b_p$ sont des réels et $\mathcal{C}$ désigne l'ensemble des solutions du système linéaire

$$\begin{cases} g_1(x) = b_1 \\ \quad\vdots \\ g_p(x) = b_p \end{cases}$$

Ce système linéaire peut être écrit sous la forme $AX = B$ où A est une matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ et B est une matrice de $\mathcal{M}_{p,1}(\mathbb{R})$.
On note $\mathcal{H}$ l'ensemble des solutions du système homogène associé, c'est-à-dire l'ensemble des solutions du système

$$\begin{cases} g_1(x) = 0 \\ \quad\vdots \\ g_p(x) = 0 \end{cases}$$

Définition 23.40

On dit que f admet un extremum (local ou global) sous la contrainte $AX = B$ si f admet un extremum (local ou global) sur $\mathcal{C}$.

Théorème 23.41

Si f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$ et si f admet un extremum en x_0 sur $\mathcal{C}$ (*i.e.* sous la contrainte $AX = B$), alors :

$$\nabla f(x_0) \in \mathcal{H}^\perp$$

Proposition 23.42

Les dérivées partielles de $g_1, \ldots, g_p$ sont constantes, donc les applications $x \mapsto \nabla g_i(x)$ $(1 \leqslant i \leqslant n)$ aussi et :

$$\forall x \in \mathbb{R}^n,\ \mathcal{H}^\perp = \mathrm{Vect}\,(\nabla g_1(x), \ldots, \nabla g_p(x))$$

Preuve $g_1, \ldots, g_p$ étant des applications linéaires sur $\mathbb{R}^n$, on note :

$$\forall i \in [\![1,p]\!]\,,\ \forall x = (x_i)_{1\leqslant i\leqslant n},\ g_i(x) = \sum_{j=1}^n a_{i,j}x_j$$

On a alors :

$$\forall i \in [\![1,p]\!]\,,\ \forall x \in \mathbb{R}^n,\ \nabla g_i(x)(a_{i,j})_{1\leqslant j\leqslant n}$$

Par ailleurs, pour tout $x \in \mathbb{R}^n$, on a, en notant $x = (x_i)_{1\leqslant i\leqslant n}$:

$$\begin{aligned} x \in \mathcal{H} &\iff \forall i \in [\![1,p]\!]\,,\ \sum_{j=1}^n a_{i,j}x_j = 0 \\ &\iff \forall i \in [\![1,p]\!]\,,\ \langle x, \nabla g_i(x)\rangle = 0 \\ &\iff x \in [\mathrm{Vect}\,(\nabla g_1(x), \ldots, \nabla g_p(x))]^\perp \end{aligned}$$

et on a donc :

$$\mathcal{H}^\perp = \mathrm{Vect}\,(\nabla g_1(x), \ldots, \nabla g_p(x))$$

□

Définition 23.43

Si f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$, on appelle **point critique** pour l'optimisation sous contrainte $AX = B$ tout élément x_0 de $\mathcal{C}$ tel que : $\nabla f(x_0) \in \mathcal{H}^\perp$.

Remarque Il est essentiel de bien comprendre que, selon que l'on recherche des extrema sur un ouvert ou sous contrainte, les points critiques (seuls points où la fonction pourra atteindre un extremum) ne sont pas définis de la même façon.

Exemple 23.4 On se propose de déterminer les extremums de $f : (x,y,z) \mapsto x^2 + y^2 + z^2$ sur $\mathbb{R}^3$ sous la contrainte $x + y + z = 3$.
f est de classe $\mathcal{C}^1$ sur l'ouvert $\mathbb{R}^3$ donc, si elle admet un extremum en (x,y,z) sous la contrainte $x + y + z = 3$, alors :

$$\nabla f(x,y,z) \in \mathcal{H}^\perp \quad \text{où} \quad \mathcal{H} = \{(x,y,z) \in \mathbb{R}^3 \ /\ x + y + z = 0\}$$

Or on a :

$$\mathcal{H} = \{(x,y,z) \in \mathbb{R}^3 \ /\ \langle (x,y,z),(1,1,1)\rangle = 0\} = [\mathrm{Vect}((1,1,1))]^\perp$$

et donc :

$$\mathcal{H}^\perp = \mathrm{Vect}((1,1,1))$$

Par ailleurs, on a :

$$\forall (x,y,z) \in \mathbb{R}^3,\ \nabla f(x,y,z) = (2x,2y,2z)$$

donc, pour tout $(x,y,z)\in\mathbb{R}^3$:

$$\nabla f(x,y,z)\in\mathcal{H}^\perp \iff x=y=z$$

donc, sous la contrainte $x+y+z=3$, $(1,1,1)$ est l'unique point critique de f. Enfin, on peut remarquer que, d'après l'inégalité de Cauchy-Shwarz, on a :

$$\forall (x,y,z)\in\mathbb{R}^3\ /\ x+y+z=3,\ \langle (1,1,1),(x,y,z)\rangle^2 \leqslant ||(1,1,1)||^2\,||(x,y,z)||^2$$

c'est-à-dire :

$$\forall (x,y,z)\in\mathbb{R}^3\ /\ x+y+z=3,\ (x+y+z)^3 \leqslant 9f(x,y,z)$$

et finalement :

$$\forall (x,y,z)\in\mathbb{R}^3\ /\ x+y+z=3,\ f(x,y,z)\geqslant 3$$

Comme $f(1,1,1)=3$, on en déduit finalement que f admet un minimum global sous la contrainte $x+y+z=3$, égal à 3 et atteint uniquement en $(1,1,1)$.

Remarque Le programme ne donne aucune condition suffisante pour savoir si une fonction f présente un extremum en un point critique sous contrainte d'égalités linéaires. Une fois les points critiques déterminés, il s'agira donc, pour tout point critique x_0, d'étudier le signe de $f(x)-f(x_0)$ pour tout point x vérifiant la contrainte. Pour cela, on pourra faire une étude directe ou, dans le cas où f est de classe $\mathcal{C}^2$ et où les valeurs propres de la matrice hessienne sont toutes de même signe en tout point, on pourra raisonner comme dans la preuve de la proposition 23.39.

E. Correction des exercices

Correction de l'exercice 23-1

◇ On a :

$$\begin{aligned}\forall (x,y)\in[0,1]^2,\ ||(x,y)|| &= \sqrt{x^2+y^2}\\ &\leqslant \sqrt{2}\end{aligned}$$

donc l'ensemble $[0,1]^2$ est borné.

◇ On a, pour tout $(x,y)\in\mathbb{R}^2$:

$$\begin{aligned}(x,y)\in E &\iff \begin{cases}x\geqslant 0\\ y\geqslant 0\\ x+y\leqslant 1\end{cases}\\ &\implies \begin{cases}0\leqslant x\leqslant 1-y\\ 0\leqslant y\leqslant 1-x\end{cases}\\ &\implies ||(x,y)||\leqslant\sqrt{2}\end{aligned}$$

donc E est borné.

◇ On a :

$$F=\left\{(x,y)\in\mathbb{R}^2\ /\ y\leqslant x\leqslant y+2\right\}$$

En particulier, on en déduit que, pour tout $M\in\mathbb{R}$, $a=(M,M)$ appartient à F et donc :

$$\forall M\in\mathbb{R}^+,\ \exists\, a\in F\ /\ ||a||>M$$

ce qui prouve que F n'est pas borné.

Correction de l'exercice 23-2

La fonction $(x,y) \mapsto 1+x^2+y^2$ est une fonction polynôme de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et prend ses valeurs dans $\mathbb{R}_+^*$. Comme la fonction ln est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$, on en déduit que la fonction $(x,y) \mapsto \ln(1+x^2+y^2)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. De plus la fonction $(x,y) \mapsto 2x^3-4x^2y$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ comme fonction polynôme définie sur $\mathbb{R}^2$, donc f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. On peut alors écrire :

$$\forall (x,y) \in \mathbb{R}^2,\ \partial_1 f(x,y) = \frac{2x}{1+x^2+y^2} + 6x^2 - 8xy \quad \text{et} \quad \partial_2 f(x,y) = \frac{2y}{1+x^2+y^2} - 4x^2$$

Comme les fonctions $(x,y) \mapsto 1+x^2+y^2$, $(x,y) \mapsto 2x$, $(x,y) \mapsto 2y$, $(x,y) \mapsto 6x^2-8xy$ et $(x,y) \mapsto 4x^2$ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et comme la fonction $(x,y) \mapsto 1+x^2+y^2$ ne s'annule pas sur $\mathbb{R}^2$, on en déduit que $\partial_1 f$ et $\partial_2 f$ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$, donc f a des dérivées partielles d'ordre 2 en tout point et :

$$\forall (x,y) \in \mathbb{R}^2,\ \begin{cases} \partial^2_{1,1} f(x,y) = \dfrac{2-2x^2+2y^2}{(1+x^2+y^2)^2} + 12x - 8y \\ \partial^2_{1,2} f(x,y) = \partial^2_{2,1} f(x,y) = -\dfrac{4xy}{(1+x^2+y^2)^2} - 8x \\ \partial^2_{2,2} f(x,y) = \dfrac{2+2x^2-2y^2}{(1+x^2+y^2)^2} \end{cases}$$

Correction de l'exercice 23-3

f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ et $\mathbb{R}^2$ est ouvert, donc si f admet un extremum en $a \in \mathbb{R}^2$, alors : $\nabla f(a) = 0$. Or on a, pour tout $(x,y) \in \mathbb{R}^2$:

$$\partial_1 f(x,y) = 2x \quad \text{et} \quad \partial_2 f(x,y) = -2y$$

donc :

$$\nabla f(x,y) = 0 \Longleftrightarrow x = y = 0$$

donc $(0,0)$ est l'unique point critique de f. De plus, on a, pour tout $(x,y) \in \mathbb{R}^2$:

$$\partial^2_{1,1} f(x,y) = 2, \quad \partial^2_{1,2} f(x,y) = \partial^2_{2,1} f(x,y) = 0 \quad \text{et} \quad \partial^2_{2,2} f(x,y) = -2,$$

donc :

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

On constate que $\nabla^2 f(0,0)$ admet une valeur propre strictement positive et une valeur propre strictement négative, donc f n'a pas d'extremum local.

Correction de l'exercice 23-4

f est une fonction polynôme donc elle est de classe $\mathcal{C}^2$ sur l'ouvert $\mathbb{R}^2$. Elle ne peut donc admettre d'extremum qu'en un point annulant sont gradient.
Or on a :

$$\forall (x,y,z) \in \mathbb{R}^3,\ \nabla f(x,y,z) = (2x, 4y, 2z)$$

donc $(0,0,0)$ est l'unique point critique de f. Ainsi $f(0,0,0) = 0$ est l'unique extremum possible de f.
Enfin on a :

$$\forall (x,y,z) \in \mathbb{R}^3,\ \nabla^2 f(x,y,z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

d'où :

$$\forall (x,y,z) \in \mathbb{R}^3,\ \mathrm{Sp}(\nabla^2 f(x,y,z)) = \{2,4\} \subset \mathbb{R}^+$$

ce qui nous permet de conclure, d'après 23.39, que f admet un minimum global en $(0,0,0)$, égal à 0 (ce qui n'est pas une grosse surprise !).
Montrer que la fonction $f : (x,y,z) \mapsto x^2 + 2y^2 + z^2$ admet un extremum global et préciser sa nature et sa valeur.

Sommaire