Chapitre 6

Extrema

Trouver les *extrema* d'une fonction de la variable réelle et à valeurs réelles se fait en général en utilisant son tableau de variations. Pour cela, on dérive la fonction (après s'être assuré que celle-ci est dérivable), on détermine ses *points critiques* (c'est-à-dire les valeurs pour lesquelles la dérivée s'annule) et on décide du sens de variation de la fonction sur chaque intervalle borné par deux points critiques. Cette méthode permet d'obtenir les *extrema locaux* et les *extrema globaux* de la fonction.

Dans ce chapitre, on veut obtenir les extrema de fonctions d'une variable vectorielle. Pour que cela ait un sens, elle sera à valeurs réelles. On va dans un premier temps suivre la méthode de la dimension 1 en déterminant les points critiques de la fonction, mais la suite n'est en revanche pas adaptable. On distinguera deux cas, suivant si le domaine de définition de la fonction est ouvert ou compact.

Dans toute la suite, E désigne un $\mathbb{R}$-espace vectoriel de dimension finie et U un ouvert de E.

6.1 Dérivées partielles d'ordre supérieur à 1

6.1.1 Dérivées partielles successives

Définition 1. *Soient $f : U \subset E \longrightarrow F$ et $\mathcal{B}$ une base de E.*
La fonction f est appelée dérivée partielle d'ordre 0 de f.
Pour $k \in \mathbb{N}$, et sous réserve d'existence, on appelle dérivées partielles d'ordre $k+1$ de f les dérivées partielles des dérivées partielles d'ordre k de f (dans la base $\mathcal{B}$).

Remarque. Si l'on note $x_1, \ldots, x_n$ les coordonnées dans la base $\mathcal{B}$ de la variable $x \in E$, on note

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} = \partial_{i_1}(\cdots(\partial_{i_k} f)\cdots).$$

Exemple : Considérons la fonction $f : (x, y) \in \mathbb{R}^2 \longmapsto xe^{xy}$.

Exemple : Reprenons l'application $f : z \in \mathbb{C}^* \longmapsto \dfrac{1}{z}$. Considérons la base $\mathcal{B} = (1, i)$ de $\mathbb{C}$ (en tant que $\mathbb{R}$-espace vectoriel).

On a déjà vu que f admet des dérivées partielles d'ordre 1

$\partial_1 f : z \in \mathbb{C}^* \mapsto \dfrac{1}{-z^2}$

$\partial_2 f : z \in \mathbb{C}^* \mapsto \dfrac{-i}{z^2}$

On pose $\widetilde{\partial_1 f} : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{C}$

$(x,y) \mapsto -\dfrac{1}{(x+iy)^2}$

par opérations, $\widetilde{\partial_1 f}$ admet des dérivées partielles sur $\mathbb{R}^2 \setminus \{(0,0)\}$

$\forall z = x+iy$

$\partial_1(\partial_1 f)(z) = \dfrac{2}{z^3}$

$\partial_2(\partial_1 f)(z) = \dfrac{2i}{z^3}$

6.1.2 Classe d'une fonction

Définition 2. *Soit $k \in \mathbb{N}$. On dit qu'une fonction $f : U \subset E \longrightarrow F$ est de classe $\mathcal{C}^k$ (sur U) si ses dérivées partielles d'ordre k dans une base de E existent et sont continues (sur U).*
On dit que f est de classe $\mathcal{C}^\infty$ si elle de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.

Remarque. On peut montrer que cette définition ne dépend pas du choix de la base de E utilisée pour définir les dérivées partielles de f.

Exemples :

① Les fonctions constantes sont C^∞ sur leur domaine de définition

② Soit $u : E \to F$ linéaire, on a vu que u est de classe C^1 sur E et $du(a) = u \; \forall a \in E$
Donc si $\dim E = n$ et $B = (e_1, \dots, e_n)$ une base de E, on a $\forall i \in [\![1;n]\!]$, la i-ème dérivée partielle d'ordre 1 en $a \in E$ est

$\partial_i u(a) = du(a)(e_i) = u(e_i)$ $\Rightarrow$ classe C^∞

Les fonctions projecteurs sont C^∞ sur $\mathbb{R}^n$

Proposition 1. *Soit $k \in \mathbb{N} \cup \{\infty\}$. Si $f : U \subset E \longrightarrow F$ est de classe $\mathcal{C}^{k+1}$ sur U, alors f est de classe $\mathcal{C}^k$ sur U.*

Démonstration. Si f est de classe $\mathcal{C}^{k+1}$ sur U, alors elle admet des dérivées partielles (dans une base fixée de E) d'ordre k (puisque les dérivées partielles d'ordre $k+1$ existent). De plus ses dérivées partielles d'ordre k sont de classe $\mathcal{C}^1$ puisque leurs dérivées partielles premières existent et sont continues, donc elles sont continues. Ceci démontre que f est de classe $\mathcal{C}^k$ sur U. $\square$

6.1.3 Opérations sur les fonctions de classe $\mathcal{C}^k$

Pour ne pas alourdir les notations dans les résultats ci-dessous, on se donne une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E et on considérera dans ce paragraphe que les dérivées partielles seront calculées relativement à cette base $\mathcal{B}$.

Lemme 1. *Soient $k \in \mathbb{N} \cup \{\infty\}$ et $f : U \subset E \longrightarrow F$. On a équivalence entre :*

(i). *f est de classe $\mathcal{C}^{k+1}$ sur U,*

(ii). *les dérivées partielles premières de f existent et sont de classe $\mathcal{C}^k$ sur U.*

Démonstration. Les dérivées partielles d'ordre $k+1$ de f sont les dérivées partielles des dérivées partielles d'ordre k de f. □

Proposition 2. *Soient $f : U \subset E \longrightarrow F$, $k \in \mathbb{N} \cup \{\infty\}$ et $\mathcal{B}' = (e'_1, \ldots, e'_p)$ une base de F. On a équivalence entre :*

(i). *f est de classe $\mathcal{C}^k$ sur U,*

(ii). *les fonctions coordonnées $f_1, \ldots, f_p$ de f dans $\mathcal{B}'$ sont de classe $\mathcal{C}^k$ sur U.*

Démonstration. On démontre le résultat par récurrence sur k, sachant que l'on a déjà démontré que f admet des dérivées partielles si et seulement si ses fonctions coordonnées en admettent, et pour tous i, j,

$$\partial_i(f_j) = (\partial_i f)_j.$$

□

Exemple : La fonction $f : \begin{array}{l} \mathbb{R}^2 \to \mathbb{R} \\ (x,y) \mapsto x + ie^{yx} \end{array}$ est C^2 sur $\mathbb{R}^2$

En effet, dans la base $\mathcal{B} = (1, i)$ de $\mathbb{C}$, les fonctions coordonnées de f sont : $f_1 : \begin{array}{l} \mathbb{R}^2 \to \mathbb{R} \\ (x,y) \mapsto x \end{array}$ et $f_2 : \begin{array}{l} \mathbb{R}^2 \to \mathbb{R} \\ (x,y) \mapsto xe^{xy} \end{array}$

↳ linéaire donc C^∞ ↳ C^2 par l'exemple précédent

Proposition 3. *Soient $k \in \mathbb{N} \cup \{\infty\}$ et $f, g : U \subset E \longrightarrow F$ de classe $\mathcal{C}^k$ sur U. Pour tous $\lambda, \mu \in \mathbb{R}$, la fonction $\lambda f + \mu g$ est aussi de classe $\mathcal{C}^k$ sur U.*

Démonstration. On raisonne par récurrence sur $k \in \mathbb{N}$.

- Le cas $k = 0$ a déjà été traité lorsque l'on a étudié les propriétés des fonctions continues.
- Supposons le résultat vrai au rang k pour un certain $k \in \mathbb{N}$. Soit f, g de classe $\mathcal{C}^{k+1}$ sur U.Alors f et g sont en particulier de classe $\mathcal{C}^1$ sur U et pour tout $i \in [\![1; n]\!]$,

$$\partial_i(\lambda f + \mu g) = \lambda \partial_i f + \mu \partial_i g.$$

Puisque $\partial_i f$ et $\partial_i g$ sont de classe $\mathcal{C}^k$ sur U, l'hypothèse de récurrence entraîne que $\partial_i(\lambda f + \mu g)$ est de classe $\mathcal{C}^k$. Ainsi les dérivées partielles premières de $\lambda f + \mu g$ existent et sont de classe $\mathcal{C}^k$ ce qui prouve que $\lambda f + \mu g$ est de classe $\mathcal{C}^{k+1}$.

Si f et g sont de classe $\mathcal{C}^\infty$ sur U, alors elles sont de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$. On peut alors appliquer le résultat que l'on vient de démontrer : $\lambda f + \mu g$ est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$. □

Proposition 4. *Soient $k \in \mathbb{N} \cup \{\infty\}$, $f : U \subset E \longrightarrow F$ et $\alpha : U \subset E \longrightarrow \mathbb{R}$ deux fonctions de classe $\mathcal{C}^k$ sur U. Alors le produit αf est aussi de classe $\mathcal{C}^k$ sur U.*

Démonstration. On raisonne encore une fois par récurrence sur $k \in \mathbb{N}$. Le cas $k = 0$ a été traité. Supposons l'hypothèse vraie au rang $k \in \mathbb{N}$. Soient $f : U \longrightarrow F$ et $\alpha : U \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^{k+1}$. Alors f et α sont en particulier de classe $\mathcal{C}^1$ sur U et pour tout $i \in [\![1; n]\!]$:

$$\partial_i(\alpha f) = \partial_i \alpha \times f + \alpha \times \partial_i f$$

D'après l'hypothèse de récurrence et par opérations, $\partial_i(\alpha f)$ est de classe $\mathcal{C}^k$ ce qui entraîne αf de classe $\mathcal{C}^{k+1}$ ce qui achève la récurrence. □

Exemple :

Théorème 1 (Composition).
Soient $k \in \mathbb{N} \cup \{\infty\}$, $f : U \subset E \longrightarrow F$ et $g : V \subset F \longrightarrow G$ avec U, V deux ouverts tels que $f(U) \subset V$. Si f et g sont de classe $\mathcal{C}^k$ (resp. sur U et V), alors la composée $g \circ f$ est elle-aussi de classe $\mathcal{C}^k$ (sur U).

Démonstration. La démonstration se fait par récurrence sur $k \in \mathbb{N}$.

- pour $k = 0$, c'est le théorème de composition des fonctions continues.
- Supposons le résultat démontré au rang k ($k \in \mathbb{N}$ fixé). Soient f, g de classe $\mathcal{C}^{k+1}$ définies comme dans l'énoncé. Alors on sait en particulier que f et g sont de classe $\mathcal{C}^1$ et donc $g \circ f$ aussi par le théorème de dérivation d'une composée. De plus, pour tout $a \in U$,

$$\forall i \in [\![1; n]\!], \quad \partial_i(g \circ f)(a) = \sum_{l=1}^{p} \partial_i f_l(a) \partial_l g(f(a))$$

  où l'on a noté $f_1, \ldots, f_p$ les fonctions coordonnées de f dans une base $\mathcal{B}'$ de F. Puisque f et g sont de classe $\mathcal{C}^{k+1}$, les dérivées partielles $\partial_i f_l$ et $\partial_l g$ sont de classe $\mathcal{C}^k$, et comme f est aussi de classe $\mathcal{C}^k$, on peut affirmer d'après l'hypothèse de récurrence et par les théorèmes d'opérations que pour tout $i \in [\![1; n]\!]$, $\partial_i(g \circ f)$ est de classe $\mathcal{C}^k$. Ainsi, les dérivées partielles de $g \circ f$ existent et sont de classe $\mathcal{C}^k$, donc les dérivées partielles d'ordre $k + 1$ de $g \circ f$ existent et sont continues, ce qui démontre que $g \circ f$ est de classe $\mathcal{C}^{k+1}$ et termine la récurrence.

□

6.1.4 Théorème de Schwarz

Théorème 2 (Théorème de Schwarz(admis)).
Soit $f : U \subset E \longrightarrow F$ une fonction de classe $\mathcal{C}^2$. Alors pour tous $i, j \in [\![1; n]\!]$,

$$\partial_i(\partial_j f) = \partial_j(\partial_i f)$$

où les dérivées partielles de f sont calculées relativement à une base $\mathcal{B}$ de E.

Corollaire 1. *Soit $f : U \subset E \longrightarrow F$ de classe $\mathcal{C}^k$ sur U. Pour tous $i_1, \ldots, i_k \in [\![1; n]\!]$, pour toute permutation $\sigma \in S_k$ (groupe des permutations de $[\![1; k]\!]$), on a*

$$\partial_{i_{\sigma(1)}}(\partial_{i_{\sigma(2)}}(\cdots(\partial_{i_{\sigma(k)}} f)\cdots)) = \partial_{i_1}(\cdots(\partial_{i_k} f)\cdots)$$

c'est-à-dire

$$\frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \cdots \partial x_{i_{\sigma(k)}}} = \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}.$$

où les dérivées partielles de f sont calculées relativement à une base $\mathcal{B}$ de E.

Exemple : Déterminer la classe de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par

$$f(x, y) = \frac{xy^3}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \quad \text{et} \quad f(0, 0) = 0.$$

- Sur l'ouvert $\mathbb{R}^2 \setminus \{(0,0)\}$, f est quotient de 2 fonctions polynomiales dont le dénominateur ne s'annule pas, donc f est C^∞ sur $\mathbb{R}^2 \setminus \{(0,0)\}$
- Etudions la classe de f sur un voisinage de $(0,0)$

$\to C^0$?

Soit $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, $\exists r > 0$, $\theta \in \mathbb{R}$ tq $(x,y) = (r\cos\theta, r\sin\theta)$

alors $0 \leqslant |f(x,y)| = \left| \frac{r\cos\theta\,(r\sin\theta)^3}{r^2} \right|$

$= r^2 |\cos\theta| |\sin\theta|^3$

$\leqslant r^2 \xrightarrow[(x,y)\to(0,0)]{} 0$

Par encadrement, $f(x,y) \xrightarrow[(x,y)\to(0,0)]{} 0 = f(0,0)$

donc f est C^0 sur $\mathbb{R}^2$

$\to$ Existence de $\partial_1 f(0,0)$ et $\partial_2 f(0,0)$

On sait que

$\partial_1 f(0,0)$ existe

$\Leftrightarrow D_{(1,0)} f(0,0)$ existe

$\Leftrightarrow \frac{1}{t}\left(f((0,0)+t(1,0)) - f(0,0)\right)$ admet une limite finie quand $t \to 0$

Soit $t \in \mathbb{R}^*$

$\frac{1}{t}(f(t,0)) = \frac{1}{t} \times 0 = 0 \xrightarrow[t\to 0]{} 0$

donc $\partial_1 f(0,0) = \frac{\partial f}{\partial x}(0,0)$ existe et vaut 0

De même, $\partial_2 f(0,0)$ existe $\Leftrightarrow D_{(0,1)} f(0,0)$ existe

or pour $t \in \mathbb{R}^*$

$\frac{1}{t} f(0,t) = \frac{1}{t} \times 0 = 0 \xrightarrow[t\to 0]{} 0$

donc $\partial_2 f(0,0) = \frac{\partial f}{\partial y}(0,0)$ existe et vaut 0

Or on a vu que f est de classe C^∞ sur l'ouvert $\mathbb{R}^2 \setminus \{(0,0)\}$

et $\forall (x,y) \in \mathbb{R}^2$,

$$\partial_1 f(x,y) = \frac{y^3(x^2+y^2) - xy^3 \times 2x}{(x^2+y^2)^2} = \frac{y^5 - x^2y^3}{(x^2+y^2)^2}$$

$$\partial_2 f(x,y) = \frac{3xy^2(x^2+y^2) - xy^3 \times 2y}{(x^2+y^2)} = \frac{3x^3y^2 + xy^4}{(x^2+y^2)^2}$$

$\to$ Classe C^2

Existence des dérivées partielles croisées en $(0,0)$?

On a $\partial_1(\partial_2 f(0,0)) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(0,0)$

Soit $t \in \mathbb{R}^*$,

$\frac{1}{t}\left(\frac{\partial f}{\partial y}((0,0) + t(1,0)) - \frac{\partial f}{\partial y}(0,0)\right)$

$= \frac{1}{t}(0 - 0) = 0 \xrightarrow[t\to 0]{} 0$

donc $\partial_1(\partial_2 f)(0,0) = 0$

Existence de $\partial_2(\partial_1 f)(0,0)$?

Soit $t \in \mathbb{R}^*$

$\frac{1}{t}\left(\frac{\partial f}{\partial x}((0,0)+t(1,0)) - \frac{\partial f}{\partial x}(0,0)\right)$

$= \frac{1}{t}\left(\frac{\partial f}{\partial x}(0,t) - 0\right)$

$= \frac{1}{t} \times \frac{t^5}{(t^2)^2} = 1 \to 1$

$\to$ contraposée du thm de Schwarz

Etudions la continuité de f en $(0,0)$.

Soit $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, $\exists r > 0$, $\theta \in \mathbb{R}$ tq $(x,y) = (r\cos\theta, r\sin\theta)$

$$0 \leq |\partial_1 f(x,y)| = r\,|\sin^5(\theta) - \cos^3\theta \sin^2\theta| \leq r(|\sin^5\theta| + |\cos^3\theta||\sin^2\theta|) \leq 2r \xrightarrow[(x,y)\to(0,0)]{} 0$$

donc $\partial_1 f$ est continue en $(0,0)$ et on fait de même pour $\partial_2 f(x,y)$

6.2 Définitions

On commence par donner les définitions des extrema. On distingue les extrema globaux (ceux dont la valeur majore ou minore la fonction sur tout le domaine de définition) des extrema locaux (ceux dont la valeur majore ou minore la fonction au voisinage de l'extremum).

Définition 3. *Soit D une partie quelconque de E et $f : D \longrightarrow \mathbb{R}$ une fonction à valeurs réelles.*

(i) La fonction f admet un ***minimum global*** *en $a \in D$ si $f(x) \geq f(a)$ pour tout $x \in D$.*

(ii) La fonction f admet un ***maximum global*** *en $a \in D$ si $f(x) \leq f(a)$ pour tout $x \in D$.*

(iii) La fonction f admet un ***minimum local*** *en $a \in D$ si $f(x) \geq f(a)$ dans un voisinage de a (c'est-à-dire s'il existe $r > 0$ tel que $f(x) \geq f(a)$ pour tout $x \in B(a, r) \cap D$).*

(iv) La fonction f admet un ***maximum local*** *en $a \in D$ si $f(x) \leq f(a)$ dans un voisinage de a (c'est-à-dire s'il existe $r > 0$ tel que $f(x) \leq f(a)$ pour tout $x \in B(a, r) \cap D$).*

Un ***extremum*** *(local ou global) de f est un minimum ou un maximum (local ou global). Il est dit strict si l'inégalité est stricte pour $x \neq a$ (c'est-à-dire $f(x) > f(a)$ pour un minimum strict et $f(x) < f(a)$ pour un maximum strict).*

Remarque :

Les extrema globaux (qu'on appelle parfois extrema absolus) sont en particuliers des extrema locaux (qu'on appelle parfois extrema relatifs) mais la réciproque est fausse. Par exemple, le point d'abscisse 1 est un minimum local non global ; le point d'abscisse 3 est un maximum local non global et le point d'abscisse -2 est un maximum local et global.

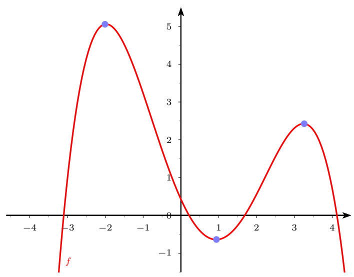

FIGURE 6.1 – Des extrema locaux non globaux.

6.3 Extrema sur un ouvert

6.3.1 Points critiques

Le but de cette section est de déterminer des conditions pour que f admette un extremum local en un point d'un ouvert de E. Commençons par rappeler ce qu'il se passe pour les fonctions d'une seule variable réelle.

Proposition 5 (Extremum d'une fonction dérivable)**.** *Soit I un intervalle* ***ouvert*** *et $f : I \longrightarrow \mathbb{R}$ une fonction dérivable. Si la fonction f admet un extremum local en $c \in I$, alors $f'(c) = 0$.*

Démonstration. L'idée est de regarder le signe du taux d'accroissement de f en l'extremum, de calculer la limite par la gauche et par la droite pour montrer que celle-ci est nécessairement nulle. $\square$

Définition 4. *On dit qu'une application $f : U \subset E \longrightarrow \mathbb{R}$ différentiable admet un* ***point critique*** *en $a \in U$ si $\mathrm{d}f(a)$ est l'application nulle de E dans $\mathbb{R}$, notée $0_{\mathcal{L}(E,\mathbb{R})}$.*

Proposition 6. *Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E, $f : U \subset E \longrightarrow \mathbb{R}$ différentiable et $a \in U$. On a équivalence entre :*

(i). a est un point critique de f,

(ii). $\forall i \in [\![1; n]\!]$, $\partial_i f(a) = 0$,

(iii). la Jacobienne de f en a est la matrice nulle

en calculant les dérivées partielles et la Jacobienne de f relativement à la base $\mathcal{B}$.

Démonstration.

On a a est un point critique de f

$\iff df(a) = 0_{\mathcal{L}(E,\mathbb{R})}$

$\iff \mathrm{Mat}_{\mathcal{B},(1)}(df(a)) = 0_{M_{1,n}(\mathbb{R})}$

$\iff \mathrm{Jac}_f(a) = 0_{1,n}(\mathbb{R})$

$\iff (\partial_1 f(a), \ldots, \partial_n f(a)) = (0, \ldots, 0)$

$\iff \forall i \in [\![1;n]\!],\ \partial_i f(a) = 0$

□

La proposition suivante est une généralisation de la proposition 5 dans le cas des fonctions d'une variable vectorielle.

Proposition 7 (Condition nécessaire d'extremum sur un ouvert)**.** *Soit U un* ***ouvert*** *de E et $f : U \longrightarrow \mathbb{R}$ une fonction différentiable. Si f admet un extremum local en $a \in U$, alors a est un point critique de f.*

Démonstration. Sq f admet un extremum local en a

Soit $h \in E$, considérons $\varphi : t \in \mathbb{R} \mapsto f(a+th)$

Alors φ est définie au moins sur un voisinage $]-\alpha, \alpha[$ de 0, avec $\alpha > 0$

Comme f admet un extremum local en a, φ admet un extremum local en 0

et donc $\varphi'(0) = 0$ (car φ est dérivable sur $]-\alpha;\alpha[$

Or par définition d'une dérivée directionnelle,

$0 = \varphi'(0) = D_h(f(a)) = df(a)(h)$

f est différentiable en a

□

Remarque.

① Le résultat devient faux si U n'est pas ouvert.

par ex, $f : [0;1] \to \mathbb{R}$, $x \mapsto x$ admet un max global en 1 mais $f'(1) \neq 0$

Dans $\mathbb{R}^2$, soit $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2+y^2$

Posons $C = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\}$

Considérons $f|_C$

$\forall (x,y) \in C$,

$\frac{\partial f}{\partial x} = 2x \qquad \frac{\partial f}{\partial y} = 2y$

donc $\begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases} \Longleftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases} \Longleftrightarrow (x,y) = (0,0)$

or $(0,0) \notin C$

Exemple : Déterminer les extrema (locaux et globaux) de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x,y) = x^2 + y^2 + xy - 1$.

Exemple : Déterminer les extrema (locaux et globaux) de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x,y) = x^2 + y^2 + 4xy - 1$.

6.3.2 Matrice hessienne

Dans toute la suite, on supposera que $E = \mathbb{R}^n$ pour simplifier. On pourrait énoncer des résultats similaires dans le cas d'un espace vectoriel normé de dimension finie en identifiant la fonction avec la fonction de $\mathbb{R}^n$ donnée par

$$\widetilde{f} : (x_1, \ldots, x_n) \longmapsto f\left(\sum_{i=1}^{n} x_i e_i\right).$$

Dans les exemples précédents, nous avons cherché à la main les extrema d'une fonction relativement simple. La méthode utilisée sera parfois impraticable sur des exemples plus complexes. Nous allons donc voir un critère donnant des conditions suffisantes pour qu'un point critique soit un extremum local.

Définition 5. *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \longrightarrow \mathbb{R}$ une application. La **matrice hessienne** de f au point $a \in U$ est, si elle est définie, la matrice des dérivées partielles secondes de f en a :*

$$\begin{aligned} H_f(a) &= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{1 \le i,j \le n} \\ &= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \ldots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix} \end{aligned}$$

Remarque.

Exemple : Déterminer la matrice hessienne de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2+y^2+4xy-1$ en tout point de $\mathbb{R}^2$.

On admet la *formule de Taylor-Young* à l'ordre 2 pour les fonctions de plusieurs variables. Elle est similaire à la formule de Taylor-Young pour les fonctions d'une variable :

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + o\left(h^2\right) \text{ quand } h \to 0 \\ &= f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + h^2\varepsilon(h) \text{ avec } \lim_{h\to 0} \varepsilon(h) = 0 \end{aligned}$$

Cette formule va nous être utile pour décider si un point critique est un extremum.

Théorème 3 (Taylor-Young à l'ordre 2)**.** *Soit U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}$ une fonction de classe C^2 et $a \in U$. Alors il existe une fonction ε, définie sur un voisinage de $0_{\mathbb{R}^n}$ et à valeurs dans $\mathbb{R}$, vérifiant :*

$$f(a+h) = f(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \|h\|^2 \varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^n}} \varepsilon(h) = 0.$$

Remarque.

Remarque.

Exemple : Écrire la formule de Taylor-Young à l'ordre 2 de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 + 4xy - 1$ en tout point de $\mathbb{R}^2$ puis en $(0, 0)$. Écrire la formule matricielle.

Comme la matrice Hessienne de f en un point est symétrique réelle, elle est orthodiagonalisable par le théorème spectral. Cela va permettre d'obtenir une condition suffisante pour avoir un extremum local en un point critique, en utilisant la formule de Taylor-Young à l'ordre 2.

Proposition 8 (Condition suffisante d'extremum sur un ouvert)**.** *Soit U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ et $a \in U$ un point critique de f.*

(i) Si toutes les valeurs propres de $H_f(a)$ sont strictement positives, alors f admet un minimum local strict en a.

(ii) Si toutes les valeurs propres de $H_f(a)$ sont strictement négatives, alors f admet un maximum local strict en a.

(*iii*) *Si $H_f(a)$ admet au moins une valeur propre strictement positive et une strictement négative, alors f n'admet pas d'extremum local en a. On dit que a est un* ***point col****, ou* ***point selle****.*

Démonstration. La formule de Taylor-Young matricielle à l'ordre 2 donne en $a \in U$:

$$f(a+h) = f(a) + J_f(a)H + \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^n}} \varepsilon(h) = 0.$$

Comme a est un point critique de f, la proposition 7 montre que la matrice jacobienne $J_f(a)$ de f en a est nulle. D'où :

$$f(a+h) - f(a) = \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^n}} \varepsilon(h) = 0.$$

Le signe de $f(a+h) - f(a)$ est donné par le signe de $\frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h)$ au voisinage de $0_{\mathbb{R}^n}$.

Comme la matrice $H_f(a)$ est symétrique et à coefficients réels, il existe une matrice orthogonale $P \in O_n(\mathbb{R})$ telle que $P^{-1}H_f(a)P = D$ est diagonale, de coefficients diagonaux $\lambda_1, \ldots, \lambda_n$. Soit $h \in \mathbb{R}^n$. On note H sa matrice coordonnées dans la base canonique de $\mathbb{R}^n$ et H' sa matrice coordonnées dans la base orthonormée formés par les vecteurs propres de $H_f(a)$ (qui sont aussi les colonnes de la matrice P). La formule de changement de base pour les vecteurs donne $H = PH'$. Comme P est une matrice orthogonale, les vecteurs H et H' ont la même norme euclidienne. Alors :

$$\begin{aligned} q(h) &= \frac{1}{2}{}^tHH_f(a)H \\ &= \frac{1}{2}{}^t(PH')H_f(a)(PH') \\ &= \frac{1}{2}{}^tH'{}^tPH_f(a)PH' \\ &= \frac{1}{2}{}^tH'DH' \\ &= \frac{1}{2}\left(\lambda_1^2h_1'^2 + \cdots + \lambda_n^2h_n'^2\right) \end{aligned}$$

- Si toutes les valeurs propres de $H_f(a)$ sont strictement positives, c'est-à-dire $\lambda_i > 0$ pour tout $i \in \{1, \ldots, n\}$, alors en supposant par exemple $\lambda_1 = \min(\lambda_1, \ldots, \lambda_n)$:

$$\begin{aligned} q(h) &= \frac{1}{2}\left(\lambda_1h_1'^2 + \cdots + \lambda_nh_n'^2\right) \\ &\geq \frac{1}{2}\left(\lambda_1h_1'^2 + \cdots + \lambda_1'h_n'^2\right) \\ &= \frac{1}{2}\lambda_1\|H'\|_2^2 \\ &= \frac{1}{2}\lambda_1\|H\|_2^2 \\ &= \frac{1}{2}\lambda_1\|h\|_2^2 \end{aligned}$$

Alors :

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \\ &\geq \frac{1}{2}\lambda_1\|h\|_2^2 + \|h\|^2\varepsilon(h) \\ &= \|h\|_2^2\left(\frac{1}{2}\lambda_1 + \varepsilon(h)\right) \\ &\geq 0 \end{aligned}$$

pour h assez proche de $0_{\mathbb{R}^n}$. Ainsi la fonction f atteint un minimum local en a.

- Si toutes les valeurs propres de $H_f(a)$ sont strictement négatives, c'est-à-dire $\lambda_i < 0$ pour tout $i \in \{1,\dots,n\}$, alors en supposant par exemple $\lambda_n = \max(\lambda_1,\dots,\lambda_n)$:

$$\begin{aligned} q(h) &= \frac{1}{2}\left(\lambda_1 h_1'^2 + \cdots + \lambda_n h_n'^2\right) \\ &\leq \frac{1}{2}\left(\lambda_n h_1'^2 + \cdots + \lambda_n h_n'^2\right) \\ &= \frac{1}{2}\lambda_n \|H'\|_2^2 \\ &= \frac{1}{2}\lambda_n \|H\|_2^2 \\ &= \frac{1}{2}\lambda_n \|h\|_2^2 \end{aligned}$$

  Alors :

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}{}^t H H_f(a) H + \|h\|^2 \varepsilon(h) \\ &\leq \frac{1}{2}\lambda_n \|h\|_2^2 + \|h\|^2\varepsilon(h) \\ &= \|h\|_2^2 \left(\frac{1}{2}\lambda_n + \varepsilon(h)\right) \\ &\leq 0 \end{aligned}$$

  pour h assez proche de $0_{\mathbb{R}^n}$. Ainsi la fonction f atteint un maximum local en a.

- Si par exemple $\lambda_1 < 0$ et $\lambda_2 > 0$, alors d'une part en notant h_1 un vecteur propre normé de $H_f(a)$ associé à la valeur propre λ_1 (dont on note H_1 la matrice coordonnées dans la base canonique de $\mathbb{R}^n$) et $t \in \mathbb{R}$ assez petit :

$$\begin{aligned} f(a+th_1) - f(a) &= \frac{1}{2}{}^t(tH_1) H_f(a)(tH_1) + \|th_1\|_2^2 \varepsilon(th_1) \\ &= \frac{1}{2}t^2\, {}^t H_1 \lambda_1 H_1 + |t|^2 \|h_1\|_2^2 \varepsilon(th_1) \\ &= \frac{1}{2}t^2 \lambda_1 \|h_1\|_2^2 + t^2 \varepsilon(th_1) \\ &= t^2 \left(\frac{\lambda_1}{2} + \varepsilon(th_1)\right) \\ &\leq 0 \end{aligned}$$

  D'autre part en notant h_2 un vecteur propre normé de $H_f(a)$ associé à la valeur propre λ_2 (dont on note H_2 la matrice coordonnées dans la base canonique de $\mathbb{R}^n$) et $t \in \mathbb{R}$ assez petit :

$$\begin{aligned} f(a+th_2) - f(a) &= \frac{1}{2}{}^t(tH_2) H_f(a)(t) H_2 + \|th_2\|_2^2 \varepsilon(th_2) \\ &= \frac{1}{2}t^2\, {}^t H_2 \lambda_2 H_2 + |t|^2 \|h_2\|_2^2 \varepsilon(th_2) \\ &= \frac{1}{2}t^2 \lambda_2 \|h_2\|_2^2 + t^2 \varepsilon(th_2) \\ &= t^2 \left(\frac{\lambda_2}{2} + \varepsilon(th_2)\right) \\ &\geq 0 \end{aligned}$$

  Ainsi a n'est pas un extremum de f.

□

Remarque :

On peut utiliser la proposition 8 dans $\mathbb{R}^2$ sans même calculer les valeurs propres de la matrice hessienne.

Corollaire 2. *Soit U un ouvert de $\mathbb{R}^2$, $f : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction de classe C^2 et a un point critique de f.*

(i) Si $\det(H_f(a)) > 0$ et $\mathrm{Tr}(H_f(a)) > 0$, alors f admet un minimum local strict en a.

(ii) Si $\det(H_f(a)) > 0$ et $\mathrm{Tr}(H_f(a)) < 0$, alors f admet un maximum local strict en a.

(iii) Si $\det(H_f(a)) < 0$, alors f n'admet pas d'extremum en a.

Démonstration.

□

Remarque.

Exemple : Déterminer les extrema locaux de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^3 + y^3 - 3xy$.

Exemple : Déterminer les extrema locaux de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2\,(1 + x + y)$.

6.3.3 Plan tangent à une surface

La formule de Taylor-Young à l'ordre 2 pour une fonction de classe C^2 sur un ouvert U de $\mathbb{R}^2$ donne :

$$\begin{aligned} f((x,y)+(h_1,h_2)) \quad = \quad & f(x,y) + \frac{\partial f}{\partial x}(x,y)h_1 + \frac{\partial f}{\partial y}(x,y)h_2 + \\ & \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(x,y)h_1^2 + \frac{\partial^2 f}{\partial x \partial y}(x,y)h_1h_2 + \frac{\partial^2 f}{\partial y \partial x}(x,y)h_2h_1 + \frac{\partial^2 f}{\partial y^2}(x,y)h_2^2\right) + \\ & \|(h_1,h_2)\|^2 \varepsilon(h_1,h_2) \text{ avec } \lim_{(h_1,h_2)\to(0,0)} \varepsilon(h_1,h_2) = 0 \end{aligned}$$

Pour $(x_0, y_0) \in U$, on aurait pu l'écrire sous la forme :

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + \\ & \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0,y_0)(x-x_0)^2 + \frac{\partial^2 f}{\partial x \partial y}(x,y)(x-x_0)(y-y_0) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(x_0,y_0)(y-y_0) + \\ & \|(x-x_0, y-y_0)\|^2 \varepsilon(x,y) \text{ avec } \lim_{(x,y)\to(x_0,y_0)} \varepsilon(x,y) = 0 \end{aligned}$$

Il s'agit en fait d'un analogue d'un développement limité en dimension 2. Comme pour définir la tangente à une courbe $y = f(x)$ dans le plan où on prend le premier terme du développement limité de f, on procède de même pour la surface $z = f(x, y)$.

Définition 6. *Soit U un ouvert de $\mathbb{R}^2$, $f : U \longrightarrow \mathbb{R}$ une fonction de classe C^1 et S la surface d'équation $z = f(x, y)$. Si $M_0 = (x_0, y_0, z_0)$ est un point de S, alors une équation du plan tangent à S en M_0 est :*

$$z = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0).$$

Exemple : soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^2+y^2$ et S la surface d'équation $z = f(x, y)$.

Si f est de classe C^2, alors la position de S par rapport à son plan tangent en M_0 dépend du signe des valeurs propres de la matrice hessienne de f en M_0.

Proposition 9 (Position d'une surface et de son plan tangent)**.** *Soit U un ouvert de $\mathbb{R}^2$, $f : U \longrightarrow \mathbb{R}$ de classe C^2 et S la surface d'équation $z = f(x, y)$. Soit $M_0 = (x_0, y_0, z_0)$ un point de S et P le plan tangent à la surface S au point M_0.*

(i) Si les deux valeurs propres de $H_f(x_0, y_0)$ sont strictement positives, alors la surface S est au-dessus de son plan tangent P au voisinage de M_0.

(ii) Si les deux valeurs propres de $H_f(x_0, y_0)$ sont strictement négatives, alors la surface S est au-dessous de son plan tangent P au voisinage de M_0.

(iii) Si $H_f(x_0, y_0)$ admet une valeur propre strictement positive et une strictement négative, alors la surface S traverse son plan tangent.

Démonstration. On veut déterminer la position, au voisinage de M_0, de la surface S d'équation $z = f(x, y)$ et de son plan tangent P d'équation $z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$. On pose donc la fonction $g : U \longrightarrow \mathbb{R}$ définie par :

$$g(x,y) \;=\; f(x,y) - \left(f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) \right)$$

La fonction g est de classe C^2 sur l'ouvert U. On veut savoir si g est à valeurs positives ou à valeurs négatives au voisinage de M_0. Or :

$$\begin{aligned} g(x_0,y_0) &= f(x_0,y_0) - \left(f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x_0-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y_0-y_0) \right) \\ &= 0 \end{aligned}$$

Ainsi on veut savoir si (x_0, y_0) est un extremum local de g. Tout d'abord, on calcule :

$$\begin{aligned} \frac{\partial g}{\partial x}(x_0,y_0) &= \frac{\partial f}{\partial x}(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial g}{\partial y}(x_0,y_0) &= \frac{\partial f}{\partial y}(x_0,y_0) - \frac{\partial f}{\partial y}(x_0,y_0) \\ &= 0 \end{aligned}$$

Ainsi (x_0, y_0) est un point critique de g. On peut donc appliquer la proposition 8 à g. On calcule donc la hessienne de g en (x_0, y_0) :

$$\begin{aligned} H_g(x_0,y_0) &= \begin{pmatrix} \frac{\partial^2 g}{\partial x^2}(x_0,y_0) & \frac{\partial^2 g}{\partial x \partial f}(x_0,y_0) \\ \frac{\partial^2 g}{\partial x \partial f}(x_0,y_0) & \frac{\partial^2 g}{\partial y^2}(x_0,y_0) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0,y_0) & \frac{\partial^2 f}{\partial x \partial f}(x_0,y_0) \\ \frac{\partial^2 f}{\partial x \partial f}(x_0,y_0) & \frac{\partial^2 f}{\partial y^2}(x_0,y_0) \end{pmatrix} \\ &= H_f(x_0,y_0) \end{aligned}$$

(i) Si les deux valeurs propres de $H_f(x_0, y_0)$ sont strictement positives, alors les deux valeurs propres de $H_g(x_0, y_0)$ sont strictement positives. Donc la proposition 8 montre que (x_0, y_0) est un minimum local de g. Autrement dit, $g(x, y) \geq g(x_0, y_0)$, c'est-à-dire $g(x, y) \geq 0$ pour (x, y) au voisinage de (x_0, y_0). Autrement dit, la surface S est au-dessus de son plan tangent P au voisinage de M_0.

(ii) Si les deux valeurs propres de $H_f(x_0, y_0)$ sont strictement négatives, alors les deux valeurs propres de $H_g(x_0, y_0)$ sont strictement négatives. Donc la proposition 8 montre que (x_0, y_0) est un maximum local de g. Autrement dit, $g(x, y) \leq g(x_0, y_0)$, c'est-à-dire $g(x, y) \leq 0$ pour (x, y) au voisinage de (x_0, y_0). Autrement dit, la surface S est au-dessous de son plan tangent P au voisinage de M_0.

(iii) Si $H_f(x_0, y_0)$ admet une valeur propre strictement positive et une strictement négative, alors $H_g(x_0, y_0)$ admet une valeur propre strictement positive et une strictement négative. Donc la proposition 8 montre que (x_0, y_0) n'est pas un extremum de g. Autrement dit, la surface S traverse son plan tangent.

□

Exemple : Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^2 + y^2$ et S la surface d'équation $z = f(x, y)$. Déterminer la position du plan tangent à la surface S en $(0, 0, 0)$ et en $(1, 1, 2)$ au voisinage de ces deux points.

Exemple : Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^4 - x^3 + xy - y^2$ et S la surface d'équation $z = f(x, y)$. Déterminer le plan tangent P à la surface S en $(0, 0, 0)$ ainsi que la position relative de S et de son plan tangent au voisinage de $(0, 0, 0)$.

6.4 Extrema sur un compact de $\mathbb{R}^n$

Rappelons le théorème donnant l'existence des extrema pour une fonction continue sur un compact.

Théorème 4 (Théorème des bornes atteintes)**.** *Soit K un compact non vide de $\mathbb{R}^n$ et $f : K \longrightarrow \mathbb{R}$ une fonction continue. Alors f admet un minimum global et un maximum global sur K.*

Remarque : le théorème 4 montre que les extrema d'une fonction continue sur un compact K de $\mathbb{R}^n$ sont donc atteints :

- soit à l'intérieur de K (que l'on note K°) qui est ouvert (c'est même le plus grand ouvert inclus dans K), et ils font alors partie des points critiques de f ;
- soit sur la frontière de K, $\mathrm{Fr}(K) = \overline{K} \backslash K^\circ = K \backslash K^\circ$ puisque K est fermé.

Méthode d'étude :

- On se place sur le plus grand ouvert incus dans K, à savoir K°, et on trouve les points critiques de f sur K°.
- On étudie les points d'extrema sur la frontière en paramétrant celle-ci (quitte à décomposer la frontière en plusieurs morceaux) : si $x \in \mathrm{Fr}(K)$ est un point d'extremum local de f sur K, la restriction $f_{|_{\mathrm{Fr}(K)}}$ admet alors un extremum local en ce point (par exemple, si $n = 2$, on se retrouve alors avec des fonctions d'une seule variable que l'on sait étudier).
- On compare les valeurs trouvées en ces différents points pour savoir où se situe le maximum/minimum.

Exemple : Soit $T = \big\{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0,\ x + y \leq 1\big\}$ et une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = xy(1 - x - y)$. Montrer que f admet un maximum sur T que l'on déterminera.