

Chapitre X

Calcul différentiel

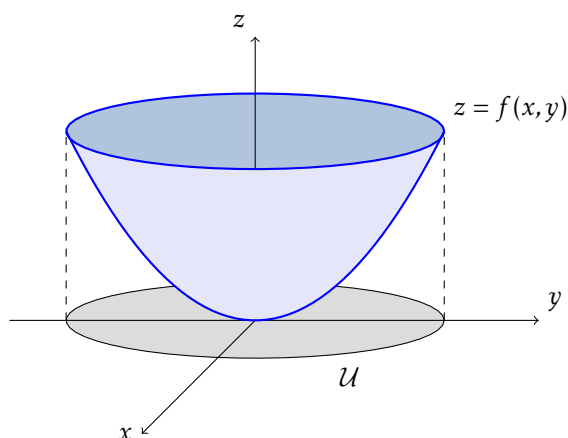
Jusqu'à présent, nous nous sommes cantonnés à l'étude de fonctions d'une *variable*, d'abord à valeurs réelles ou complexes puis à valeurs vectorielles (dans \mathbb{R}^n), ces fonctions étaient systématiquement définies sur un intervalle I de \mathbb{R} . Nous allons maintenant nous intéresser aux fonctions de *plusieurs variables*, c'est à dire définies sur une partie \mathcal{U} de \mathbb{R}^p , à valeurs dans \mathbb{R}^n :

$$f : \left(\begin{array}{l} \mathcal{U} \subset \mathbb{R}^p \longrightarrow \mathbb{R}^n \\ x = (x_1, \dots, x_p) \longmapsto f(x) = (f_1(x_1, \dots, x_p), \dots, f_n(x_1, \dots, x_p)) \end{array} \right)$$

Dans ce cours, nous aurons l'occasion, comme pour les fonctions vectorielles, de montrer que l'étude d'une telle fonction se ramène à celle de ses fonctions coordonnées f_1, \dots, f_n et ainsi nous ramener à l'étude des fonctions à valeurs réelles (autrement dit prendre $n = 1$). Pour des raisons pratiques, nos exemples se cantonneront le plus souvent à des fonctions à deux ou trois variables ($p = 2$ ou 3).

Ainsi, lorsque l'on a $p = 2$ et $n = 1$, le graphe $z = f(x, y)$ d'une telle fonction est une nappe paramétrée que l'on peut visualiser et ainsi fournir un support à une interprétation géométrique :

Exemple. $\mathcal{U} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$, et $f : \left(\begin{array}{l} \mathcal{U} \longrightarrow \mathbb{R} \\ (x, y) \longmapsto x^2 + y^2 \end{array} \right)$



D'un point de vue historique, on peut noter que la notion de fonction à plusieurs variables apparaît très tôt en physique, où l'on étudie souvent des quantités dépendants de plusieurs paramètres. Citons par exemple :

- en mécanique des fluides, la *pression* p est un champ¹ scalaire qui associe à un point du fluide la pression en ce point ; mathématiquement, cela correspond à une application d'une partie \mathcal{U} de \mathbb{R}^3 (ou de \mathbb{R}^4 , si on tient compte du temps) dans \mathbb{R} :

$$p : \left(\begin{array}{l} \mathcal{U} \longrightarrow \mathbb{R} \\ M = (x, y, z) \longmapsto p(M) \end{array} \right)$$

- en électromagnétisme, la *densité de courant* \vec{j} est un champ vectoriel qui associe à tout point de l'espace considéré un vecteur qui décrit le courant électrique qui circule à l'échelle locale ; mathématiquement, cela correspond à une application d'une partie \mathcal{U} de \mathbb{R}^3 dans \mathbb{R}^3 :

$$\vec{j} : \left(\begin{array}{l} \mathcal{U} \longrightarrow \mathbb{R}^3 \\ M = (x, y, z) \longmapsto \vec{j}(M) \end{array} \right)$$

Mais avant de débiter l'étude du concept de différentiabilité, nous allons revenir un instant sur les notions de limite et de continuité, déjà abordées dans le chapitre consacré aux espaces vectoriels normés.

1. en mathématiques, un champ est une application qui associe aux points de l'espace une valeur, scalaire ou vectorielle.

1. Calcul différentiel

Dans cette section, E et F désigneront deux \mathbb{R} -espaces vectoriels normés de dimensions finies, la norme étant notée $\|\cdot\|$ indépendamment de l'espace, mais le plus souvent nous auront $E = \mathbb{R}^2$ ou \mathbb{R}^3 et $F = \mathbb{R}$. \mathcal{U} désignera une partie de E (le plus souvent un ouvert), et $f : \mathcal{U} \rightarrow F$ une fonction à plusieurs variables.

1.1 Étude locale d'une application

Dans le chapitre consacré aux espaces vectoriels normés nous avons donné la définition de la limite de f en un point adhérent à \mathcal{U} :

$f(x)$ admet $\ell \in F$ pour limite lorsque x tend vers a lorsque :

$$\forall \epsilon > 0, \exists \eta > 0 \mid \forall x \in \mathcal{U}, \quad \|x - a\| \leq \eta \Rightarrow \|f(x) - \ell\| \leq \epsilon$$

En outre, lorsque $a \in \mathcal{U}$ et $\ell = f(a)$, f est dite continue en a .

Observons sur deux exemples comment se traduit cette définition dans le cadre des fonctions à plusieurs variables.

Exemple. Considérons la fonction $f_1 : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$ définie par $f_1(x, y) = \frac{x^2 y}{x^2 + y^2}$, et utilisons la norme euclidienne canonique sur $\mathbb{R}^2 : \|(x, y)\|_2 = \sqrt{x^2 + y^2}$.

Sachant que $|x| \leq \|(x, y)\|$ et $|y| \leq \|(x, y)\|$, nous pouvons affirmer que $|f_1(x, y)| \leq \|(x, y)\|$, ce qui implique :

$$\lim_{(x,y) \rightarrow (0,0)} f_1(x, y) = 0.$$

Exemple. Considérons la fonction $f_2 : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$ définie par $f_2(x, y) = \frac{x^2 y}{x^4 + y^2}$.

Supposons que cette fonction possède une limite ℓ en $(0,0)$. Les théorèmes de composition des limites impliquent que pour toute fonction vectorielle $\phi : I \subset \mathbb{R} \rightarrow \mathcal{U}$ pour laquelle $\lim_{t \rightarrow 0} \phi(t) = (0,0)$ nous avons : $\lim_{t \rightarrow 0} f_2 \circ \phi(t) = \ell$.

Or $\lim_{t \rightarrow 0} f_2(t, t) = \lim_{t \rightarrow 0} \frac{t}{t^2 + 1} = 0$ et $\lim_{t \rightarrow 0} f_2(t, t^2) = \frac{1}{2}$. Ainsi f_2 ne possède pas de limite en $(0,0)$.

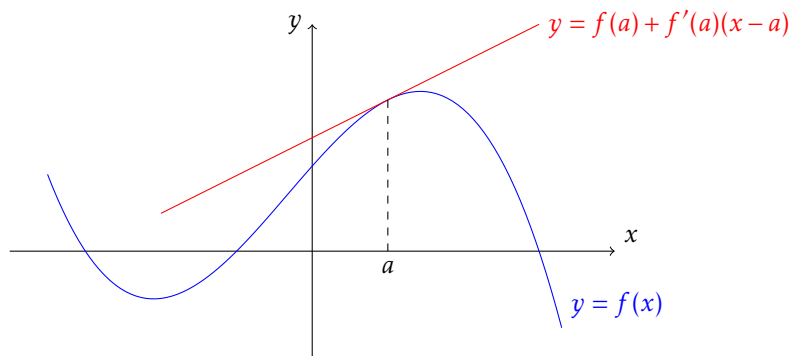
Exercice 1

Déterminer si les fonctions suivantes, définies sur $\mathbb{R}^2 \setminus \{(0,0)\}$, ont une limite finie en $(0,0)$:

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}, \quad g(x, y) = \frac{|x + y|}{x^2 + y^2}, \quad h(x, y) = (x + y) \sin\left(\frac{1}{x^2 + y^2}\right).$$

1.2 Applications différentiables

Pour comprendre comment nous allons généraliser la notion de dérivée aux fonctions à plusieurs variables, observons l'interprétation géométrique qu'on peut faire de la dérivée en $a \in I$ d'une fonction à une variable $f : I \rightarrow \mathbb{R}$.



Au voisinage de a , le graphe de f est approché par une droite, sa tangente. Autrement dit, f est localement approchée par la fonction affine $x \mapsto f(a) + f'(a)(x - a)$, ce qui se traduit par le développement limité suivant :

$$f(x) \underset{a}{=} f(a) + f'(a)(x - a) + o(x - a).$$

qu'on peut écrire de façon équivalente :

$$f(a + h) \underset{0}{=} f(a) + f'(a)h + o(h).$$

Cette approximation affine est formée d'une constante $f(a)$ et d'une application linéaire $h \mapsto f'(a)h$. Ceci nous conduit à adopter la définition suivante :

DÉFINITION. — Soient E et F deux \mathbb{R} -espaces vectoriels normés de dimensions finies, \mathcal{U} un ouvert de E et $f : \mathcal{U} \rightarrow F$ une application.

On dira que f est différentiable en $a \in \mathcal{U}$ lorsqu'il existe une application linéaire $u \in \mathcal{L}(E, F)$ telle que :

$$f(a + h) \underset{0_E}{=} f(a) + u(h) + o(\|h\|).$$

Dans ce cas, l'application linéaire u est appelée la *différentielle* de f en a , et sera notée $df(a)$. Ainsi, on écrira :

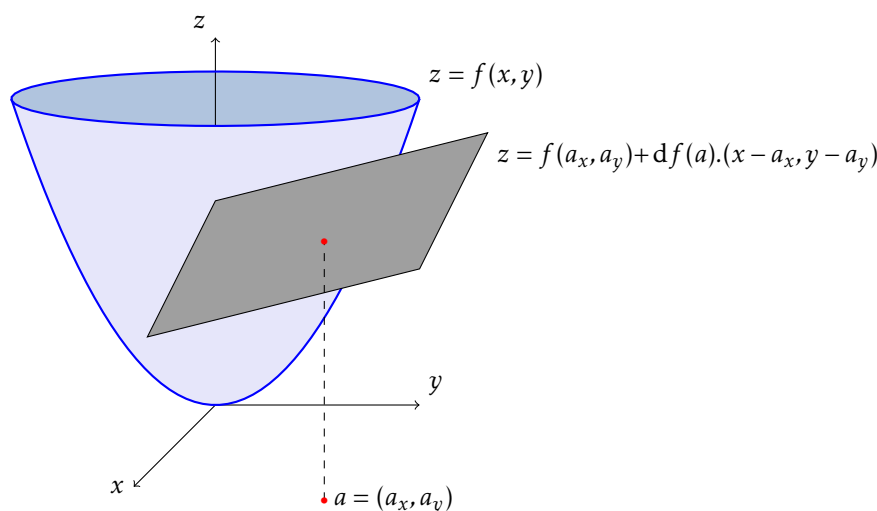
$$\boxed{f(a + h) \underset{0_E}{=} f(a) + df(a).h + o(\|h\|)} \quad \text{ou encore :} \quad \boxed{f(x) \underset{a}{=} f(a) + df(a).(x - a) + o(\|x - a\|)}$$

Notons que la différentiabilité de f en a entraîne *a fortiori* la continuité de f en a , puisqu'une application linéaire est continue (en dimension finie).

Exemple. Dans le cas où $E = \mathbb{R}^2$ et $F = \mathbb{R}$, les fonctions affines prennent la forme suivante :

$$\tilde{u} : \begin{pmatrix} \mathbb{R}^2 & \mapsto & \mathbb{R} \\ (x, y) & \mapsto & \alpha + \beta x + \gamma y \end{pmatrix}$$

et le graphe de cette fonction affine est le plan affine d'équation $z = \alpha + \beta x + \gamma y$. Autrement dit, la nappe d'équation $z = f(x, y)$ est localement approchée par un plan.



Exercice 2

Soit $f : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathcal{M}_n(\mathbb{R})$ l'application définie par $f(M) = M^2$. Montrer que f est différentiable en tout point $A \in \mathcal{M}_n(\mathbb{R})$, et déterminer l'application linéaire $df(A)$.

Remarque. Lorsque f est une application linéaire, l'égalité $f(x) = f(a) + f(x - a)$ montre que sa différentielle en a est égale à elle-même : pour tout $a \in E$, $df(a) = f$.

différence entre dérivée et différentielle

Dans le cas des fonctions numériques, la dérivée de f en a est le réel $f'(a)$, alors que la différentielle de f en a est l'application linéaire $x \mapsto f'(a)x$. En effet, les applications linéaires de \mathbb{R} dans \mathbb{R} s'écrivent de manière unique sous la forme : $x \mapsto \lambda x$, avec $\lambda \in \mathbb{R}$.

On peut se rapprocher encore de la notion de dérivée lorsque $F = \mathbb{R}$: dans ce cas, la différentielle $df(a)$ de f en a est une application linéaire de E dans \mathbb{R} , c'est à dire une *forme linéaire*. Or lorsque E est un espace euclidien, nous avons vu que les formes linéaires sur E s'écrivent de manière unique sous la forme $x \mapsto \langle \ell | x \rangle$, avec $\ell \in E$. Cela conduit à la définition :

DÉFINITION. — Lorsque $f : \mathcal{U} \subset E \rightarrow \mathbb{R}$ est différentiable en $a \in E$, il existe un unique vecteur de E , noté $\nabla f(a)$ tel que :

$$\forall h \in E, \quad df(a).h = \langle \nabla f(a) | h \rangle.$$

Le vecteur $\nabla f(a)$ est appelé le gradient de f en a .

Exercice 3

Soit E un espace euclidien, et $f : E \rightarrow \mathbb{R}$ défini par $f(x) = \langle x | x \rangle$. Montrer que f est différentiable en tout $a \in E$, et déterminer le vecteur $\nabla f(a)$.

1.3 Dérivée directionnelle et dérivées partielles

Nous allons maintenant nous attacher à voir comment calculer une différentielle (ou un gradient) dans le cas où $E = \mathbb{R}^p$ et $F = \mathbb{R}$. L'idée que nous allons suivre est d'essayer, autant que faire se peut, de se ramener à des calculs de dérivées de fonctions d'une seule variable.

Sachant que f est définie sur un ouvert \mathcal{U} de \mathbb{R}^p , l'application partielle $t \mapsto f(a + tv)$ est, quel que soit le vecteur $v \in \mathbb{R}^p$, $v \neq 0$, une fonction vectorielle définie au voisinage de 0. Lorsque cette fonction est dérivable en 0, on note $D_v f(a)$ cette quantité, qu'on appelle la *dérivée en a selon le vecteur v* .

En particulier, lorsque $v = e_k$ est le k^e vecteur de la base canonique de \mathbb{R}^p , cette application prend la forme suivante :

$$t \mapsto f(a_1, a_2, \dots, a_{k-1}, a_k + t, a_{k+1}, \dots, a_p)$$

et sa dérivée en 0 est notée $\partial_k f(a)$ ou $\frac{\partial f}{\partial x_k}(a)$, quantité qu'on appelle la *k^e dérivée partielle d'ordre 1 de f en a* . Autrement dit :

$$\forall k \in \llbracket 1, p \rrbracket, \quad \partial_k f(a) = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_k + t, \dots, a_p) - f(a_1, \dots, a_k, \dots, a_p)}{t}.$$

PROPOSITION 1.1 — Lorsque f est différentiable en a , f admet en a des dérivées partielles d'ordre 1 et pour tout

$$h = (h_1, \dots, h_p) \in \mathbb{R}^p, \quad df(a).h = \sum_{k=1}^p \partial_k f(a) h_k. \quad \text{Autrement dit,} \quad \nabla f(a) = (\partial_1 f(a), \dots, \partial_p f(a)).$$

La réciproque de ce résultat est fautive : une fonction peut posséder des dérivées partielles en a sans être différentiable en ce point. Cependant, en renforçant un peu les hypothèses on dispose du résultat suivant, avec lequel nous allons maintenant justifier l'existence de la différentielle :

THÉORÈME 1.2 — Soit \mathcal{U} un ouvert de \mathbb{R}^p , et $f : \mathcal{U} \rightarrow \mathbb{R}$ une fonction. On suppose que :

- (i) f possède pour tout $k \in \llbracket 1, p \rrbracket$ et en tout point a de \mathcal{U} une dérivée partielle $\partial_k f(a)$;
- (ii) pour tout $k \in \llbracket 1, p \rrbracket$, l'application $\partial_k f : \mathcal{U} \rightarrow \mathbb{R}$ est continue sur \mathcal{U} .

Alors f est différentiable en tout point a de \mathcal{U} .

Une application f vérifiant ces hypothèses sera dorénavant dite de classe \mathcal{C}^1 sur \mathcal{U} .

Exemple. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ défini par $f(x, y) = x^2y$. f admet en tout point (x, y) des dérivées partielles $\frac{\partial f}{\partial x}(x, y) = 2xy$ et $\frac{\partial f}{\partial y}(x, y) = x^2$ à l'évidence continues, donc f est de classe \mathcal{C}^1 et $\nabla f(x, y) = 2xy\vec{e}_1 + x^2\vec{e}_2$.

La différentielle s'écrit donc $df(x, y) : (h_1, h_2) \mapsto 2xyh_1 + x^2h_2$.

Exercice 4

On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = \frac{x^2y}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$. La fonction f est-elle de classe \mathcal{C}^1 ?

Expression de la dérivée directionnelle pour une fonction de classe \mathcal{C}^1

Lorsque f est de classe \mathcal{C}^1 , on a donc $f(a + tv) \underset{t \rightarrow 0}{=} f(a) + df(a).(tv) + o(\|tv\|) \underset{t \rightarrow 0}{=} f(a) + t df(a).(v) + o(t)$ donc

$$D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = df(a).(v)$$

1.4 Règle de la chaîne

Considérons une fonction $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , I un intervalle et $x_k : I \rightarrow \mathbb{R}$, $1 \leq k \leq p$ des fonctions elles aussi de classe \mathcal{C}^1 telles que pour tout $t \in I$, $(x_1(t), \dots, x_p(t)) \in \mathcal{U}$. On peut dès lors définir la fonction $\phi : I \rightarrow \mathbb{R}$ par $\phi(t) = f(x_1(t), \dots, x_p(t))$.

Le résultat suivant, appelé *règle de la dérivation en chaîne*, ou plus simplement *règle de la chaîne*, donne la règle de dérivation d'une telle fonction.

THÉORÈME 1.3 (règle de la chaîne) — Si f et toutes les fonctions x_1, \dots, x_p sont de classe \mathcal{C}^1 , il en est de même de la fonction ϕ , et

$$\forall t \in I, \quad \phi'(t) = \sum_{k=1}^p x'_k(t) \partial_k f(x_1(t), \dots, x_p(t)) = \sum_{k=1}^p x'_k(t) \frac{\partial f}{\partial x_k}(x_1(t), \dots, x_p(t)).$$

Exercice 5

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 , et $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par : $\forall (r, \theta) \in \mathbb{R}^2$, $g(r, \theta) = f(r \cos \theta, r \sin \theta)$. Calculer les dérivées partielles de g en fonction de celles de f , et en déduire l'expression du gradient en coordonnées polaires.

PROPOSITION 1.4 — Soit \mathcal{U} un ouvert convexe de \mathbb{R}^p , et $f : \mathcal{U} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . Alors f est constante sur \mathcal{U} si et seulement si pour tout $a \in \mathcal{U}$, $df(a) = 0$, autrement dit si et seulement si les fonctions $\partial_1 f, \dots, \partial_p f$ sont nulles sur \mathcal{U} .

1.5 Dérivées partielles d'ordre deux

Soit \mathcal{U} un ouvert de \mathbb{R}^p , et $f : \mathcal{U} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . Les applications $\partial_i f = \frac{\partial f}{\partial x_i}$ sont des applications définies et continues de \mathcal{U} dans \mathbb{R} et à ce titre peuvent être elles-mêmes de classe \mathcal{C}^1 . Lorsque c'est le cas, on note $\partial_j \partial_i f = \frac{\partial^2 f}{\partial x_j \partial x_i}$ la dérivée partielle par rapport à la j^{e} variable de $\partial_i f$.

Remarque. Dans le cas particulier où $i = j$ on utilisera plutôt la notation $\partial_i^2 f$ ou $\frac{\partial^2 f}{\partial x_i^2}$.

DÉFINITION. — Une application $f : \mathcal{U} \rightarrow \mathbb{R}$ est dite de classe \mathcal{C}^2 lorsqu'elle est de classe \mathcal{C}^1 et lorsque pour tout $i \in \llbracket 1, p \rrbracket$, $\partial_i f$ est de classe \mathcal{C}^1 sur \mathcal{U} .

A priori, l'expression $\partial_j \partial_i f$ signifie que l'on dérive d'abord par rapport à x_i , puis par rapport à x_j . cependant, le théorème suivant, que nous admettrons, montre qu'il n'en est rien dans le cas d'une fonction de classe \mathcal{C}^2 .

THÉORÈME 1.5 (Théorème de Schwarz) — Soit $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 . Alors pour $(i, j) \in \llbracket 1, p \rrbracket^2$, $\partial_j \partial_i f = \partial_i \partial_j f$.

Exercice 6

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 , et $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par : $\forall (r, \theta) \in \mathbb{R}^2, g(r, \theta) = f(r \cos \theta, r \sin \theta)$.

a. Calculer les dérivées partielles secondes de g en fonction de celles de f .

On appelle *laplacien* de f la quantité : $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$.

b. Dédurre des calculs précédents l'expression du laplacien en coordonnées polaires (c'est à dire en fonction des dérivées de g).

■ Équations aux dérivées partielles

En sciences physiques il est fréquent d'avoir à résoudre une équation mêlant les dérivées partielles d'ordre 1 ou 2 d'une même fonction. Le phénomène de propagation des ondes peut par exemple être modélisé par l'équation de d'Alembert $\Delta f - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0$. En dimension 1, cette équation s'écrit $\frac{\partial^2 f}{\partial x^2}(x, t) - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2}(x, t) = 0$.

Il n'y a pas de méthode générale de résolution, mais celle-ci passe le plus souvent par l'utilisation d'un changement de variable.

On appelle *changement de variable de classe \mathcal{C}^k* une application bijective $\phi : \mathcal{U} \rightarrow \mathcal{V}$ entre deux ouverts \mathcal{U} et \mathcal{V} de \mathbb{R}^p telle que ϕ et ϕ^{-1} soient toutes deux de classe \mathcal{C}^k . Nous nous restreindrons à deux types de changement de variable : les changements de variables affines et le changement de variable en coordonnées polaires.

Un changement de variable *affine* consiste simplement à poser $\begin{cases} u = ax + by + e \\ v = cx + dy + f \end{cases}$ qui se traduit matriciellement par $Y = AX + B$ avec $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et $B = \begin{pmatrix} e \\ f \end{pmatrix}$. Pour qu'il soit bijectif, il faut et il suffit que $\det A \neq 0$, et il s'agit alors d'un changement de variable de classe \mathcal{C}^∞ .

Par exemple, pour résoudre l'équation de propagation en dimension 1, on posera $\begin{cases} u = x + ct \\ v = x - ct \end{cases}$ et $g(u, v) = f(x, t)$.

Cette dernière équation introduit une nouvelle fonction inconnue g , et peut être comprise de deux façons :

- $g(u, v) = f\left(\frac{u+v}{2}, \frac{u-v}{2c}\right)$ (pour définir g à partir de f);
- $f(x, t) = g(x+ct, x-ct)$ (pour retrouver f une fois déterminé g).

On calcule successivement à l'aide de la règle de la chaîne :

$$\begin{aligned} \frac{\partial f}{\partial x}(x, t) &= \frac{\partial g}{\partial u}(x+ct, x-ct) + \frac{\partial g}{\partial v}(x+ct, x-ct) & \frac{\partial f}{\partial t}(x, t) &= c \frac{\partial g}{\partial u}(x+ct, x-ct) - c \frac{\partial g}{\partial v}(x+ct, x-ct) \\ \frac{\partial^2 f}{\partial x^2}(x, t) &= \frac{\partial^2 g}{\partial u^2}(u, v) + 2 \frac{\partial^2 g}{\partial u \partial v}(u, v) + \frac{\partial^2 g}{\partial v^2}(u, v) & \frac{\partial^2 f}{\partial t^2}(x, t) &= c^2 \frac{\partial^2 g}{\partial u^2}(u, v) - 2c^2 \frac{\partial^2 g}{\partial u \partial v}(u, v) + c^2 \frac{\partial^2 g}{\partial v^2}(u, v) \end{aligned}$$

et alors : $\frac{\partial^2 f}{\partial x^2}(x, t) - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2}(x, t) = 0 \iff 4 \frac{\partial^2 g}{\partial u \partial v}(u, v) = 0$.

Cette équation est maintenant aisément résoluble : la fonction $\frac{\partial g}{\partial v}$ est indépendante de u donc ne dépend que de v . La fonction g s'écrit donc sous la forme $g(u, v) = \phi(u) + \psi(v)$, où ϕ et ψ sont deux fonctions quelconques de classe \mathcal{C}^2 . Il reste à revenir à f en concluant que $f(x, t) = \phi(x+ct) + \psi(x-ct)$.

Le changement de variables en coordonnées polaires

En toute rigueur, pour que le changement de variable $\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$ soit bijectif et de classe \mathcal{C}^∞ ainsi que sa réciproque, il est nécessaire d'imposer $r \in]0, +\infty[$ et $\theta \in]\alpha, \alpha + 2\pi[$, ce qui restreint (x, y) à $\mathbb{R}^2 \setminus \mathcal{D}_\alpha$, où \mathcal{D}_α est la demi-droite fermée issue de l'origine et d'angle α par rapport à l'axe Ox .

Une fois ceci précisé, l'équation $f(x, y) = g(r, \theta)$ peut s'écrire $g(r, \theta) = f(r \cos \theta, r \sin \theta)$. En revanche, on observera que faute d'une expression simple de θ en fonction de x et de y , il sera plus difficile d'exprimer $f(x, y)$ en fonction de g, x et y , aussi va-t-on dans ce cas calculer les dérivées partielles de g en fonction de celles de f :

$$\frac{\partial g}{\partial r}(r, \theta) = \cos \theta \frac{\partial f}{\partial x}(x, y) + \sin \theta \frac{\partial f}{\partial y}(x, y) \quad \text{et} \quad \frac{\partial g}{\partial \theta}(r, \theta) = -r \sin \theta \frac{\partial f}{\partial x}(x, y) + r \cos \theta \frac{\partial f}{\partial y}(x, y)$$

Par exemple, l'équation aux dérivées partielles $y \frac{\partial f}{\partial x}(x, y) - x \frac{\partial f}{\partial y}(x, y) = 0$ s'écrit $\frac{\partial g}{\partial \theta}(r, \theta) = 0$ en coordonnées polaires, soit $g(r, \theta) = \phi(r)$ où ϕ est une fonction de classe \mathcal{C}^1 , et $f(x, y) = \phi(\sqrt{x^2 + y^2})$.

De même, l'équation $x \frac{\partial f}{\partial x}(x, y) + y \frac{\partial f}{\partial y}(x, y) = 0$ s'écrit $r \frac{\partial g}{\partial r}(r, \theta) = 0$ en coordonnées polaires, soit $g(r, \theta) = \psi(\theta)$ où ψ est une fonction de classe \mathcal{C}^1 . Ici, faute d'une expression convenable pour la fonction θ , on se contentera de conclure que $f(x, y) = \psi(\theta(x, y))$.

■ Matrice Hessienne

Nous avons montré que la formule $f(a+h) \underset{h \rightarrow 0}{=} f(a) + hf'(a) + o(h)$ valable pour une fonction numérique de classe \mathcal{C}^1 se généralise au cas d'une fonction $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 par la formule :

$$f(a+h) \underset{h \rightarrow 0}{=} f(a) + \langle \nabla f(a) | h \rangle + o(\|h\|) \underset{h \rightarrow 0}{=} f(a) + \nabla f(a)^T h + o(\|h\|)$$

en identifiant les vecteurs de \mathbb{R}^p et les matrices colonnes de $\mathcal{M}_{p,1}(\mathbb{R})$.

Nous admettrons que la formule $f(a+h) \underset{h \rightarrow 0}{=} f(a) + hf'(a) + \frac{h^2}{2} f''(a) + o(h^2)$ se généralise pour une fonction $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 par la formule :

$$f(a+h) \underset{h \rightarrow 0}{=} f(a) + \nabla f(a)^T h + \frac{1}{2} h^T H_f(a) h + o(\|h\|^2)$$

où $H_f(a) \in \mathcal{M}_p(\mathbb{R})$ est la matrice des dérivées partielles secondes de f en a , autrement dit :

$$[H_f(a)]_{i,j} = \partial_i \partial_j f(a)$$

On notera que cette matrice, appelée *matrice hessienne* de f en a , est une matrice symétrique (d'après le théorème de Schwarz).

1.6 Extremums locaux

Dans cette section, nous considérons un domaine (non nécessairement ouvert) \mathcal{U} de \mathbb{R}^p ainsi qu'une fonction $f : \mathcal{U} \subset \mathbb{E} \rightarrow \mathbb{R}$.

DÉFINITION. — On dit que f présente en un point $a \in \mathcal{U}$ un maximum local lorsqu'il existe un réel $r > 0$ tel que pour tout $x \in B(a, r) \cap \mathcal{U}$, $f(x) \leq f(a)$.

On dit que f présente en $a \in \mathcal{U}$ un maximum global lorsque pour tout $x \in \mathcal{U}$, $f(x) \leq f(a)$.

On définit de la même façon les notions de minimum local et de minimum global.

Remarque. De cette définition il résulte immédiatement que tout extremum global est un extremum local, la réciproque n'étant bien évidemment pas vraie.

THÉORÈME 1.6 — Soit \mathcal{U} un ouvert de \mathbb{R}^p , $f : \mathcal{U} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 , et $a \in \mathcal{U}$ un point en lequel f présente un extremum local. Alors : $df(a) = 0$ (la forme linéaire nulle).

Remarque. La condition $df(a) = 0$, qui peut s'écrire $\nabla f(a) = 0_E$, est donc une condition *nécessaire* mais non *suffisante* pour que f présente un extremum local en a dans l'ouvert \mathcal{U} .

Un point a en lequel $\nabla f(a) = 0_E$ est appelé un *point critique* de f .

Évidemment, la question de la réciproque se pose : un point critique est-il nécessairement un extremum local ? La réponse est négative : ne serait-ce qu'en dimension 1, la fonction $t \mapsto t^3$ présente en 0 un point critique qui n'est pas un extremum local.

Il va néanmoins être possible par l'affirmative dans certains cas, grâce à la formule de Taylor à l'ordre 2. Ainsi, supposons f de classe \mathcal{C}^2 , et considérons un point critique $a \in \mathcal{U}$ de f . Nous avons alors :

$$f(a+h) = f(a) + \frac{1}{2} h^T H_f(a) h + o(\|h\|^2)$$

Nous savons par ailleurs que la matrice hessienne $H_f(a)$ est symétrique ; elle est donc ortho-diagonalisable. Soit donc (e) une base orthonormée formée de vecteurs propres de $H_f(a)$. Si on décompose le vecteur h dans cette

$$\text{base : } h = \sum_{k=1}^p h_k e_k \text{ alors } h^T H_f(a) h = \sum_{k=1}^p \lambda_k h_k^2.$$

THÉORÈME 1.7 — Si f est de classe \mathcal{C}^2 et a un point critique de f , alors :

- si $\text{Sp}(H_f(a)) \subset \mathbb{R}_+^*$ (autrement dit si $H_f(a) \in \mathcal{S}_p^{++}(\mathbb{R})$), f présente en a un minimum local strict ;
- si $\text{Sp}(H_f(a)) \subset \mathbb{R}_-^*$, f présente en a un maximum local strict ;
- si $H_f(a)$ possède deux valeurs propres de signes différents alors f ne présente pas d'extremum en a .

Remarque. Notons que cette étude n'est pas exhaustive ; en particulier lorsque la matrice hessienne n'est pas inversible, $H_f(a)$ admet 0 pour valeur propre et on ne peut conclure.

Le cas de la dimension 2

Lorsque $p = 2$, posons $H_f(a) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$; autrement dit $r = \partial_1^2 f(a)$, $s = \partial_1 \partial_2 f(a)$ et $t = \partial_2^2 f(a)$.

Les deux valeurs propres λ et μ de $H_f(a)$ vérifient $\lambda + \mu = \text{tr } H_f(a) = r + t$ et $\lambda \mu = \det H_f(a) = rt - s^2$ donc ces deux valeurs propres sont non nulles et de même signe si et seulement si $\det H_f(a) > 0$. Ainsi, f présente en a un extremum local strict si et seulement si $\det H_f(a) > 0$, et cet extremum est :

- un minimum si $\text{tr } H_f(a) > 0$;
- un maximum si $\text{tr } H_f(a) < 0$.

Exemple. Considérons la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x^3 + y^3 - 3xy$.

Ses points critiques vérifient :

$$\begin{cases} \partial_1 f(x, y) = 0 \\ \partial_2 f(x, y) = 0 \end{cases} \iff \begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}$$

et la résolution de ce système donne deux points critiques $a = (0, 0)$ et $b = (1, 1)$.

On calcule $\partial_1^2 f(x, y) = 6x$, $\partial_1 \partial_2 f(x, y) = 6y$ et $\partial_2^2 f(x, y) = -3$ donc $H_f(a) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$ et $H_f(b) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$.

$\text{Sp } H_f(a) = \{-3, 3\}$ donc f ne présente pas d'extremum local en a (il s'agit d'un point selle).

$\text{Sp } H_f(b) = \{3, 9\}$ donc f présente en b un minimum local.

Exercice 7

Déterminer les points critiques de la fonction $f : (x, y) \mapsto y(x^2 + (\ln y)^2)$ sur $\mathbb{R} \times \mathbb{R}_+^*$, puis déterminer s'il s'agit d'extremums locaux ou pas.

■ Recherche d'extremums globaux sur une partie fermée bornée de \mathbb{R}^p

Considérons maintenant une partie \mathcal{K} fermée et bornée de \mathbb{R}^p et une application $f : \mathcal{K} \rightarrow \mathbb{R}$ continue, sur \mathcal{K} , de classe \mathcal{C}^1 sur $\overset{\circ}{\mathcal{K}}$.

Dans le chapitre consacré aux espaces vectoriels normés, nous avons admis le résultat suivant, qui va nous être de nouveau utile :

Rappel. Si \mathcal{K} est une partie fermée et bornée de \mathbb{R}^p et $f : \mathbb{R}^p \rightarrow \mathbb{R}$ une fonction continue, alors f est bornée et atteint ses bornes sur \mathcal{K} .

Ce résultat assure l'existence d'un minimum et d'un maximum global sur \mathcal{K} . Ces deux extremums se trouvent ou bien sur la frontière $\text{Fr}(\mathcal{K})$ de \mathcal{K} , ou bien dans l'intérieur $\overset{\circ}{\mathcal{K}} = \mathcal{K} \setminus \text{Fr}(\mathcal{K})$ de \mathcal{K} . En d'autres termes, les extremums globaux sont à chercher :

- sur la frontière de \mathcal{K} ;
- et parmi les points critiques de l'intérieur de \mathcal{K} .

Exercice 8

Soit $\mathcal{K} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, x + y \leq 1\}$, et $f : \mathcal{K} \rightarrow \mathbb{R}$ définie par $f(x, y) = xy(1 - x - y)$. Déterminer la valeur maximale prise par la fonction f .