

FONCTIONS DE PLUSIEURS VARIABLES

G. FICHOU

Table des matières

1	Introduction	5
2	L'espace \mathbb{R}^d	6
2.1	Produit scalaire, norme et distance dans \mathbb{R}^d	6
2.2	Produit vectoriel dans \mathbb{R}^3	8
2.3	Coordonnées polaires, cylindriques, sphériques	9
2.4	Topologie de \mathbb{R}^d	10
2.5	Suites dans \mathbb{R}^d	12
2.6	Ensembles compacts	12
3	Fonctions de plusieurs variables	14
3.1	Définitions	14
3.2	Représentation géométrique	16
3.3	Fonctions continues de \mathbb{R}^d dans \mathbb{R}	17
3.4	Etude de certaines surfaces quadratiques	20
3.5	Fonctions continues sur un ensemble compact	22
4	Courbes paramétrées	24
4.1	Introduction : courbes dans le plan	24
4.1.1	Définition, exemples	24
4.1.2	Longueur de courbes	24
4.1.3	La courbe de Peano, la courbe de Koch	24
4.2	Courbes paramétrées	25
4.2.1	Définitions	25
4.2.2	Quelques exemples	28
5	Dérivées des fonctions de plusieurs variables	29
5.1	Dérivées partielles des fonctions à valeurs réelles	29
5.1.1	Rappels	29
5.1.2	Dérivée partielle	29

5.1.3	Interprétation géométrique	30
5.1.4	Gradient	30
5.1.5	Dérivées partielles et continuité	30
5.1.6	Dérivation composée	31
5.1.7	Accroissements finis	31
5.1.8	Dérivée selon un vecteur	31
5.2	La différentielle d'une fonction à valeurs réelles	32
5.2.1	Règle de différentiation	34
5.2.2	Remarques	34
5.2.3	Dérivées partielles successives	34
5.3	La différentielle d'une fonction à valeurs vectorielles	35
5.3.1	Propriétés de la différentielle	36
5.3.2	Différentielles des fonctions composées	36
6	Sous-ensembles de \mathbb{R}^n et fonctions	38
6.1	Nappes paramétrées	38
6.2	Tangentes aux courbes (surfaces, hypersurfaces) de niveau	39
6.3	Fonctions implicites – Inversion locale	39
6.3.1	Inversion locale	39
6.3.2	Fonctions implicites : cas $f(x, y) = 0$	40
6.3.3	Fonctions implicites : cas $f(x_1 \dots x_n) = 0$	41
7	Optimisation libre et sous contraintes	42
7.1	En dimension 1	42
7.2	Extrema locaux de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$	42
7.3	Formule de Taylor	43
7.4	Extrema de f sur un compact $K \subset \mathbb{R}^2$	44
7.5	Extrema liés (multiplicateur de Lagrange)	44
7.5.1	Une seule contrainte	44
7.5.2	Plusieurs contraintes	45
8	Intégration des fonctions de \mathbb{R}^n dans \mathbb{R}	47

8.1	Intégration des fonctions d'une variable	47
8.2	Volume de parties bornées de \mathbb{R}^d	48
8.3	Intégration des fonctions de deux variables	49
8.4	Calcul des intégrales doubles	50
8.5	Intégration sur les régions bornées de \mathbb{R}^2	50
8.6	Intégrale double et changement de variables	51
8.7	Intégrales triples	53
8.8	Quelques calculs classiques	53
8.8.1	L'aire d'un disque	53
8.8.2	Le volume de la boule	54
8.8.3	Solides de révolution	54
8.8.4	Le volume d'une pyramide	55

1 Introduction

Le cours porte sur les fonctions de plusieurs variables. Le terme n'est pas très précis. Prenons quelques exemples. En physique : le temps la distance la charge... En économie : le prix en fonction du capital et du travail. On peut imaginer des fonctions dont les variables ou les valeurs sont discrètes ou qualitatives. Dans le cours nous nous intéresserons au cas des fonctions définies sur des parties de \mathbb{R}^d à valeurs dans \mathbb{R}^d .

2 L'espace \mathbb{R}^d

2.1 Produit scalaire, norme et distance dans \mathbb{R}^d

Définition 2.1. Si $x = (x_1 \dots x_d)$ et $y = (y_1 \dots y_d)$ sont deux vecteurs de \mathbb{R}^d , on définit leur produit scalaire par :

$$\langle x, y \rangle = x_1 y_1 + \dots + x_d y_d$$

Définition 2.2. On appelle **norme** de x (ou longueur) $\|x\| = \langle x, x \rangle^{1/2}$ et la **distance** entre deux vecteurs $d(x, y) = \|x - y\|$.

Proposition 2.3. On a les propriétés suivantes :

- (1) $\langle x, y \rangle = \langle y, x \rangle$
- (2) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$
- (3) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- (4) $\langle x, x \rangle \geq 0$ avec $\langle x, x \rangle = 0$ si et seulement si $x = 0$

Théorème 2.4. Le produit scalaire vérifie l'**inégalité de Cauchy-Schwarz** $\langle x, y \rangle^2 \leq \|x\|^2 \|y\|^2$ avec égalité si et seulement si x et y sont colinéaires.

Démonstration

Soient x et y deux vecteurs et λ un nombre réel. Le nombre $\langle \lambda x + y, \lambda x + y \rangle$ est positif ou nul. On peut développer ce produit scalaire en appliquant les propriétés de linéarité. On obtient :

$$\begin{aligned} \langle \lambda x + y, \lambda x + y \rangle &= \langle \lambda x, \lambda x + y \rangle + \langle y, \lambda x + y \rangle \\ &= \lambda \langle x, \lambda x + y \rangle + \langle y, \lambda x + y \rangle \\ &= \lambda \langle x, \lambda x \rangle + \lambda \langle x, y \rangle + \langle y, \lambda x \rangle + \langle y, y \rangle \\ &= \lambda^2 \langle x, x \rangle + \lambda \langle x, y \rangle + \lambda \langle y, x \rangle + \langle y, y \rangle \\ &= \lambda^2 \langle x, x \rangle + 2\lambda \langle x, y \rangle + \langle y, y \rangle \end{aligned}$$

On obtient donc un polynôme de degré 2 en λ qui est positif ou nul pour tout λ . Cela signifie que le discriminant de ce polynôme est négatif ou nul. Autrement dit on a

$$\langle x, y \rangle^2 - \langle y, y \rangle \langle x, x \rangle \leq 0.$$

Dire que ce discriminant est nul est équivalent à dire qu'il existe λ pour lequel le polynôme est nul ou encore qu'il existe λ pour lequel $\|\lambda x + y\|^2 = 0$ ce qui est équivalent à $\lambda x + y = 0$. Autrement dit l'égalité ne peut être vérifiée que dans les cas où x et y sont colinéaires. On vérifie facilement que l'égalité est satisfaite lorsque x et y sont colinéaires (on a donc l'équivalence). \square

Donnons une autre façon de voir les choses. L'identité de Lagrange

$$\left(\sum_{i=1}^d x_i y_i\right)^2 + \sum_{1 \leq i < j \leq d} (x_i y_j - x_j y_i)^2 = \left(\sum_{i=1}^d x_i^2\right) \left(\sum_{i=1}^d y_i^2\right)$$

s'écrit ici

$$\langle x, y \rangle^2 + \sum_{1 \leq i < j \leq d} (x_i y_j - x_j y_i)^2 = \|x\|^2 \|y\|^2$$

et donne donc une justification à notre énoncé (la somme $\sum_{1 \leq i < j \leq n} (x_i y_j - x_j y_i)^2$ est positive ou nulle). Établissons l'identité de Lagrange.

$$\begin{aligned} \left(\sum_{i=1}^d x_i y_i\right)^2 &= \sum_{i=1}^d x_i^2 y_i^2 + \sum_{1 \leq i < j \leq d} x_i y_i x_j y_j + \sum_{1 \leq j < i \leq d} x_i y_i x_j y_j \\ &= \sum_{i=1}^d x_i^2 y_i^2 + 2 \sum_{1 \leq i < j \leq d} x_i y_i x_j y_j \\ &= \sum_{i=1}^d x_i^2 y_i^2 + \sum_{1 \leq i < j \leq d} [x_i^2 y_j^2 + x_j^2 y_i^2 - (x_i y_j - x_j y_i)^2] \\ &= \sum_{i=1}^d x_i^2 y_i^2 + \sum_{1 \leq i < j \leq d} [x_i^2 y_j^2 + x_j^2 y_i^2] - \sum_{1 \leq i < j \leq d} (x_i y_j - x_j y_i)^2 \\ &= \left(\sum_{i=1}^d x_i^2\right) \left(\sum_{i=1}^d y_i^2\right) - \sum_{1 \leq i < j \leq d} (x_i y_j - x_j y_i)^2 \end{aligned}$$

□

Théorème 2.5. La norme définie précédemment s'appelle **norme euclidienne** et vérifie :

- (i) $\|x\| = 0$ si et seulement si $x = 0$
- (ii) $\|x\| > 0$ si $x \neq 0$
- (iii) $\|\alpha x\| = |\alpha| \|x\|$
- (iv) $\|x + y\| \leq \|x\| + \|y\|$

Démonstration Si $\|x\| = 0$ alors la somme $\sum_{i=1}^n x_i^2$ vaut 0. Comme tous les carrés sont positifs ou nuls, pour que leur somme soit nulle il faut qu'il soient tous nuls, autrement dit que tous les x_i soient nuls ce qui signifie bien que $x = 0$. Réciproquement, si $x = 0$ alors $\|x\| = 0$. On a

$$\|\alpha x\|^2 = \sum_{i=1}^n (\alpha x_i)^2 = \alpha^2 \sum_{i=1}^n x_i^2 = \alpha^2 \|x\|^2.$$

En prenant la racine carrée on obtient

$$\|\alpha x\| = \sqrt{\|\alpha x\|^2} = \sqrt{\alpha^2 \|x\|^2} = \sqrt{\alpha^2} \sqrt{\|x\|^2} = |\alpha| \|x\|.$$

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y \rangle \\
&= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&= \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \\
&\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

□

L'inégalité de Cauchy Schwarz permet aussi de définir la mesure d'un angle géométrique entre deux vecteurs : comme $\langle x, y \rangle \leq \|x\|\|y\|$, si aucun des deux vecteurs n'est nul, alors le quotient $\frac{\langle x, y \rangle}{\|x\|\|y\|}$ est un nombre compris entre -1 et 1.

Définition 2.6. La mesure d'un **angle** entre deux vecteurs non nuls est $\theta \in [0, \pi]$ vérifiant $\cos \theta = \frac{\langle x, y \rangle}{\|x\|\|y\|}$.

Définition 2.7. x et y de \mathbb{R}^n sont dits **orthogonaux** lorsque $\langle x, y \rangle = 0$.

Définition 2.8. (plan dans \mathbb{R}^3)

Soient $A = (x_0, y_0, z_0)$ un point de \mathbb{R}^3 et $N = (a, b, c)$ un vecteur non nul. Le plan passant par A et orthogonal à N est $P = \{x \in \mathbb{R}^3 / (x - A) \cdot N = 0\}$.

2.2 Produit vectoriel dans \mathbb{R}^3

Définition 2.9. Si $x = (x_1, x_2, x_3)$ et $y = (y_1, y_2, y_3)$ sont deux vecteurs de \mathbb{R}^3 , on définit le **produit vectoriel** de x et de y par : $x \wedge y = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - y_1 x_2)$.

Voici une autre définition du produit vectoriel. Supposons x et y fixés et considérons l'application

$$z = (z_1, z_2, z_3) \mapsto \det(x, y, z).$$

C'est une application linéaire de \mathbb{R}^3 dans \mathbb{R} . Il existe donc des coefficients a_1, a_2, a_3 tels que, pour tous x, y on ait :

$$\det(x, y, z) = a_1 z_1 + a_2 z_2 + a_3 z_3.$$

Si on note a le vecteur (a_1, a_2, a_3) alors cela s'écrit encore

$$\det(x, y, z) = \langle a, z \rangle.$$

On vérifie que le vecteur a ainsi défini a pour coordonnées $(x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - y_1 x_2)$, c'est le produit vectoriel de x et y .

Pour qui sait calculer un déterminant 3x3 en développant par rapport à la troisième colonne cela donne une façon simple de retrouver les coordonnées d'un produit vectoriel :

$$\det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} = (x_2 y_3 - x_3 y_2)z_1 - (x_1 y_3 - x_3 y_1)z_2 + (x_1 y_2 - y_1 x_2)z_3.$$

Théorème 2.10. *On a les propriétés suivantes :*

- (1) $x \wedge y = -y \wedge x$
- (2) $x \wedge (y + z) = x \wedge y + x \wedge z$
- (3) $\alpha x \wedge y = x \wedge \alpha y = \alpha(x \wedge y)$
- (4) $\langle x, x \wedge y \rangle = 0$ et $\langle y, x \wedge y \rangle = 0$
- (5) $\|x \wedge y\|^2 = \|x\|^2 \|y\|^2 - (x \cdot y)^2$ (identité de Lagrange ; déjà vu)

Démonstration : Je donne une démonstration basée sur les propriétés des déterminants (vous n'avez peut-être pas encore vu ces propriétés en AL2 ; cela viendra). Vous pouvez aussi vérifier ces propriétés en utilisant la définition analytique et en développant les expressions obtenues.

- (1) $\langle x \wedge y, z \rangle = \det(x, y, z) = -\det(y, x, z) = -\langle y \wedge x, z \rangle$
- (2) $\langle x \wedge (y + z), u \rangle = \det(x, y + z, u) = \det(x, y, u) + \det(x, z, u) = \langle x \wedge y, u \rangle + \langle x \wedge z, u \rangle = \langle x \wedge y + x \wedge z, u \rangle$
- (3) $\langle \alpha x \wedge y, u \rangle = \det(\alpha x, y, u) = \alpha \det(x, y, u) = \alpha \langle x \wedge y, u \rangle$
- (4) $\langle x, x \wedge y \rangle = \det(x, y, x) = 0$ et $\langle y, x \wedge y \rangle = \det(x, y, y) = 0$

□

Interprétation géométrique de $x \wedge y$

$\|x \wedge y\| = \|x\| \|y\| \sin \theta$ est l'aire du parallélogramme engendré par x et y .

Démonstration : L'aire d'un parallélogramme est donnée par le produit longueur de la base fois hauteur. Un peu de trigonométrie montre que c'est égal à $\|x\| \|y\| \sin \theta$. Le point (5) ci-dessus donne

$$\|x \wedge y\|^2 = \|x\|^2 \|y\|^2 - (x \cdot y)^2 = \|x\|^2 \|y\|^2 - \|x\|^2 \|y\|^2 \cos^2 \theta = \|x\|^2 \|y\|^2 \sin^2 \theta.$$

□

2.3 Coordonnées polaires, cylindriques, sphériques

Plutôt que de repérer un point (x, y) du plan \mathbb{R}^2 par ses coordonnées cartésiennes dans le repère orthonormé formé par la base canonique, on peut le faire au moyen de sa distance à l'origine et de l'angle formé par le premier vecteur de la base canonique et le vecteur (x, y) .

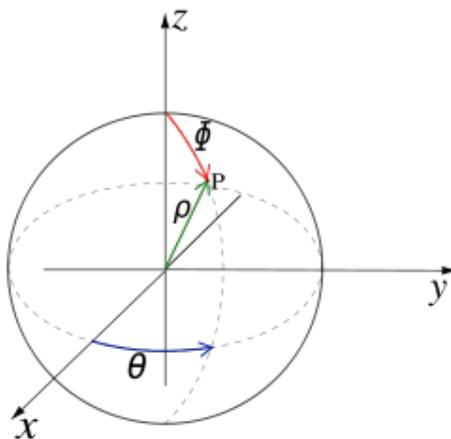
La distance à l'origine est définie au moyen du produit scalaire comme ci-dessus. L'angle n'est pas déterminé de manière unique. Plusieurs choix sont possibles. On peut ainsi définir les coordonnées polaires d'un point du plan au moyen de l'application suivante :

$$]0, +\infty[\times]0, 2\pi[\rightarrow \mathbb{R}^2 \quad (\rho, \theta) \mapsto (\rho \cos \theta, \rho \sin \theta).$$

On aurait pu choisir (le choix est tout aussi bon) de faire varier θ dans $[-\pi, \pi[$. On n'attribue généralement pas de coordonnées polaires au point origine : il est facile de définir sa distance à l'origine, l'angle n'aurait pas de sens.

Dans \mathbb{R}^3 on définit les coordonnées sphériques d'un point au moyen de l'application

$$]0, +\infty[\times]0, 2\pi[\times]0, \pi[\rightarrow \mathbb{R}^3 \quad (\rho, \theta, \phi) \mapsto (\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi).$$



Le couple $(\rho \sin \phi, \theta)$ forme les coordonnées polaires de la projection du point sur le plan d'équation $z = 0$. Là encore on aurait pu choisir d'autres intervalles pour domaines de θ et ϕ . En géographie par exemple la latitude qui correspond à ϕ varie de -90 à 90 degrés et c'est l'angle avec le plan de l'équateur qui la définit (pas l'angle avec l'axe pôle sud pôle nord). Pour une illustration très parlante des coordonnées sphériques on pourra regarder le premier chapitre du film dimensions¹

2.4 Topologie de \mathbb{R}^d

Définition 2.11. Soient $a \in \mathbb{R}^d$ et $r > 0$.

On appelle $B(a, r) = \{x \in \mathbb{R}^d / \|x - a\| < r\}$ la **boule ouverte** de centre a et de rayon r .

Exemple

Dans \mathbb{R} , \mathbb{R}^2 ou \mathbb{R}^3 on retrouve les intervalles, les disques, les boules ouvertes.

1. http://www.dimensions-math.org/Dim_fr.htm

Proposition 2.12. Soient $A \subset \mathbb{R}^d$, $x \in \mathbb{R}^d$.

Alors une des trois conditions suivantes est vérifiée :

- (i) $\exists r > 0$ tel que $B(x, r) \subset A$
- (ii) $\exists r > 0$ tel que $B(x, r) \subset A^c$ où $A^c = \mathbb{R}^n \setminus A$
- (iii) $\forall r > 0$, $B(x, r)$ contient des points de A et de A^c .

Définition 2.13. L'intérieur de A (noté $\text{int}(A)$ ou $\overset{\circ}{A}$) est l'ensemble des points de \mathbb{R}^d vérifiant (i).

L'extérieur de A (noté $\text{ext } A$) est l'ensemble des points de \mathbb{R}^d vérifiant la condition (ii).

La frontière (ou le bord) de A (notée ∂A) est l'ensemble des points de \mathbb{R}^d vérifiant la condition (iii).

La fermeture de A (notée \overline{A}) est la réunion de A et de ∂A .

Exemples dans \mathbb{R}^2

$$A = \{x \in \mathbb{R}^2 / \|x\| < 1\}$$

$$A = \{(n, 0) / n \in \mathbb{Z}\}$$

Définition 2.14. Un ensemble A de \mathbb{R}^d est :

- (i) **ouvert** si $\forall a \in A$, $\exists r > 0$ tel que $B(a, r) \subset A$
- (ii) **fermé** si A^c est ouvert.

Proposition 2.15. A est ouvert si et seulement si $\overset{\circ}{A} = A$.

A est fermé si et seulement si $\overline{A} = A$.

La frontière de A est égale à $\partial A = \overline{A} \setminus \overset{\circ}{A}$.

Démonstration. Détailler le dernier point. \square

Exemples

$A_1 = \{(x, y) / x^2 + y^2 < 1\}$ est ouvert.

$A_2 = \{(x, y) / x^2 + y^2 \leq 1\}$ est fermé.

$A_3 = A_1 \cup \{(1, 0)\}$ n'est ni ouvert ni fermé.

$]0, 1[\subset \mathbb{R}$ est ouvert dans \mathbb{R} .

$]0, 1[\times \{0\} \subset \mathbb{R}^2$ n'est ni ouvert ni fermé.

$[0, 1] \subset \mathbb{R}$ est fermé dans \mathbb{R} .

$[0, 1] \times \{0\} \subset \mathbb{R}^2$ est fermé dans \mathbb{R}^2 .

Proposition 2.16. 1. \mathbb{R}^d et \emptyset sont ouverts (et donc aussi fermés).

2. Toute réunion d'ouverts est un ouvert.

3. Toute intersection finie d'ouverts est un ouvert.

2.5 Suites dans \mathbb{R}^d

Définition 2.17. Une suite dans \mathbb{R}^d est une famille de vecteurs $x_n = (x_{1,n}, \dots, x_{d,n})$ indexée par l'ensemble des entiers naturels $(x_n)_{n \in \mathbb{N}}$. Chaque terme de la suite x_n est un vecteur avec ses d coordonnées.

Définition 2.18. Une suite $(x_n)_{n \in \mathbb{N}}$ **converge** dans \mathbb{R}^d vers $b \in \mathbb{R}^d$ si $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ tel que $n \geq N$ entraîne $\|x_n - b\| < \varepsilon$.

De manière équivalente on peut définir la convergence d'une suite de vecteurs (x_n) par la convergence de chacune des suites réelles données par les coordonnées $x_{i,n}$, i allant de 1 à d , n variant dans \mathbb{N} (les suites des coordonnées sont indexées par n et il y en a d : $(x_{i,n})_{n \in \mathbb{N}}$).

Une autre façon de dire que la suite (x_n) tend vers b est de dire que la suite réelle de nombre positifs ou nuls $(d(x_n, b))_{n \in \mathbb{N}}$ tend vers 0.

Remarques

1. On dit que b est **la limite** de la suite (x_n) et on note $x_n \rightarrow b$.
2. $x_n \rightarrow b$ si et seulement si $\forall \varepsilon > 0$ la boule $B(b, \varepsilon)$ contient toute la suite sauf un nombre fini de x_n .

Proposition 2.19. A est fermé si et seulement si pour toute suite convergente contenue dans A et convergente, la limite est dans A .

Cette proposition fournit un critère pour démontrer qu'un ensemble A n'est pas fermé : il suffit de trouver une suite de points de A convergeant vers un point n'appartenant pas à A .

Théorème 2.20. Soit (x_n) une suite bornée. Il existe une sous-suite de (x_n) convergeant dans \mathbb{R}^d .

2.6 Ensembles compacts

Définition 2.21. $X \subset \mathbb{R}^d$ est compact si X est fermé et borné (borné veut dire qu'il existe $R > 0$ tel que $X \subset B(0, R)$).

Exemples

$[0, 23]$ est un compact dans \mathbb{R} .

$\{(x, y) / x^2 + (y - 2)^2 \leq 6\}$ est un compact de \mathbb{R}^2 .

$[2, 3] \times [1, 3] \times [5, 7]$ est un compact dans \mathbb{R}^3 .

Théorème 2.22. (*Bolzano-Weierstrass*)

Soit $X \subset \mathbb{R}^d$ compact.

Alors toute suite $(x_n) \subset X$ contient une sous-suite (x_{l_n}) qui converge vers un point de X .

3 Fonctions de plusieurs variables

3.1 Définitions

Définition 3.1. Une fonction f définie sur un sous-ensemble D de \mathbb{R}^d à valeurs dans \mathbb{R} (où D est un) s'appelle fonction numérique de n variables.

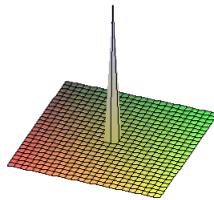
D est le domaine de définition de f .

$\{f(x) / x \in D\}$ est l'image de f .

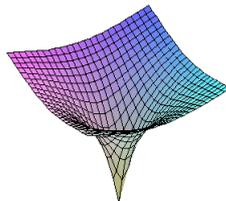
$\{(x, f(x)) / x \in D\} \subseteq \mathbb{R}^d \times \mathbb{R}$ est appelé graphe de f .

Exemples ²

$$f(x, y) = \frac{1}{\sqrt{x^2 + y^2}}$$



$$f(x, y, z) = \text{Ln}(1 + x^2 + y^2)$$



Définition 3.2. Soient D et E deux parties de \mathbb{R}^d telles que $D \subset E$ et f et g deux fonctions définies respectivement sur D et E . On dit que g est un prolongement de f à E si pour tout $x \in D$ on a $f(x) = g(x)$. Dans cette situation, on dit aussi que f est la restriction de g à D .

Exemple

$f(x, y) = \frac{x^3}{x^2 + y^2}$ qu'on prolonge en une fonction g définie sur \mathbb{R}^2 en posant $g(0, 0) = a$,

2. Les images données sont obtenues avec le logiciel Maple.

où $a \in \mathbb{R}$.

Définition 3.3. Soient D une partie de \mathbb{R}^d , f une fonction définie sur D à valeurs dans \mathbb{R} et $x_0 \in D$.

On dit que f est majorée sur D s'il existe $M \in \mathbb{R}$ tel que pour tout $x \in D$ on ait $f(x) \leq M$.

On dit que f est minorée sur D s'il existe $m \in \mathbb{R}$ tel que pour tout $x \in D$ on ait $f(x) \geq m$.

On dit que f est bornée sur D si elle est à la fois majorée et minorée. Cela revient à dire qu'il existe $M \geq 0$ tel que pour tout $x \in D$ on ait $|f(x)| \leq M$.

Exemple

La fonction g précédente est bornée sur \mathbb{R}^2 .

Définition 3.4. On dit que f a un minimum en x_0 si pour tout $x \in D$ on a $f(x) \geq f(x_0)$.

On dit que f a un maximum en x_0 si pour tout $x \in D$ on a $f(x) \leq f(x_0)$.

On dit que f a un minimum strict en x_0 si pour tout $x \in D$, $x \neq x_0$, on a $f(x) > f(x_0)$.

On dit que f a un maximum strict en x_0 si pour tout $x \in D$, $x \neq x_0$, on a $f(x) < f(x_0)$.

Exemple

Pour g , le point $(0, 1)$ est un minimum non strict, alors que $(1, 0)$ est maximum non strict.

Définition 3.5. On dit que f a un minimum local en x_0 s'il existe $r > 0$ tel que pour tout $x \in B(x_0, r)$ on a $f(x) \geq f(x_0)$.

On dit que f a un maximum local en x_0 s'il existe $r > 0$ tel que pour tout $x \in B(x_0, r)$ on a $f(x) \leq f(x_0)$.

On dit que f a un minimum local strict en x_0 s'il existe $r > 0$ tel que pour tout $x \in B(x_0, r)$, $x \neq x_0$, on a $f(x) > f(x_0)$.

On dit que f a un maximum local strict en x_0 s'il existe $r > 0$ tel que pour tout $x \in B(x_0, r)$, $x \neq x_0$, on a $f(x) < f(x_0)$.

Définitions

Soient D une partie de \mathbb{R}^d , f une fonction définie sur D à valeurs dans \mathbb{R} .

Si f est majorée, on appelle borne supérieure de f sur D le plus petit des majorants de f , soit le nombre réel noté $\sup_D f$ ou $\sup_{x \in D} f(x)$ défini par :

$$\forall x \in D \quad f(x) \leq \sup_D f, \quad \forall M < \sup_D f, \quad \exists x \in D, \quad f(x) > M.$$

Si f est minorée, on appelle borne inférieure de f sur D le plus petit des minorants de f , soit le nombre réel noté $\inf_D f$ ou $\inf_{x \in D} f(x)$ défini par :

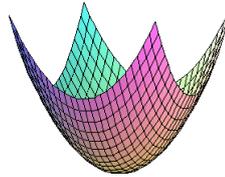
$$\forall x \in D \quad f(x) \geq \inf_D f, \quad \forall m > \inf_D f, \quad \exists x \in D, \quad f(x) < m.$$

Si f est majorée sur D , on dit que f atteint sa borne supérieure s'il existe $x \in D$ tel que $f(x) = \sup_D f$.

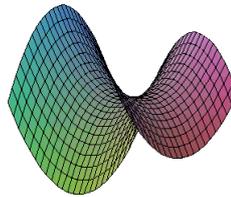
Si f est minorée sur D , on dit que f atteint sa borne inférieure s'il existe $x \in D$ tel que $f(x) = \inf_D f$.

3.2 Représentation géométrique

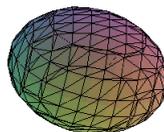
- Cas d'une fonction de deux variables $f(x, y)$
 - a) On considère le graphe $G(f) = \{(x, y), f(x, y)\} / (x, y) \in \mathbb{C}\} \subseteq \mathbb{R}^3$.
Exemples : $f(x, y) = x^2 + y^2$.
Le graphe est un parabolôïde de révolution.



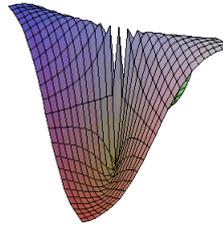
Le graphe $f(x, y) = y^2 - x^2$ est un parabolôïde hyperbolique :



- b) On considère les courbes de niveau $\{(x, y) \in D / f(x, y) = C\}$.
Dans les exemples précédents, les courbes de niveau sont des cercles $\{(x, y) / x^2 + y^2 = C \geq 0\}$ et des hyperboles $\{(x, y) / x^2 - y^2 = C \geq 0\}$.
- Cas d'une fonction de plus de deux variables
Le graphe étant dans \mathbb{R}^4 , on ne peut le dessiner.
Si $n = 3$, on utilise les surfaces de niveau $\{(x, y, z) \in D / f(x, y, z) = C\}$.
Par exemple la surface de niveau 5 de la fonction $f(x, y, z) = 2x^2 + 3y^2 + z^2$ ressemble à peu près à



Remarque : il arrive qu'un logiciel ne donne pas de représentation fidèle de ce qui se passe au voisinage d'un point. Par exemple il est difficile de se rendre compte de ce qui se passe au voisinage de $(0, 0)$ pour la fonction définie par $f(x, y) = \frac{xy}{x^2+y^2}$ en regardant l'image ;



3.3 Fonctions continues de \mathbb{R}^d dans \mathbb{R}

Définition 3.6. Une fonction $f : D \rightarrow \mathbb{R}$ (où $D \subset \mathbb{R}^d$) a pour limite b en x_0 si $x_0 \in \overline{D}$ et si $\forall \varepsilon > 0, \exists \delta > 0$ tel que :
 $x \in D, \|x - x_0\| < \delta \Rightarrow |f(x) - b| < \varepsilon$.

Notation

Dans ce cas $b = \lim_{x \rightarrow x_0} f(x)$.

Définition 3.7. (i) $f : D \rightarrow \mathbb{R}$ est **continue** en $x_0 \in D$ si et seulement si $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

(ii) f est **continue sur** D si et seulement si elle est continue en tout point de D .

Théorème 3.8. Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$. Les propositions suivantes sont équivalentes :

- (i) f est continue sur \mathbb{R}^d .
- (ii) $\forall b \in \mathbb{R}, \forall x_p$ avec $x_p \rightarrow b$ on a : $f(x_p) \rightarrow f(b)$ dans \mathbb{R} .
- (iii) $\forall F$ fermé de $\mathbb{R}, f^{-1}(F) = \{x \in \mathbb{R}^d / f(x) \in F\}$ est un fermé de \mathbb{R}^d .
- (iv) $\forall \theta$ ouvert de $\mathbb{R}, f^{-1}(\theta) = \{x \in \mathbb{R}^d / f(x) \in \theta\}$ est un ouvert de \mathbb{R}^d .

Remarque

Si $D \neq \mathbb{R}^d$, il faut modifier les points (iii) et (iv), et dire que $f^{-1}(\theta)$ est un ouvert de D et $f^{-1}(F)$ est un fermé de D .

Démonstration

Il suffit de démontrer les implications successives (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i).

(i) \Rightarrow (ii) : On se donne une fonction f continue et une suite (x_k) convergeant vers b et il s'agit de montrer que $(f(x_k))$ converge vers $f(b)$. Écrivons les définitions de l'hypothèse et de ce que nous cherchons à démontrer.

Hypothèses :

- Convergence de la suite $(x_k) : \forall \epsilon > 0 \exists K \forall k > K d(x_k, b) < \epsilon$.
- Continuité de la fonction (en b) : $\forall \alpha > 0 \exists \delta > 0 (d(x, b) < \delta \Rightarrow |f(x) - f(b)| < \alpha)$.

À montrer :

- Convergence de la suite $(f(x_k)) : \forall \epsilon > 0 \exists K \forall k > K d(f(x_k), f(b)) < \epsilon$.

Soit $\epsilon > 0$. Prenons $\delta > 0$ pour que si $d(x, b) < \delta$ alors $|f(x) - f(b)| < \epsilon$ (c'est possible grce à la deuxième hypothèse). Prenons ensuite K tel que si $k > K$ alors $d(x_k, b) < \delta$ (c'est possible grce à la première hypothèse). Alors si $k > K$ on a $d(x_k, b) < \delta$, donc $|f(x_k) - f(b)| < \epsilon$.

(ii) \Rightarrow (iii) : On suppose que (ii) est vraie et on cherche à montrer qu'alors (iv) est vraie. Soit F un fermé. On veut montrer que $f^{-1}(F)$ est fermé. Autrement dit on veut montrer que toute suite convergente d'éléments de $f^{-1}(F)$ a sa limite dans $f^{-1}(F)$. Soit (x_k) une telle suite, x sa limite. Comme (ii) est vraie, $(f(x_k))_k$ converge vers $f(x)$. D'autre part, pour tout k , x_k appartient à $f^{-1}(F)$. Autrement dit, pour tout k , $f(x_k)$ appartient à F . Mais F est fermé donc toute suite convergente d'éléments de F a sa limite dans F . Or $(f(x_k))_k$ converge vers $f(x)$. On a donc $f(x) \in F$ soit encore $x \in f^{-1}(F)$, ce qu'on voulait montrer.

(iii) \Rightarrow (iv) : Commençons par remarquer que pour toute partie E de \mathbb{R} , l'égalité

$$f^{-1}({}^c E) = {}^c f^{-1}(E),$$

est valable.

Supposons $f^{-1}(F)$ est fermé quand F est ouvert et donnons-nous un ouvert θ . Alors ${}^c \theta$ est fermé, donc $f^{-1}({}^c \theta)$ est fermé. Mais comme $f^{-1}({}^c \theta) = {}^c f^{-1}(\theta)$ cela signifie que ${}^c f^{-1}(\theta)$ est fermé, autrement dit que $f^{-1}(\theta)$ est ouvert.

(iv) \Rightarrow (i) : On suppose que (iii) est vraie et on cherche à montrer qu'alors (i) est vraie.

À montrer : $\forall b \in \mathbb{R}^d, \forall \alpha > 0 \exists \delta > 0 (d(x, b) < \delta \Rightarrow |f(x) - f(b)| < \alpha)$.

Soient $b \in \mathbb{R}^d, \alpha > 0$. L'intervalle $]f(b) - \alpha, f(b) + \alpha[$ est un ouvert de \mathbb{R} . Par (iii) $f^{-1}(]f(b) - \alpha, f(b) + \alpha[)$ est un ensemble ouvert (auquel b appartient évidemment). Dire que cet ensemble est ouvert, c'est en particulier, dire qu'il existe $\delta > 0$ tel que, $B(b, \delta) \subset f^{-1}(]f(b) - \alpha, f(b) + \alpha[)$. Mais cette inclusion signifie que si $d(x, b) < \delta$ alors $|f(x) - f(b)| < \alpha$. C'est ce que nous voulions montrer. \square

Théorème 3.9. *Soit D un sous-ensemble de \mathbb{R}^d et f et g des fonctions continues sur D . Alors $f + g$ et fg sont continues sur D . De plus $\frac{f}{g}$ est continue en tout point où g ne s'annule pas.*

Démonstration Le plus simple est peut-être d'utiliser la caractérisation séquentielle de la continuité : nous avons à montrer que, pour tout point x en lequel f et g sont

définies, pour toute suite $(y_k)_k$ convergeant vers x , les suites $(f(y_k)+g(y_k))_k$, $(f(y_k)g(y_k))_k$, $(f(y_k)/g(y_k))_k$ convergent respectivement vers $f(x) + g(x)$, $f(x)g(x)$, $f(x)/g(x)$. Par hypothèse f et g sont continues donc $(f(y_k))_k$ et $(g(y_k))_k$ tendent vers $f(x)$ et $g(x)$. Or, on sait que si deux suites de nombres réels $(u_k)_k$, $(v_k)_k$ tendent vers l et l' alors les suites $(u_k + v_k)_k$, $(u_k.v_k)_k$, $(u_k/v_k)_k$ convergent respectivement vers $l + l'$, ll' , l/l' . Il suffit donc d'appliquer ce résultat pour $u_k = f(y_k)$ et $v_k = g(y_k)$.

Il y a quand même un petit problème avec le quotient quand $g(x)$ vaut 0. Ce n'est pas parce que $g(x)$ vaut 0 que le quotient n'a pas de limite en x ; cela peut arriver mais dépend des fonctions considérées. Ce qui est sûr en revanche c'est que si $g(x)$ n'est pas nul, alors le quotient f/g est continu en x . Il faut être plus précis dans ce cas. Si $g(x)$ n'est pas nul, alors comme $(g(y_k))_k$ tend vers $g(x)$, à partir d'un certain rang $g(y_k)$ est différent de 0, le quotient $(f(y_k)/g(y_k))_k$ est alors défini à partir d'un certain rang et tend vers $f(x)/g(x)$. \square

Théorème 3.10. Soit D un sous-ensemble de \mathbb{R}^d et E un sous-ensemble de \mathbb{R}^n . Si $f : D \rightarrow \mathbb{R}$ et $g : E \rightarrow \mathbb{R}^d$ sont continues, avec l'image de E par g incluse dans D , alors $f \circ g$ est continue sur E .

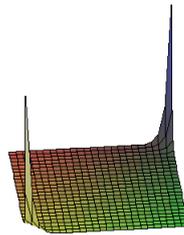
Exemple d'application de ce résultat Comme $|x - x'| \leq \|(x, y) - (x', y')\|$ et $|y - y'| \leq \|(x, y) - (x', y')\|$, les applications définies par $(x, y) \mapsto x$, $(x, y) \mapsto y$ sont continues sur \mathbb{R}^2 .

D'après le théorème précédent les applications définies par $(x, y) \mapsto x + y$, $(x, y) \mapsto xy$, puis $(x, y) \mapsto x^2 + 3xy$ et toutes les fonctions polynôme en deux variables x et y sont continues sur \mathbb{R}^2 .

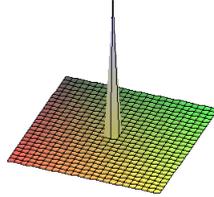
De la même façon toutes les fractions rationnelles en deux variables sont continues là où elles sont définies.

Autres exemples

$f(x, y) = e^{xy}$ continue sur \mathbb{R}^2



$$f(x, y) = \frac{1}{\sqrt{x^2 + y^2}} \text{ continue sur } \mathbb{R}^2 \setminus \{0\}$$



Remarque : vous voyez qu'il n'est pas évident à la vue des deux images précédentes de deviner que la première fonction est continue sur \mathbb{R}^2 et l'autre non.

3.4 Etude de certaines surfaces quadratiques

On cherche à étudier les polynômes quadratiques de la forme $z = Lx^2 + 2Mxy + Ny^2$.

Pour ce faire nous allons utiliser l'identité remarquable vue au collège : $(x + \alpha)^2 = x^2 + 2\alpha x + \alpha^2$ en écrivant quand il le faudra $x^2 + 2\alpha x = (x + \alpha)^2 - \alpha^2$. Si L n'est pas nul, on peut écrire

$$\begin{aligned} Lx^2 + 2Mxy + Ny^2 &= L(x^2 + 2Mxy/L + Ny^2/L) \\ &= L((x + My/L)^2 - M^2y^2/L^2 + Ny^2/L) \\ &= L((x + My/L)^2 + (NL - M^2)y^2/L^2). \end{aligned}$$

Ceci permet de voir que si $NL - M^2$ est strictement positif alors dans la parenthèse on trouve la somme de deux carrés et le graphe est un paraboloides elliptique ("tourné" vers le haut si $L > 0$, vers le bas si $L < 0$). Si $NL - M^2$ est strictement négatif alors dans la parenthèse on trouve la différence de deux carrés et le graphe est un paraboloides hyperbolique.

Si N n'est pas nul, en procédant de la même façon, on peut écrire

$$\begin{aligned} Lx^2 + 2Mxy + Ny^2 &= N(Lx^2/N + 2Mxy/N + y^2) \\ &= N(Lx^2/N + (y + Mx/N)^2 - M^2x^2/N^2) \\ &= N((LN - M^2)x^2/N^2 + (y + Mx/N)^2). \end{aligned}$$

On obtient que si $LN - M^2$ est strictement positif alors le graphe est un paraboloides elliptique ("tourné" vers le haut si $N > 0$, vers le bas si $N < 0$). Si $LN - M^2$ est strictement négatif alors dans la parenthèse on trouve la différence de deux carrés et le graphe est un paraboloides hyperbolique.

Dans le cas où L et N ne sont pas nuls tous les deux, ces deux façons de faire donne bien les mêmes résultats. En effet, si $LN - M^2$ est strictement positif alors en particulier $LN > 0$ autrement dit L et N ont le même signe (l'orientation du paraboloides elliptique

est bien déterminée de la même façon).

Si L et N sont tous les deux nuls, et M est différent de 0, alors le graphe est un parabolôïde hyperbolique. Pour le voir il suffit d'écrire (encore une identité remarquable!) :

$$2Mxy = M(x+y)^2/2 - M(x-y)^2/2.$$

Quelque soit le signe de M nous avons une différence de deux carrés.

Enfin reste le cas où $LN - M^2$ est nul. On a alors

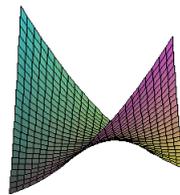
$$Lx^2 + 2Mxy + Ny^2 = L(x + My/L)^2$$

(quand $L \neq 0$ par exemple), et le graphe est un cylindre parabolique.

Exemples :

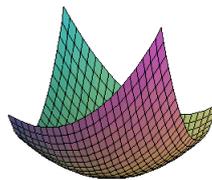
- $2x^2 - 6xy + y^2$

On calcule $2 \cdot 1 - 3^2 = -7 < 0$. Le graphe est un parabolôïde hyperbolique :



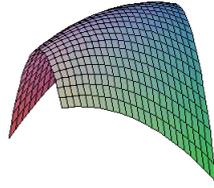
- $2x^2 - 2xy + 3y^2$

On calcule $2 \cdot 3 - 2^2 = 2 > 0$ et $2 > 0$. Le graphe est un parabolôïde elliptique tourné vers le haut :



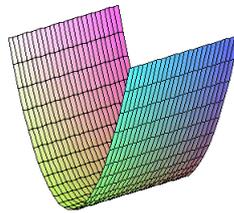
- $-x^2 + 2xy - 3y^2$

On calcule $(-1) \cdot (-3) - 1^2 = 1 > 0$ et $-1 < 0$. Le graphe est un parabolôïde elliptique tourné vers le bas :



- $x^2 + 2xy + y^2$

On calcule $1.1 - 1^2 = 0$. Le graphe est un cylindre parabolique :



Remarque : comme vous le voyez, il n'est pas toujours facile de reconnaître un parabolôïde hyperbolique ou un parabolôïde elliptique sur l'image donnée par l'ordinateur.

Proposition 3.11. *Il existe des coordonnées orthogonales X, Y, Z dans lesquelles $Z = k_1 X^2 + k_2 Y^2$.*

3.5 Fonctions continues sur un ensemble compact

Théorème 3.12. *Soit $f : X \rightarrow \mathbb{R}$ (avec X compact) continue. Alors :*

(i) *f est bornée sur X .*

(ii) *f atteint ses bornes inférieure et supérieure.*

Démonstration Utilisons encore les suites. Supposons que f ne soit pas bornée. Alors on peut trouver une suite (x_k) telle que $(f(x_k))$ tende vers $+\infty$ ou $-\infty$. Considérons par exemple le cas où $(f(x_k))$ tend vers $+\infty$. Comme X est compact, il existe une sous-suite $(x_{k_l})_l$ convergeant vers un point x de X . On a alors $\lim_l f(x_{k_l}) = +\infty$ et $\lim_l f(x_{k_l}) = f(x)$ (par continuité de f). Contradiction. L'hypothèse faite est absurde : f est bornée.

Reste à montrer qu'elle atteint ses bornes. On procède de manière analogue. Montrons le pour la borne inférieure.

Il existe une suite (x_k) telle que $(f(x_k))$ tende vers $\inf_X f$. X est compact, il existe une sous-suite $(x_{k_l})_l$ convergeant vers un point x de X . On a alors $\lim_l f(x_{k_l}) = \inf_X f$ et

$\lim_l f(x_{k_l}) = f(x)$ (par continuité de f). Conclusion : $f(x) = \inf_X f$, autrement dit $\inf_X f$ est atteinte. \square

Définition 3.13. $f : \mathbb{R}^d \rightarrow \mathbb{R}$ est uniformément continue si $\forall \varepsilon > 0, \exists \delta > 0$ tel que $\|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Théorème 3.14. Soit f continue de X dans \mathbb{R} avec X compact. Alors f est uniformément continue sur X .

Démonstration Nous allons raisonner par l'absurde. Considérons une fonction f continue sur un ensemble compact X . Supposons que f ne soit pas uniformément continue. Alors la phrase

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tel que } \|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

est fausse, c'est-à-dire que

$$\exists \varepsilon > 0, \forall \delta > 0 \exists x, y \text{ avec } \|x - y\| < \delta \text{ et } |f(x) - f(y)| \geq \varepsilon.$$

En particulier pour tout $n \in \mathbb{N}$, on peut trouver x_n et y_n deux points de X distants de moins de $1/n$ tels que $|f(x_n) - f(y_n)| \geq \varepsilon$. On obtient ainsi deux suites d'éléments de X . Comme X est compact, on peut extraire de la suite (x_n) une sous-suite convergente. Soient $(x_{n_k})_k$ une telle suite et x_* sa limite. Alors $(y_{n_k})_k$ converge aussi vers x_* car $\|x_{n_k} - y_{n_k}\| < 1/n_k$. Comme f est continue en x_* on a $\lim_k f(x_{n_k}) = \lim_k f(y_{n_k}) = f(x_*)$. Mais par construction on a aussi, pour tout k , $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon$. Contradiction. \square

4 Courbes paramétrées

4.1 Introduction : courbes dans le plan

4.1.1 Définition, exemples

On appelle courbe dans le plan une application de \mathbb{R} dans \mathbb{R}^2 ou d'un intervalle de \mathbb{R} dans \mathbb{R}^2 . Une telle courbe est définie par ses applications coordonnées dans le repère $((1, 0), (0, 1))$:

$$f : \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto (x(t), y(t)).$$

Si on change de repère l'écriture de la courbe change. Nous ne voulons pas étudier en détail la notion de courbes et leurs différentes représentations. Nous supposons donc qu'une courbe est donnée par ses coordonnées dans la base précédente.

Remarque : on appelle généralement courbe l'image de l'application plutôt que l'application elle-même.

4.1.2 Longueur de courbes

Supposons que les deux applications x et y soit C^1 et définie sur $[0, 1]$. On appelle longueur de la courbe définie par x et y la quantité

$$l(\mathcal{C}) = \int_0^1 \sqrt{x'(t)^2 + y'(t)^2} dt.$$

Lorsque les fonctions ne sont pas dérivables il n'est pas toujours possible de définir la longueur de la courbe. Elle peut très bien être infinie.

4.1.3 La courbe de Peano, la courbe de Koch

Peano a ainsi montré que l'on peut très bien voir le carré unité comme une courbe. Il existe une application continue surjective de $[0, 1]$ dans $[0, 1]^2$. Mais il existe aussi des courbes de longueur infinie mais d'aire nulle. Les plus célèbres de ces courbes présentent des propriétés d'autosimilarité. Ce sont ce qu'on appelle des courbes fractales. Ce type de courbe a été longtemps considéré comme des objets étranges et peu naturels. Ce n'est plus le cas aujourd'hui. Les trajectoires de ce qu'on appelle le mouvement brownien, par exemple, sont des courbes sans tangente, de longueur infinie.

Les objets fractales font l'objet de nombreuses recherches.

Remarquons que si la courbe est C^1 alors elle est négligeable dans le plan. Autrement dit son aire est nulle ou encore pour tout $\epsilon > 0$ on peut la recouvrir par une réunion finie de

petits disques dont la somme des aires est inférieure à ϵ . Montrons-le. Donnons-nous une fonction C^1

$$f : [0, 1] \rightarrow \mathbb{R}^2, t \mapsto (x(t), y(t)).$$

Comme les dérivées x' et y' sont continues sur le segment $[0, 1]$ elles sont toutes les deux bornées. Soit M tel que $|x'(t)|$ et $|y'(t)|$ soient toutes deux inférieures à M pour tout t dans $[0, 1]$. Le théorème des accroissements finis assure alors que pour tous $t, t' \in [0, 1]$ on a

$$|x(t) - x(t')| \leq M|t - t'|, \quad |y(t) - y(t')| \leq M|t - t'|.$$

La distance entre l'image de t et celle de t' est donc inférieure à $\sqrt{2}M|t - t'|$.

Soit $\epsilon > 0$.

Prenons $k = E(\sqrt{2}M)/\epsilon + 1$ points sur $[0, 1]$ espacés de moins de $\epsilon/\sqrt{2}M : t_0, \dots, t_k$. Alors pour tout t dans $[0, 1]$ le point $f(t)$ appartient à la réunion des disques de rayons $M\epsilon/\sqrt{2}M = \epsilon/\sqrt{2}$ centrés en les points $f(t_i)$. Cette réunion a une aire inférieure à $k\pi\epsilon^2/2 \leq (\sqrt{2}M/\epsilon + 1)\pi\epsilon^2/2 \leq \pi M\epsilon$.

L'exemple de la courbe de Peano montre que si f est supposée seulement continue ce résultat n'est plus vrai.

4.2 Courbes paramétrées

4.2.1 Définitions

Définition 4.1. Une fonction $F : \mathbb{R} \rightarrow \mathbb{R}^d$ s'appelle fonction vectorielle ou courbe paramétrée.

On exprime $F(t)$ à l'aide des fonctions coordonnées $F(t) = (f_1(t), \dots, f_d(t))$.

Exemples

- (1) Soit F la fonction définie par $F(t) = (x_0 + ta, y_0 + tb, z_0 + tc)$ avec $t \in \mathbb{R}$. La droite passant par (x_0, y_0, z_0) et dirigée par (a, b, c) .
- (2) Soit F la fonction définie par $F(t) = (R \cos t, R \sin t)$ avec $t \in [0, 2\pi]$. Un cercle.
- (3) Soit F la fonction définie par $F(t) = (\cos t, \sin t, t)$ avec $t \in \mathbb{R}$. Une hélice.

Définition 4.2. Soit $F : \mathbb{R} \rightarrow \mathbb{R}^d$.

- (1) $\lim_{t \rightarrow p} F(t) = \left(\lim_{t \rightarrow p} f_1(t), \dots, \lim_{t \rightarrow p} f_d(t) \right)$
- (2) $F'(t) = (f_1'(t), \dots, f_d'(t))$
- (3) $\int_a^b F(t) dt = \left(\int_a^b f_1(t) dt, \dots, \int_a^b f_d(t) dt \right)$

Théorème 4.3. Si $F, G : \mathbb{R} \rightarrow \mathbb{R}^d$ et $u : \mathbb{R} \rightarrow \mathbb{R}$ sont dérivables alors :

- (i) $(F + G)'(t) = F'(t) + G'(t)$
- (ii) $(uF)'(t) = u'(t)F(t) + u(t)F'(t)$
- (iii) $\langle F, G \rangle'(t) = \langle F'(t), G(t) \rangle + \langle F(t), G'(t) \rangle$
- (iv) $(F \wedge G)'(t) = F'(t) \wedge G(t) + F(t) \wedge G'(t)$ si $d = 3$
- (v) $F(u(t))' = u'(t)F'(u(t))$

Corollaire 4.4. Si une courbe paramétrée $F(t)$ est dérivable et si $\|F(t)\|$ est constante alors $\langle F(t), F'(t) \rangle = 0$. (Autrement dit, si la courbe $F(t)$ est sur une sphère centrée en 0, alors $F(t)$ et $F'(t)$ sont orthogonaux).

Exemple

$$F(t) = (\cos t, \sin t)$$

Si $F(t) = (x_1(t), \dots, x_n(t))$ est dérivable et $F'(t) \neq 0$, on voit que $F'(t) = \lim_{h \rightarrow 0} \frac{F(t+h) - F(t)}{h}$.

Interprétation graphique du vecteur dérivée.

Définition 4.5. Soit C la courbe tracée par F . Si $F'(t_0) \neq 0$ alors la droite passant par $F(t_0)$ de vecteur directeur $F'(t_0)$ est appelée **droite tangente** à C en $F(t_0)$.

$F'(t_0)$ est un vecteur tangent à C en $F(t_0)$. Le point $F(t_0)$ est dit régulier.

Lorsque $F'(t_0) = 0$, le point est dit stationnaire et il nécessite une étude plus poussée (via la formule de Taylor).

Nous voulons maintenant définir la longueur d'un arc de courbe régulière.

Le cas d'un segment.

Le cas du cercle.

Le cas des courbes planes. Soit F une fonction C^1 définie sur $[a, b]$ à valeurs dans \mathbb{R}^2 . On cherche à approcher l'arc par une ligne brisée comportant de plus en plus de segments. Soit n un entier naturel. Pour i variant de 0 à n posons : $t_i = a + i(b-a)/n$, $x_i = F(t_i)$.

Calculons la longueur de la courbe brisée dont les sommets sont les points x_i . Grâce à la définition de la dérivée on a :

$$\begin{aligned} x_{i+1} - x_i &= F(t_i + (b-a)/n) - F(t_i) \\ &= F'(t_i)(b-a)/n + \epsilon(1/n)/n \end{aligned}$$

où $\epsilon(1/n)$ tend vers 0 quand n tend vers l'infini. En utilisant l'inégalité triangulaire, on obtient que la longueur du segment $[x_i, x_{i+1}]$ est donc très proche de $\|F'(t_i)(b-a)/n\|$:

$$\left| \|x_{i+1} - x_i\| - \|F'(t_i)(b-a)/n\| \right| \leq \|\epsilon(1/n)\|/n.$$

On obtient donc une estimation de la longueur de la courbe brisée :

$$\begin{aligned} \sum_{i=0}^{n-1} \|x_{i+1} - x_i\| - \sum_{i=0}^{n-1} \|F'(t_i)(b-a)/n\| &\leq \sum_{i=0}^{n-1} \left| \|x_{i+1} - x_i\| - \|F'(t_i)(b-a)/n\| \right| \\ &\leq \sum_{i=0}^{n-1} \|\epsilon(1/n)\|/n \\ &\leq \|\epsilon(1/n)\|. \end{aligned}$$

La différence entre la longueur de la courbe brisée et la somme $\sum_{i=0}^{n-1} \|F'(t_i)(b-a)/n\|$ tend donc vers 0 quand n tend vers l'infini. Mais lorsque n tend vers l'infini la somme $\sum_{i=0}^{n-1} \|F'(t_i)(b-a)/n\|$ converge vers l'intégrale $\int_a^b \|F'(t)\| dt$. C'est donc cette quantité qu'on appelle longueur de l'arc de courbe défini par F .

Le cas général.

Définition 4.6. La longueur de l'arc de la courbe $F(t)$ entre $t = a$ et $t = b$ est donnée par $\int_a^b \|F'(t)\| dt$.

Exemple

La longueur du graphe d'une fonction f de classe C^1 définie sur un intervalle $[a, b]$ est donnée par

$$l = \int_a^b \sqrt{1 + (f'(x))^2} dx$$

Soit $F : \mathbb{R} \rightarrow \mathbb{R}^d$ une courbe paramétrée.

On suppose F de classe C^1 et que $F'(t) \neq 0$ pour tout t . On dit que F est **régulière**.

Alors :

(i) $l(t) = \int_{t_0}^t \|F'(u)\| du$ est la longueur de l'arc entre t_0 et t .

(ii) $\frac{dl}{dt} = \|F'(t)\| > 0$

Ainsi, la fonction l est de classe C^1 et de dérivée positive strictement : elle admet une fonction réciproque $s \rightarrow t(s)$ dont la dérivée est donnée par $t'(s) = \frac{1}{\|F'(t(s))\|}$.

On note $G(s)$ la fonction $G(s) = F(t(s))$. Elle définit la même courbe, mais avec un paramétrage différent. On l'appelle **paramétrisation unitaire** de F car on a $\|G'(s)\| = 1$.

Définition 4.7. La courbure de $G(s)$ est donnée par $\rho(s) = \|G''(s)\|$.

Proposition 4.8. Si F n'est pas une paramétrisation unitaire alors la courbure est donnée par $\rho(t) = \frac{\|F'(t) \wedge F''(t)\|}{\|F'(t)\|^3}$.

4.2.2 Quelques exemples

Étude détaillée de la courbe de Lissajous définie par $t \mapsto (\cos(3t), \sin(2t))$ avec réduction de l'intervalle à $[0, \pi/2]$ via la périodicité et les symétries, étude des variations, tangentes et tracer de la courbe.

Nous n'avons malheureusement pas plus de temps à consacrer à l'étude des courbes paramétrées. Pour voir d'autres exemples :

<http://fr.wikipedia.org/wiki/Clothode>

<http://fr.wikipedia.org/wiki/Cyclode>

<http://www.mathcurve.com>

5 Dérivées des fonctions de plusieurs variables

Les fonctions de plusieurs variables sont des fonctions de chacune de leurs variables. Si elles sont dérivables par rapport à chaque variables comme fonctions d'une variable alors elles admettent des dérivées partielles. Le calcul des dérivées partielles se fait donc comme le calcul des dérivées des fonctions réelles de la variable réelles (les autres variables sont considérées comme des constantes). Mais en dimension supérieure, dire qu'une fonction est dérivable, n'est pas seulement dire qu'elle a des dérivées partielles. On dit qu'une fonction f est dérivable ou différentiable en un point si elle a une bonne approximation linéaire (ou affine) en ce point. Lorsque f et g sont dérivables alors les propriétés habituelles sont vérifiées. Mais $f'(x)$ n'est pas un nombre mais une matrice. Dans la formule de la dérivation des fonctions composées par exemple l'ordre a alors une grande importance.

Remarque : on écrit la plupart du temps on écrit les variables d'une fonction à plusieurs variables en ligne mais dans l'écriture du développement de Taylor on considère des vecteurs colonnes.

5.1 Dérivées partielles des fonctions à valeurs réelles

5.1.1 Rappels

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$. La dérivée de f en x , si elle existe, est : $f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$.

Exemple La fonction définie sur \mathbb{R} par $f(x) = x^2$ est dérivable de dérivée $f'(x) = 2x$. En effet la limite quand h tend vers 0 de

$$\frac{(x+h)^2 - x^2}{h} = 2x + h$$

existe et vaut $2x$.

5.1.2 Dérivée partielle

Définition 5.1. Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$. On dit que f admet une dérivée partielle par rapport à la variable x_i au point $a = (a_1, \dots, a_d)$ si la fonction d'une variable

$$x_i \mapsto f(a_1, \dots, x_i, a_{i+1}, \dots, a_d)$$

est dérivable au point a_i . Dit autrement, on définit la dérivée partielle de f par rapport à x_i au point $a = (a_1, \dots, a_d)$ par

$$\lim_{h \rightarrow 0} \frac{f(a_1, \dots, x_i + h, a_{i+1}, \dots, a_d) - f(a)}{h}$$

si cette limite existe.

Notation

Cela se note $\frac{\partial f}{\partial x_i}(x_1, \dots, x_n)$, $f_{x_i}(x_1, \dots, x_n)$, $D_i f(x_1, \dots, x_n)$.

Dans le cas de deux variables on a :

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}$$

$$\frac{\partial f}{\partial y}(x, y) = \lim_{k \rightarrow 0} \frac{f(x, y+k) - f(x, y)}{k}$$

Exemple

$$(1) f(x, y) = e^{xy^2}$$

$$(2) f(x, y) = x^2 + 3y^2 - 2xy$$

$$(3) f(x, y) = \sqrt{1 - x^2 - y^2}$$

$$(4) f(x, y, z) = xy^2 + z$$

5.1.3 Interprétation géométrique

$\frac{\partial f}{\partial x}(x_0, y_0)$ est la pente de la tangente à la courbe $z = f(x, y_0)$ en (x_0, y_0) .

5.1.4 Gradient

Définition 5.2. Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction admettant des dérivées partielles.

Son gradient en $a \in \mathbb{R}^d$, noté $\nabla f(a)$ est le vecteur $\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_d}(a) \right)$.

Exemple

$$(1) f(x, y) = x^2 y^3$$

$$(2) f(x, y, z) = x^2 \sin(yz)$$

Remarque

Le gradient peut être considéré comme un vecteur de \mathbb{R}^d mais aussi comme une matrice $1 \times d$.

Théorème 5.3. Si f et g sont deux fonctions de \mathbb{R}^d dans \mathbb{R} avec des gradients, alors :

$$(i) \nabla(f + g) = \nabla f + \nabla g$$

$$(ii) \nabla(cf) = c \nabla f \text{ où } c \in \mathbb{R}$$

5.1.5 Dérivées partielles et continuité

Une fonction peut avoir des dérivées partielles sans être continue!!

Exemples

(1) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = \begin{cases} 0 & \text{si } 0 < y < x^2 \\ 1 & \text{sinon} \end{cases}$$

a des dérivées partielles en $(0, 0)$ mais n'y est pas continue.

(2) $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{en } (0, 0) \end{cases}$

admet des dérivées partielles en tout point mais n'est pas continue en $(0, 0)$.

Pour dépasser cette difficulté, on définit la différentielle (ou dérivée totale) ou on considère les fonctions ayant des dérivées partielles continues encore appelées fonctions de classe \mathcal{C}^1 .

5.1.6 Dérivation composée

Théorème 5.4. Si g est définie sur \mathbb{R} par $g(t) = f(r(t))$ où f est \mathcal{C}^1 de \mathbb{R}^d dans \mathbb{R} et r dérivable de \mathbb{R} dans \mathbb{R}^d , alors g est dérivable et on a $g'(t) = \langle \nabla f(r(t)), r'(t) \rangle$.

Exemple

$$f(x, y) = xy^2$$

$$r(t) = (t, t^2)$$

5.1.7 Accroissements finis

On a un théorème des accroissements finis de manière similaire au cas d'une variable réelle.

Théorème 5.5. Soient $f : \mathbb{R}^d \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , $X = (x_1, \dots, x_n)$, $H = (h_1, \dots, h_n)$. Alors il existe $\theta \in]0, 1[$ tel que $f(X + H) - f(X) = \langle \nabla f(X + \theta H), H \rangle$.

5.1.8 Dérivée selon un vecteur

Définition 5.6. Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 et $V \in \mathbb{R}^n$.

La dérivée selon le vecteur V en X est définie par $D_V f(X) = \lim_{t \rightarrow 0} \frac{f(X + tV) - f(X)}{t}$.

Remarque

Si $V = e_i$, on retrouve $\frac{\partial f}{\partial x_i}$.

Proposition 5.7. On a $D_V f(X) = \nabla f(X) \cdot V$.

Interprétation géométrique du gradient : direction de plus forte pente

La variation de f est la plus forte dans la direction de $\nabla f(X)$. En effet, si g est définie sur \mathbb{R} par $g(t) = f(r(t))$ où f est \mathcal{C}^1 de \mathbb{R}^d dans \mathbb{R} et r dérivable de \mathbb{R} dans \mathbb{R}^d , alors

g représente l'évolution de f le long de la courbe r . Regardons en particulier l'évolution de f au point $x_0 \in \mathbb{R}^d$ le long de la direction $v \in \mathbb{R}^d$, en choisissant pour r la courbe $r(t) = x_0 + tv$. La variation de f dans la direction v au point x_0 est donnée par la dérivée $g'(0) = \langle \nabla f(x_0), v \rangle$. Par conséquent, la direction v dans laquelle f grandit le plus est donnée par les valeurs les plus grandes du produit scalaire $\langle \nabla f(x_0), v \rangle$, c'est-à-dire lorsque le vecteur v est proportionnel au gradient $\nabla f(x_0)$ (d'après l'inégalité de Cauchy-Schwarz).

Comme illustration, l'hiver venu sur les pistes de ski, le skieur voulant aller vite choisit comme direction en un point de la piste celle de l'inverse du gradient (soit la direction de la plus grande pente descendante en un point de la montagne).

Interprétation géométrique du gradient (bis) : espace tangent aux niveaux de $f : \mathbb{R}^d \rightarrow \mathbb{R}$

Notons $N_c = \{x \in \mathbb{R}^d \mid f(x) = c\}$ le niveau c de f . Soit $r : \mathbb{R} \rightarrow \mathbb{R}^d$ une courbe paramétrée dont l'image est incluse dans N_c , et passant par le point $r(0) = x_0$. Alors la fonction $f \circ r$ est constante et égale à c , sa dérivée en $\langle \nabla f(x_0), r'(0) \rangle$ est donc nulle : le gradient de f en x_0 est orthogonal à $r'(0)$, que l'on peut interpréter comme la direction de la tangente à la courbe r . En considérant toutes les courbes possibles passant par x_0 , on en déduit donc une définition de l'espace tangent $T_{x_0}N_c$ au niveau N_c au point x_0 comme étant l'ensemble des directions des tangentes possibles en x_0 , ou plus précisément

$$T_{x_0}N_c = \{x \in \mathbb{R}^d \mid \langle x - x_0, \nabla f(x_0) \rangle = 0\}.$$

Interprétation géométrique du gradient (ter) : espace tangent au graphe de $f : \mathbb{R}^d \rightarrow \mathbb{R}$

En procédant de la même manière avec des courbes $r : \mathbb{R} \rightarrow \mathbb{R}^{d+1}$ dont l'image est contenue dans le graphe G_f de f , on voit que l'espace tangent T_0G_f au graphe de f au point $(x_0, f(x_0)) \in \mathbb{R}^{d+1}$ est l'ensemble

$$T_0G_f = \{(x, z) \in \mathbb{R}^d \times \mathbb{R} \mid z - f(x_0) = \langle \nabla f(x_0), x - x_0 \rangle\}.$$

En effet, une telle courbe $r(t) = (x(t), z(t))$ satisfait $f(x(t)) = z(t)$ puisque qu'elle est incluse dans le graphe, et donc en dérivant on trouve que sa tangente au point $(x(0), z(0)) = (x_0, z_0)$ satisfait

$$z'(0) = \langle \nabla f(x_0), x'(0) \rangle.$$

5.2 La différentielle d'une fonction à valeurs réelles

Cas des fonctions d'une variable

- (i) f est dérivable en X_0 si $\lim_{h \rightarrow 0} \frac{f(X_0 + h) - f(X_0)}{h}$ existe. Sa valeur ℓ est notée $f'(X_0)$.

(ii) On peut, de manière équivalente, écrire $\lim_{h \rightarrow 0} \frac{f(X_0 + h) - f(X_0) - \ell h}{h} = 0$.

On remarque que $h \rightarrow L(h) = \ell h$ est une application linéaire de \mathbb{R} dans \mathbb{R} , que l'on appelle **différentielle** de f en X_0 et que l'on note $df(X_0)$.

(iii) Si f est dérivable en X_0 , alors pour h petit : $f(X_0 + h)$ est voisin de $f(X_0) + f'(X_0)h$.
Donc $h \rightarrow f(X_0) + f'(X_0)h$ est une application affine qui "approche" $f(X_0 + h)$.

Définition 5.8. f est différentiable en X s'il existe une application linéaire $L : \mathbb{R}^d \rightarrow \mathbb{R}$ telle que :

$$\lim_{\|H\| \rightarrow 0} \frac{f(X + H) - f(X) - L(H)}{\|H\|} = 0$$

L'application L est la **différentielle de f en X** et se note $df(X)$.

Remarque

On sait qu'une telle application linéaire L est donnée par $L(H) = \langle H, V \rangle$ où V est un vecteur de \mathbb{R}^d . Voir la proposition ci-dessous.

Remarque

Comme dans le cas $d = 1$ on a $f(X + H)$ "voisin" de $f(X) + df(X)(H)$, on a $f(X + H)$ est "approché" par l'application affine $f(X) + df(X)(H)$.

La différentielle, lorsqu'elle existe, est unique.

Proposition 5.9. Si f est différentiable en X , alors ses dérivées partielles existent et on a :

$$\begin{aligned} df(X)(H) &= \frac{\partial f}{\partial x_1}(X) h_1 + \dots + \frac{\partial f}{\partial x_d}(X) h_d \\ &= \langle \nabla f(X), H \rangle \end{aligned}$$

Démonstration On choisit de faire tendre H vers 0 le long des axes de coordonnées, on retrouve d'une part la définition des dérivées partielles de f , d'autre part les coordonnées du vecteur V qui décrit la différentielle de f . \square

Remarque

La matrice de l'application linéaire $df(X)$ dans la base canonique est le gradient $\nabla f(X)$.

Proposition 5.10. Si f est différentiable en X alors f est continue en X .

Démonstration Notons g la fonction définie par $g(H) = \frac{f(X+H) - f(X) - df(X)(H)}{\|H\|}$. Alors

$$f(X + H) = f(X) + df(X)(H) + \|H\|g(H)$$

et il est clair que $df(X)(H) + \|H\|g(H)$ tend vers 0 quand H tend vers le vecteur nul. Donc la limite de f en X existe et vaut $f(X)$, donc f est continue en X . \square

Remarque

L'existence des dérivées partielles de f n'implique pas la différentiabilité. Par exemple

la fonction $f(x, y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{en } (0, 0) \end{cases}$ admet des dérivées partielles en tout point mais n'est pas différentiable en $(0, 0)$. Mais :

Théorème 5.11. *Si f admet des dérivées partielles et si elles sont continues alors f est différentiable.*

On dit que f est de classe \mathcal{C}^1 .

Démonstration \square

Remarque

A contrario, une fonction peut être différentiable sans que ses dérivées partielles ne soient continues. Prendre par exemple la fonction f définie par $f(x, y) = x^2 \sin \frac{1}{x}$.

Exemples et utilisation

Formes linéaires Fonctions homogènes

5.2.1 Règle de différentiation

Proposition 5.12. *Si f et g sont différentiables on a :*

- (i) $d(f + g)(X) = df(X) + dg(X)$
- (ii) $d(\lambda f)(X) = \lambda df(X)$
- (iii) $d(fg)(X) = f(X) dg(X) + g(X) df(X)$
- (iv) $d\left(\frac{f}{g}\right)(X) = \frac{g(X) df(X) - f(X) dg(X)}{g^2(X)}$

5.2.2 Remarques

Si $f : U \rightarrow \mathbb{R}$ où U est un ouvert de \mathbb{R}^d , alors :

- (i) Si f est \mathcal{C}^1 sur U alors f est différentiable sur U et les dérivées $\frac{\partial f}{\partial x_i}$ existent sur U .

Les réciproques ne sont pas vraies !!

- (ii) Si f est différentiable en $X_0 \in U$ alors l'application affine $A(H) = f(X_0) + df(X_0) \cdot H$ a pour graphe l'espace tangent au graphe de f en X_0 .

Exemple

Droite tangente

Plan tangent

5.2.3 Dérivées partielles successives

Les dérivées partielles $\frac{\partial f}{\partial x_i}(x_1, \dots, x_d)$ sont des fonctions de x_1, \dots, x_d , et il arrive souvent qu'elles soient elles-mêmes dérivables.

Définition 5.13. Si la dérivée partielle de la fonction $\frac{\partial f}{\partial x_j}$ par rapport à la variable x_i existe, on la note $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)$ et on dit qu'il s'agit d'une **dérivée partielle seconde** de f .

Exemple

$f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^3 y^4$. Alors $\frac{\partial^2 f}{\partial x \partial y}(x, y) = 12x^2 y^3 = \frac{\partial^2 f}{\partial y \partial x}(x, y)$.

Théorème 5.14. (Schwarz)

Si toutes les dérivées partielles premières $\frac{\partial f}{\partial x_i}$ et secondes $\frac{\partial^2 f}{\partial x_i \partial x_j}$ existent et sont continues dans une boule autour de $a = (a_1, \dots, a_d)$ alors :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

5.3 La différentielle d'une fonction à valeurs vectorielles

On a déjà rencontré des fonctions à valeurs vectorielles. Quelques exemples :

1. De \mathbb{R}^2 dans \mathbb{R}^2 :

(a) $F(x, y) = (e^x \cos y, e^x \sin y)$

(b) $F(r, \theta) = (r \cos \theta, r \sin \theta)$

2. De \mathbb{R}^3 dans \mathbb{R}^3 : $F(X) = \frac{X}{\|X\|}$.

3. De \mathbb{R}^2 dans \mathbb{R}^3 : $F(x, y) = (x, y, x^2 + y^2)$.

4. De \mathbb{R}^d dans \mathbb{R}^m linéaire : $F(x_1 \dots x_d) = A(x_1 \dots x_d)$ où A est une matrice d colonnes et m lignes.

On peut exprimer F en termes de composantes $F(x_1 \dots x_d) = (f_1(x_1 \dots x_d), \dots, f_m(x_1 \dots x_d))$.

Pour la notion de continuité de les limites, on remplace dans l'espace d'arrivée \mathbb{R}^m les habituelles valeurs absolues par des normes.

Définition 5.15. F de \mathbb{R}^n dans \mathbb{R}^m est **différentiable** en $X \in \mathbb{R}^n$ s'il existe une **application linéaire** L de \mathbb{R}^n dans \mathbb{R}^m telle que :

$$\lim_{\|H\| \rightarrow 0} \frac{F(X+H) - F(X) - L(H)}{\|H\|} = 0.$$

L est la **différentielle** de F en X et se note : $dF(X)$.

Théorème 5.16. $F = (f_1, \dots, f_m)$ est différentiable en X si et seulement si ses fonctions composantes f_1, \dots, f_m sont différentiables et on a :

$$dF(X)(H) = (\langle \nabla f_1(X), H \rangle, \dots, \langle \nabla f_m(X), H \rangle).$$

Définition 5.17. La matrice

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(X) & \cdots & \frac{\partial f_1}{\partial x_n}(X) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(X) & \cdots & \frac{\partial f_m}{\partial x_n}(X) \end{bmatrix}$$

est la matrice de $dF(X)$ et est appelée **matrice jacobienne** de F en X et se note : $J(F)(X)$.

Théorème 5.18. Si F a des composantes de classe \mathcal{C}^1 alors elles sont différentiables et F est également différentiable.

Exemple

- (i) Trouver la matrice jacobienne de F en $(1, 1)$ de : $F(x, y) = (x^2 + y^2, e^{xy})$.
- (ii) Trouver la différentielle de $F(x, y, z) = (x, y, z)$.
- (iii) Trouver la différentielle de $F(r, \theta) = (r \cos \theta, r \sin \theta)$.

5.3.1 Propriétés de la différentielle

Proposition 5.19. Si F de \mathbb{R}^n dans \mathbb{R}^m est linéaire, alors $dF(X) = F$.

Proposition 5.20. Si F est différentiable en X alors F est continue en X .

5.3.2 Différentielles des fonctions composées

Si F est une fonction de \mathbb{R}^n dans \mathbb{R}^m , si G est une fonction de \mathbb{R}^m dans \mathbb{R}^q , alors $G \circ F$ est une fonction de \mathbb{R}^n dans \mathbb{R}^q .

Théorème 5.21. Si F est différentiable en X , et si G est différentiable en $F(X)$, alors $G \circ F$ est différentiable en X et on a :

$$d(G \circ F)(X) = dG(F(X)) \circ dF(X).$$

Exemple

$$F(x, y) = (x^2 + y^2, e^{xy})$$

$$G(u, v) = (xy, \sin x, x^2 y)$$

Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $g : \mathbb{R}^p \rightarrow \mathbb{R}^n$ deux fonctions différentiables. Écrivons $h = f \circ g$.

La fonction $f \circ g$ est une fonction de \mathbb{R}^p dans \mathbb{R} . Sa “dérivée” est donc un vecteur ligne à p colonnes, la transposée de son gradient :

$$\left(\frac{\partial h}{\partial x_1} \quad \frac{\partial h}{\partial x_2} \quad \cdots \quad \frac{\partial h}{\partial x_p} \right).$$

La fonction g est une fonction de \mathbb{R}^p dans \mathbb{R}^n . Sa “dérivée” est la matrice $n \times p$ composée des vecteurs transposés des gradients des coordonnées de g . Si $g(x) = (g_1(x), g_2(x), \dots, g_n(x))$ (on devrait écrire ce vecteur en colonne si on voulait se conformer en toute rigueur aux choix du cours) la dérivée de g s’écrit :

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_p} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_2}{\partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \dots & \frac{\partial g_n}{\partial x_p} \end{pmatrix}.$$

La dérivée de f est donnée par la transposée de son gradient :

$$\left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_n} \right).$$

L’égalité matricielle $h'(x) = (f \circ g)'(x) = f'(g(x)) \cdot g'(x)$ signifie donc :

$$\left(\frac{\partial h}{\partial x_1} \quad \frac{\partial h}{\partial x_2} \quad \dots \quad \frac{\partial h}{\partial x_p} \right) = \left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_n} \right) \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_p} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_2}{\partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \dots & \frac{\partial g_n}{\partial x_p} \end{pmatrix}.$$

Autrement dit pour tout $i = 1, \dots, p$ on a

$$\frac{\partial h}{\partial x_i} = \sum_{k=1}^n \frac{\partial f}{\partial x_k} \frac{\partial g_k}{\partial x_i}.$$

Un exemple

Prenons $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ deux fonctions différentiables définies par

$$f(x, y, z) = 2xy - 3(x + z),$$

$$g(x, y) = (x + y^4, y - 3x^2, 2x^2 - 3y).$$

On demande de calculer les dérivées partielles de la fonction de deux variables $h = f \circ g$.

Pour $\frac{\partial h}{\partial x}$, on obtient :

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x}(g(x, y)) \frac{\partial(x + y^4)}{\partial x} + \frac{\partial f}{\partial y}(g(x, y)) \frac{\partial(y - 3x^2)}{\partial x} + \frac{\partial f}{\partial z}(g(x, y)) \frac{\partial(2x^2 - 3y)}{\partial x},$$

Je vous laisse le calcul de la deuxième dérivée partielle de h en exercice.

6 Sous-ensembles de \mathbb{R}^n et fonctions

On a déjà parlé des courbes paramétrées. On va généraliser cette étude aux nappes paramétrées, puis à l'étude géométrique des niveaux d'une fonction.

6.1 Nappes paramétrées

Si f une fonction de deux variables son graphe est une surface incluse dans \mathbb{R}^3 :

$$\{(x, y, f(x, y)) / (x, y) \in \mathbb{R}^2\}.$$

Une telle surface est un exemple de nappe paramétrée (par f).

Si f est une fonction partout dérivable alors son graphe admet en chaque point un plan tangent. Pour trouver des vecteurs appartenant au plan tangent en $(x_0, y_0, f(x_0, y_0))$ traçons deux courbes sur la surface dans des directions données par les coordonnées :

$$t \mapsto (x_0 + t, y_0, f(x_0 + t, y_0)) \quad t \mapsto (x_0, y_0 + t, f(x_0, y_0 + t))$$

et calculons les coordonnées de leurs vecteurs tangents à l'instant $t = 0$. On obtient, par dérivation composée,

$$\left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right) \quad \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0)\right).$$

Ce sont deux vecteurs indépendants tangents à deux courbes tracées sur la nappe. Le plan tangent à la nappe paramétrée est le plan passant par (x_0, y_0) de direction engendrée par ces deux vecteurs. Pour obtenir une équation de ce plan on peut utiliser le produit vectoriel. Un vecteur normal au plan est donné par

$$\left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right) \wedge \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0)\right) = \left(-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1\right)$$

L'équation du plan tangent est donnée par :

$$-\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + (z - f(x_0, y_0)) = 0$$

Plus généralement une nappe paramétrée est un ensemble décrit par deux paramètres

$$\{(f_1(s, t), f_2(s, t), f_3(s, t)) / (s, t) \in \mathbb{R}^2\}.$$

Les vecteurs

$$\left(\frac{\partial f_1}{\partial s}(s_0, t_0), \frac{\partial f_2}{\partial s}(s_0, t_0), \frac{\partial f_3}{\partial s}(s_0, t_0)\right) \quad \left(\frac{\partial f_1}{\partial t}(s_0, t_0), \frac{\partial f_2}{\partial t}(s_0, t_0), \frac{\partial f_3}{\partial t}(s_0, t_0)\right)$$

sont des vecteurs tangents à la surface au point image de (s_0, t_0) . S'ils sont indépendants le paramétrage définit une surface en (s_0, t_0) et un vecteur normal à la surface est donné par le produit vectoriel des vecteurs précédents

$$\left(\frac{\partial f_2}{\partial s} \frac{\partial f_3}{\partial t} - \frac{\partial f_3}{\partial s} \frac{\partial f_2}{\partial t}, \frac{\partial f_3}{\partial s} \frac{\partial f_1}{\partial t} - \frac{\partial f_1}{\partial s} \frac{\partial f_3}{\partial t}, \frac{\partial f_1}{\partial s} \frac{\partial f_2}{\partial t} - \frac{\partial f_2}{\partial s} \frac{\partial f_1}{\partial t}\right)(s_0, t_0)$$

Si on note (a, b, c) les coordonnées de ce vecteur l'équation du plan tangent à la nappe paramétrée au point image de (s_0, t_0) est

$$a(x - f_1(s_0, t_0)) + b(y - f_2(s_0, t_0)) + c(z - f_3(s_0, t_0)) = 0.$$

(C'est simplement écrire que les vecteurs AM et (a, b, c) sont orthogonaux lorsque M a pour coordonnées (x, y, z) et A est le point image de (s_0, t_0) , $(f_1(s_0, t_0), f_2(s_0, t_0), f_3(s_0, t_0))$).

6.2 Tangentes aux courbes (surfaces, hypersurfaces) de niveau

Donnons une fonction de deux variables f . Que dire des ensembles $\{(x, y) / f(x, y) = c\}$? On les appelle les courbes de niveaux de la fonction f . C'est dire qu'on s'attend à ce que ces ensembles soient des courbes... Ce n'est toutefois pas toujours le cas! Si par exemple f est constante égale à 0, alors les courbes de niveau sont toutes vides sauf la courbe de niveau 0 qui est égale à \mathbb{R}^2 tout entier.

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est différentiable et son gradient n'est pas nul en (x_0, y_0) alors la courbe de niveau $f(x_0, y_0)$ définit bien une courbe au voisinage de (x_0, y_0) . Cette courbe est régulière et on a déjà vu que l'équation de sa tangente est donnée par le gradient :

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0.$$

Par exemple, la tangente en $(1/\sqrt{3}, \sqrt{2/3})$ à la courbe de niveau 1 de la fonction $f(x, y) = x^2 + y^2$ est la droite d'équation :

$$2/\sqrt{3}(x - 1/\sqrt{3}) + 2\sqrt{2/3}(y - \sqrt{2/3}) = 0.$$

On a de la même façon les équations des plans tangents aux surfaces de niveaux de fonctions de trois variables dérivables au voisinage de points en lesquels le gradient n'est pas nul. Par exemple le plan tangent à la surface $xyz = 1$ au point $(1/2, 1, 2)$ a pour équation :

$$2(x - 1/2) + (y - 1) + 1/2(z - 2) = 0.$$

On va proposer par la suite une étude plus précise des courbes de niveaux. Pour ce faire nous disposons du théorème des fonctions implicites, conséquence du théorème d'inversion local. Ce sont deux théorèmes très importants du calcul différentiel. Nous ne donnons pas la démonstration du théorème d'inversion locale. Nous donnons celle du théorème des fonctions implicites seulement en dimension 2.

6.3 Fonctions implicites – Inversion locale

6.3.1 Inversion locale

Soient U un ouvert de \mathbb{R}^n , F une application de U dans \mathbb{R}^n et $V = F(U) \subset \mathbb{R}^n$.

Définition 6.1. F est **inversible** sur U s'il existe une application G de V dans \mathbb{R}^n telle que $G \circ F = \mathbf{1}_U$ et $F \circ G = \mathbf{1}_V$.

Définition 6.2. F de \mathbb{R}^n dans \mathbb{R}^n est **localement inversible** en $X_0 \in \mathbb{R}^n$ s'il existe des ouverts U et V avec $X_0 \in U$ et $F(X_0) \in V$ et $F(U) = V$ tel que F est inversible sur U .

Exemples

- (1) $f : \mathbb{R} \rightarrow \mathbb{R}$ avec $f(x) = x^3$
- (2) $f : \mathbb{R} \rightarrow \mathbb{R}$ avec $f(x) = x^2$
- (3) Si $A \in \mathbb{R}^n$, soit F de \mathbb{R}^n dans \mathbb{R}^n avec $F(X) = X + A$.
- (4) $U = \{(r, \theta) / r > 0, 0 < \theta < \pi\}$
 $F(r, \theta) = (r \cos \theta, r \sin \theta)$

Théorème 6.3. (d'inversion locale)

Soient F définie sur un domaine D de \mathbb{R}^n à valeurs dans \mathbb{R}^n de classe \mathcal{C}^1 et X_0 un point intérieur à D . Alors si $dF(X_0)$ est inversible (en tant qu'application linéaire) F est localement inversible en F_0 . Si G désigne son inverse locale, G est aussi de classe \mathcal{C}^1 et en $Y = F(X)$, pour X proche de X_0 , on a $dG(y) = dF(X)^{-1}$ (l'exposant désigne ici l'opération d'inversion d'une matrice).

Une démonstration de ce théorème est donnée en annexe.

6.3.2 Fonctions implicites : cas $f(x, y) = 0$

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. On considère la courbe de niveau $\{f(x, y) = 0\} = N_0$.

Définition 6.4. On dit que la fonction $y = \varphi(x)$ est **définie implicitement** par $f(x, y) = 0$ si $f(x, \varphi(x)) = 0$, c'est-à-dire si $(x, \varphi(x)) \in N_0$.

Alors on dit que $y = \varphi(x)$ est une **fonction implicite** de $f(x, y) = 0$.

Exemple

$$f(x, y) = \ln(xy) - \sin x \quad \text{avec } xy > 0$$

$$f(x, y) = x^2 + y^2 - 1. \text{ Faire un dessin!}$$

Théorème 6.5. (des fonctions implicites)

Soient $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 et (x_0, y_0) un point tel que $f(x_0, y_0) = 0$.

Si $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ alors :

(i) Il existe une fonction implicite $y = \varphi(x)$ de classe \mathcal{C}^1 , définie sur l'intervalle ouvert $B(x_0, \epsilon)$, tel que pour tout $x \in B(x_0, \epsilon)$ on ait $y_0 = \varphi(x_0)$ et $f(x, \varphi(x)) = 0$.

(ii) De plus, la dérivée de φ est donnée par $\varphi'(x) = \frac{-\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))}$ en tout point de

$$B(x_0, \epsilon) \text{ où } \frac{\partial f}{\partial y}(x, \varphi(x)) \neq 0.$$

Démonstration C'est une conséquence du théorème d'inversion locale. Soit f une fonction C^1 de deux variables et (x_0, y_0) tel que $f(x_0, y_0) = 0$ et $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Considérons la fonction F définie par

$$F(x, y) = (x, f(x, y)).$$

La matrice jacobienne de F est

$$\begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}.$$

Par hypothèse $\frac{\partial f}{\partial y}$ ne s'annule pas en (x_0, y_0) . La matrice $dF(x_0, y_0)$ est donc inversible et d'après le théorème d'inversion locale, F est localement inversible en (x_0, y_0) : il existe $r > 0$ tel que F soit une bijection de la boule $B = B((x_0, y_0), r)$ sur son image et l'application inverse, appelons la G est C^1 sur l'ouvert $F(B)$. Écrivons $G(s, t) = (g_1(s, t), g_2(s, t))$ les coordonnées de G . Comme G est l'inverse de F on a, pour tout (s, t) dans $F(B)$ (en utilisant la définition de F) :

$$(s, t) = F(g_1(s, t), g_2(s, t)) = (g_1(s, t), f(g_1(s, t), g_2(s, t))).$$

On a donc les égalités : $g_1(s, t) = s$ et $f(s, g_2(s, t)) = t$. Les points (x, y) de B pour lesquels $f(x, y) = 0$ sont les points dont l'image par F est de la forme $(x, 0)$. Ce sont donc les points $G(x, 0)$ pour $(x, 0)$ dans $F(B)$, soit encore, d'après la forme de l'application G , les points $(x, g_2(x, 0))$ pour $(x, 0)$ dans $F(B)$. Or $F(B)$ est un ouvert contenant $(x_0, 0)$. Il existe donc $\alpha > 0$ tel que, pour $x \in]x_0 - \alpha, x_0 + \alpha[$, $(x, y) \in B$, l'équation $f(x, y) = 0$ équivaut à $y = g_2(x, 0)$. Il suffit d'écrire $\phi(x) = g_2(x, 0)$ pour voir qu'on a bien établi le résultat souhaité. \square

Exemple

Le cas du cercle.

Étude au point $(\lambda, 0)$ de $f(x, y) = x(x^2 + y^2) - \lambda(x^2 - y^2)$.

Remarque : On retrouve ainsi une équation de la tangente aux courbes de niveau.

6.3.3 Fonctions implicites : cas $f(x_1 \dots x_n) = 0$

L'étude est similaire pour les hypersurfaces de niveau en plusieurs variables, où on va pouvoir exprimer une variable en fonction des autres si la dérivée partielle correspondante n'est pas nulle.

Théorème 6.6. Si $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est de classe C^1 et si $\frac{\partial f}{\partial x_n}(X_0) \neq 0$ alors :

(i) La fonction implicite $x_n = \varphi(x_1 \dots x_{n-1})$ existe sur une boule ouverte $B((x_{1,0} \dots x_{n-1,0}), \varepsilon)$ et on a : $f(x_1 \dots x_{n-1}, \varphi(x_1 \dots x_{n-1})) = 0$.

$$(ii) \frac{\partial \varphi}{\partial x_i} = - \frac{\frac{\partial f}{\partial x_i}(x_1 \dots x_{n-1}, \varphi(x_1 \dots x_{n-1}))}{\frac{\partial f}{\partial x_n}(x_1 \dots x_{n-1}, \varphi(x_1 \dots x_{n-1}))}$$

7 Optimisation libre et sous contraintes

7.1 En dimension 1

Donnons-nous une fonction d'une variable f définie sur \mathbb{R} et de classe C^2 .

Théorème 7.1. (Formule de Taylor à l'ordre 2) Si f est une fonction trois fois continuellement dérivable sur \mathbb{R} , il existe une fonction ϵ tendant vers 0 en 0 telle que

$$f(x+h) = f(x) + hf'(x) + f''(x)h^2/2 + |h|^2\epsilon(h)$$

Choisissons pour x un point $x = a$ tel que $f''(a) \neq 0$. Alors pour h assez petit, le terme $f''(a)h^2/2 + |h|^2\epsilon(h)$ est du même signe que $f''(a)$. Si par exemple $f''(a) > 0$, on en déduit que f a un minimum local en x .

La recherche pratique des **extrema locaux** pour une fonction d'une variable se passe donc ainsi :

- (i) On recherche les points critiques ($f'(x) = 0$).
- (ii) On étudie la dérivée seconde f'' si a est un point critique et si
 - $f''(a) > 0$ il y a un minimum local,
 - $f''(a) < 0$ il y a un maximum local,
 - $f''(a) = 0$ il faut approfondir l'étude.

Lorsque f n'est plus définie sur \mathbb{R} entier, ou sur un intervalle ouvert, il faudra de plus étudier le comportement de f sur les bords du domaine de définition. si l'ensemble de départ est compact, on a la garantie de l'existence d'extrema globaux.

7.2 Extrema locaux de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

On suppose f de classe C^2 , c'est-à-dire que ses dérivées partielles jusqu'à l'ordre 2 existent et sont continues.

Définition 7.2. f a un **minimum** (resp. **maximum**) local en (x_0, y_0) s'il existe $\epsilon > 0$ tel que : $\forall (x, y) \in B((x_0, y_0), \epsilon)$ alors $f(x, y) \geq f(x_0, y_0)$ (resp. $f(x, y) \leq f(x_0, y_0)$).

Exemple

$$f(x, y) = -x^2 - y^2$$

$$f(x, y) = x^2 + y^2$$

$$f(x, y) = x^2 - y^2$$

Proposition 7.3. Si f admet un **extremum local** en (x_0, y_0) alors $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$.

Démonstration. La fonction de une variable $x \mapsto f(x, y_0)$ admet un extremum local en x_0 , donc sa dérivée $\frac{\partial f}{\partial x}(x_0, y_0)$ en x_0 est nulle. On fait de même avec $y \mapsto f(x_0, y)$. \square

Définition 7.4. Si en (x_0, y_0) on a $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$ on dit que (x_0, y_0) est un **point critique** de f ou un **point stationnaire** de f .

Remarque

Un extremum local est un point critique mais la réciproque n'est pas vraie.

7.3 Formule de Taylor

Définition 7.5. Soit $f(x, y)$ une fonction de classe \mathcal{C}^2 . La **matrice hessienne** de f en (x_0, y_0) est la matrice $\text{Hess}(x_0, y_0) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$

où $A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)$, $B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$, $C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$.

Exemple

Calculer la matrice Hessienne de $f(x, y) = 4xy - x^4 - y^4$.

Théorème 7.6. (Formule de Taylor à l'ordre 2, en $X = (x_0, y_0)$)

Si f une fonction de classe \mathcal{C}^2 sur \mathbb{R}^2 , il existe une fonction ϵ tendant vers 0 en $(0, 0)$ telle que

$$f(X + H) = f(X) + \langle \nabla f(X), H \rangle + \frac{1}{2} H^t \text{Hess}(x_0, y_0) H + \|H\|^2 \epsilon(H)$$

Démonstration dans le cas où la hessienne est diagonale.

Théorème 7.7. Soient $f(x, y)$ de classe \mathcal{C}^2 et (x_0, y_0) un point critique. Soit $\Delta = AC - B^2 = \det(\text{Hess}(x_0, y_0))$. Alors :

si $\Delta > 0$ et $A > 0$, f a un minimum local en (x_0, y_0)

si $\Delta > 0$ et $A < 0$, f a un maximum local en (x_0, y_0)

si $\Delta < 0$, f n'a ni maximum ni minimum, elle a un point selle

si $\Delta = 0$ on ne peut conclure (avec le seul développement à l'ordre 2).

Exemple

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x^3 + y^3 - 3xy$. Les points critiques de f sont $(0, 0)$ et $(1, 1)$. Le premier est un point col, le second un minimum local (non global).

7.4 Extrema de f sur un compact $K \subset \mathbb{R}^2$

La marche à suivre pour étudier les extrema d'une fonction différentiable sur un compact de \mathbb{R}^2 est la suivante.

Soit f une fonction différentiable définie sur un compact K de \mathbb{R}^2 . Comme f est différentiable, elle est continue. Elle est donc bornée sur K et atteint ses bornes.

En pratique (dans les exercices que je vous demanderai de résoudre en particulier) la fonction f sera donnée par une formule valable sur un certain sous-ensemble de \mathbb{R}^2 et le compact K sera inclus dans cet ensemble de définition.

On mène l'étude des extrema de f en plusieurs étapes. La première est d'étudier l'existence d'extrema locaux de f à l'intérieur de K . C'est pour cette étude qu'on utilisera le développement de Taylor à l'ordre 2 donné ci-dessus.

Mais cette étude n'est pas suffisante. Il faut aussi regarder ce qui se passe sur le bord de K . Pour cela on procède autrement.

Exemple

Etude de $f(x, y) = x^2 - y^2$ sur $K = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\}$. On procède de la manière suivante :

- (i) On cherche les points critiques et les extrema locaux dans $\text{Int}(K)$.

On trouve un seul point stationnaire en $(0, 0)$. Mais en $(0, 0)$ f a un point selle. La fonction n'a donc pas d'extremum à l'intérieur de K . Mais comme K est compact et f est continue sur K , f est bornée sur K et atteint ses bornes sur K . Ce sera donc sur le bord de K .

- (ii) On analyse f sur ∂K .

Une possibilité ici est de paramétrer le bord de K : le cercle de rayon 1 centré en $(0, 0)$. On obtient : $f(\cos t, \sin t) = \cos^2 t - \sin^2 t = \cos(2t)$. On peut alors étudier les variations de cette fonction. On obtient qu'elle est maximum égale à 1 lorsque $2t$ est égal à 0 modulo 2π , minimum égale à -1 lorsque $2t$ vaut π modulo 2π . La fonction f atteint donc son maximum 1 aux points $(1, 0)$ et $(-1, 0)$ de K , son minimum -1 aux points $(0, 1)$ et $(0, -1)$.

7.5 Extrema liés (multiplicateur de Lagrange)

7.5.1 Une seule contrainte

Il s'agit de trouver les extrema de $f(x, y, z)$ lorsque (x, y, z) appartient à une surface S définie par $g(x, y, z) = c$.

Exemple

Maximiser $x^2 y^2 z^2$ lorsque $x^2 + y^2 + z^2 = 1$.

Définition 7.8. Un point $X_0 = (x_0, y_0, z_0)$ est un minimum (resp. maximum) local pour f , lié à la contrainte $g(x, y, z) = c$ si :

(i) $g(X_0) = c$

(ii) Il existe $r > 0$ tel que $f(X_0) \leq f(X)$ (resp. $f(X_0) \geq f(X)$) pour tout $X \in S \cap B(X_0, r)$.

Théorème 7.9. (de Lagrange)

Soit $f(x, y, z)$ et $g(x, y, z)$ de classe C^1 telle que $\nabla g \neq 0$ sur S .

Alors si f admet un **extrema lié** en (x_0, y_0, z_0) on a : $\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$ où $\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.

Démonstration Soit $\gamma : \mathbb{R} \rightarrow S$ passant par X_0 en $t = t_0$. Alors $f \circ \gamma$ a un extremum local en t_0 donc sa dérivée s'annule, c'est-à-dire

$$\langle \nabla f(X_0), \gamma'(t_0) \rangle = 0.$$

Ceci étant valable pour toute courbe γ passant par X_0 , on en déduit que le gradient de f en X_0 est orthogonal au plan tangent à S en X_0 , donc est colinéaire au gradient de g en X_0 . \square

Remarque

Si P est un extremum lié, on a $\nabla f(P)$ parallèle à $\nabla g(P)$. La réciproque n'est pas vraie. Nous avons une condition nécessaire mais pas suffisante. C'est l'équivalent de la nullité de la dérivée pour les extrema libres : en un extremum libre la dérivée est nulle mais la dérivée peut être nulle sans que la fonction ait un extremum (penser à $x \mapsto x^3$ en $x = 0$).

Exemple

Sur l'exemple précédent on montre la méthode de résolution.

Autre exemple

Maximum de la distance de l'ellipsoïde donné par $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ à l'origine.

Le théorème est encore vrai en dimension plus grande.

Théorème 7.10. (de Lagrange)

Soit f et g deux fonctions de \mathbb{R}^d dans \mathbb{R} de classe C^1 telle que $\nabla g \neq 0$ sur l'ensemble S défini par $g = C$ (c'est une hypersurface).

Alors si f admet un **extrema lié** en x_0 on a : $\nabla f(x_0) = \lambda \nabla g(x_0)$ où $\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.

7.5.2 Plusieurs contraintes

On cherche les extrema d'une fonction f sur l'ensemble S défini par $g_1 = g_2 = \dots = g_k = 0$, toutes les fonctions considérées étant de classe C^1 . Pour que les choses marchent bien il faut faire l'hypothèse suivante : en tout point de S les gradients des fonctions g_i sont linéairement indépendants.

Théorème 7.11. (de Lagrange)

Soit f et g_1, \dots, g_k $k + 1$ fonctions de \mathbb{R}^d dans \mathbb{R} de classe \mathcal{C}^1 telles que les vecteurs $\nabla g_1, \dots, \nabla g_k$, soit indépendants sur l'ensemble S défini par $g_1 = \dots = g_k = 0$.

Alors si f admet un **extrema lié** sur S en x_0 le vecteur $\nabla f(x_0)$ est combinaison linéaire des vecteurs $\nabla g_i(x_0)$: il existe $\lambda_1, \dots, \lambda_k$ tels que $\nabla f(x_0) = \sum_{i=1}^k \lambda_i \nabla g_i(x_0)$. Les nombres $\lambda_1, \dots, \lambda_k$ sont appelés multiplicateurs de Lagrange.

8 Intégration des fonctions de \mathbb{R}^n dans \mathbb{R}

8.1 Intégration des fonctions d'une variable

Au lycée on définit l'intégrale d'une fonction (continue) positive comme étant l'aire sous la courbe. Pour cela, il faut savoir ce qu'est l'aire... On rappelle en quelques mots ci-dessous la définition de l'intégrale selon Riemann.

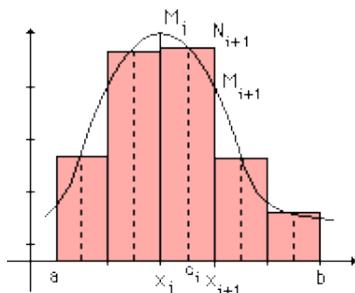
Soit $f : [a, b] \rightarrow \mathbb{R}$ bornée.

a) Cas où f est en escalier.

Il existe une partition $a = t_0 < t_1 < t_2 < \dots < t_n = b$ telle que f est constante sur chaque intervalle $]t_i, t_{i+1}[$ ($f(x) = C_i$).

$$\text{Alors } \int_a^b f(t) dt = \sum_{i=0}^{n-1} C_i (t_{i+1} - t_i).$$

b) Si f est bornée on l'approche par des fonctions en escalier.



Définition 8.1. f est **intégrable** sur $[a, b]$ s'il existe un nombre unique I tel que pour toutes fonctions en escalier u, v sur $[a, b]$ vérifiant $u(x) \leq f(x) \leq v(x)$ on a :

$$\int_a^b u(x) dx \leq I \leq \int_a^b v(x) dx$$

et si pour tout $\varepsilon > 0$ il existe des fonctions en escalier u_ε et v_ε vérifiant :

$$u_\varepsilon(x) \leq f(x) \leq v_\varepsilon(x)$$

et

$$0 \leq \int_a^b v_\varepsilon(x) dx - \int_a^b u_\varepsilon(x) dx < \varepsilon$$

Notation : I s'appelle l'intégrale de f sur $[a, b]$ et se note $\int_a^b f(x) dx$.

Théorème 8.2. Si f est continue sur $[a, b]$ alors f est intégrable sur $[a, b]$.

Démonstration : Soit f une fonction continue sur $[a, b]$. Soit ϵ un nombre réel strictement positif. Alors la fonction f est uniformément continue sur $[a, b]$: il existe $\eta > 0$ tel que, si $|x - y| < \eta$ alors $|f(x) - f(y)| < \epsilon$. Prenons n tel que $(b - a)/n < \eta$ et divisons $[a, b]$ en n intervalles de longueur $(b - a)/n < \eta$ dont les extrémités sont les points $x_k = a + k(b - a)/n$, k variant de 0 à n (k prenant $n + 1$ valeurs détermine n intervalles $[x_k, x_{k+1}]$ k variant de 0 à $n - 1$; on a $x_0 = a$ et $x_n = b$). Définissons deux fonctions en escalier

$$f^-(x) = \min\{f(t), t \in [x_k, x_{k+1}]\} \text{ si } x \in [x_k, x_{k+1}]$$

$$f^+(x) = \max\{f(t), t \in [x_k, x_{k+1}]\} \text{ si } x \in [x_k, x_{k+1}]$$

Par définition on a $f^- \leq f \leq f^+$ et on a

$$\begin{aligned} \int_a^b f^+(x)dx - \int_a^b f^-(x)dx &= \sum_{k=0}^{n-1} (x_{k+1} - x_k) \max\{f(t), t \in [x_k, x_{k+1}]\} \\ &\quad - \sum_{k=0}^{n-1} (x_{k+1} - x_k) \min\{f(t), t \in [x_k, x_{k+1}]\} \\ &= \sum_{k=0}^{n-1} (x_{k+1} - x_k) \left(\max_{t \in [x_k, x_{k+1}]} f(t) - \min_{t \in [x_k, x_{k+1}]} f(t) \right) \\ &\leq \sum_{k=0}^{n-1} (x_{k+1} - x_k) \epsilon \\ &= \epsilon \sum_{k=0}^{n-1} (x_{k+1} - x_k) \\ &= \epsilon(b - a) \end{aligned}$$

□

8.2 Volume de parties bornées de \mathbb{R}^d

On cherche à définir le volume d'une partie. Le volume d'un pavé est donné par le produit des longueurs de ses côtés. Si une partie est une réunion disjointe de pavés alors son volume est la somme des volumes des pavés intervenant dans la réunion. Que le pavé soit ouvert, fermé, ni ouvert ni fermé ne change rien à son volume. On vérifie que ces règles permettent de définir sans ambiguïté et de manière cohérente le volume de toute réunion disjointe finie de pavés. Il faut montrer que le résultat ne dépend pas de la façon de découper en pavés.

On souhaite évidemment définir le volume de parties plus générales.

Pour définir le volume d'une partie donnée E on peut ensuite procéder de la façon suivante. Supposons qu'il existe deux suites d'ensembles $(E_n^-)_{n \in \mathbb{N}}$ et $(E_n^+)_{n \in \mathbb{N}}$ telles que, pour tout

n , E_n^- et E_n^+ soit des réunions finies de pavés, et on ait :

$$E_n^- \subset E_{n+1}^- \subset E \subset E_{n+1}^+ \subset E_n^+,$$

et $\lim vol(E_n^+) - vol(E_n^-) = 0$. Alors les deux suites $vol(E_n^-)$ et $vol(E_n^+)$ sont adjacentes. Elles sont donc convergentes et ont même limite. On définit le volume de E comme étant cette limite commune.

8.3 Intégration des fonctions de deux variables

Soit $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$. On note R le rectangle $[a, b] \times [c, d] \subset \mathbb{R}^2$. On va définir l'intégrale d'une fonction sur R en deux temps : d'abord dans le cas où la fonction est en escalier, puis en approchant la fonction considérée (si c'est possible) par des fonctions en escalier.

a) f est en escalier.

Il existe une partition de $[a, b] \times [c, d]$:

$$a = s_0 < s_1 < s_2 < \dots < s_m = b$$

$$c = t_0 < t_1 < \dots < t_n = d$$

telle que f est constante à l'intérieur de chaque rectangle $]s_i, s_{i+1}[\times]t_j, t_{j+1}[$ (où elle vaut C_{ij}).

$$\text{On définit } \iint_R f(x, y) dx dy = \sum_{i,j} C_{ij} (s_{i+1} - s_i) (t_{j+1} - t_j).$$

On remarque en particulier que la valeur de l'intégrale ne dépend pas des valeurs de f sur les bords des petits rectangles.

b) f est bornée sur $R = [a, b] \times [c, d]$.

On approche f par des fonctions en escalier.

Définition 8.3. f est intégrable sur R s'il existe un nombre unique I tel que pour toutes fonctions en escalier $u(x, y)$ et $v(x, y)$, telles que $u(x, y) \leq f(x, y) \leq v(x, y)$, on a :

$$\iint_R u(x, y) dx dy \leq I \leq \iint_R v(x, y) dx dy$$

et si, pour tout $\varepsilon > 0$, il existe des fonctions en escalier u_ε et v_ε telles que :

$$u_\varepsilon(x, y) \leq f(x, y) \leq v_\varepsilon(x, y)$$

et

$$0 \leq \iint_R v_\varepsilon(x, y) dx dy - \iint_R u_\varepsilon(x, y) dx dy < \varepsilon$$

Notation : I s'appelle l'intégrale de f sur R et se note $\iint_R f(x, y) dx dy$.

Théorème 8.4. 1. Si f est continue sur R alors f est intégrable.

2. Si f est positive sur R alors $\iint_R f(x, y) dx dy$ est le volume sous le graphe de f au-dessus de R .

Proposition 8.5. (Propriétés de l'intégrale double)

- $\iint_R (\alpha f + \beta g)(x, y) dx dy = \alpha \iint_R f(x, y) dx dy + \beta \iint_R g(x, y) dx dy.$
- Si $R = R_1 \cup R_2$ avec $R_1 \cap R_2 = \emptyset$ alors :

$$\iint_R f(x, y) dx dy = \iint_{R_1} f(x, y) dx dy + \iint_{R_2} f(x, y) dx dy$$

Théorème 8.6. Si f est continue sur R alors $\iint_R f(x, y) dx dy$ existe.

8.4 Calcul des intégrales doubles

Théorème 8.7. (Fubini) Si f est continue sur $R = [a, b] \times [c, d]$ alors :

$$\iint_R f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

Exemple

$$(1) \iint_R (x^2 + y^2) dx dy \quad R = [0, 1] \times [0, 1]$$

$$(2) \iint_R (1 + x + y) dx dy \quad R = [0, 1] \times [0, 1]$$

Corollaire 8.8. Si la fonction f est le produit de deux fonctions g et h d'une variable, c'est-à-dire $f(x, y) = g(x)h(y)$, alors

$$\iint_R f(x, y) dx dy = \left(\int_a^b g(x) dx \right) \left(\int_c^d h(y) dy \right).$$

Exemple : $\iint_{[0,1] \times [0,1]} e^{x+y} dx dy$

8.5 Intégration sur les régions bornées de \mathbb{R}^2

Soit $f : D \rightarrow \mathbb{R}$ où $D \subset \mathbb{R}^2$ est non rectangulaire.

On considère un rectangle R tel que $D \subset R$ et on définit \bar{f} sur R avec :

$$\bar{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in D \\ 0 & \text{si } (x, y) \notin D \end{cases}$$

Avec les notations précédentes on pose, si cela a un sens :

$$\iint_D f(x, y) dx dy = \iint_R \bar{f}(x, y) dx dy$$

On peut se ramener à deux types de domaine D :

$$\text{Type 1 : } D = \left\{ (x, y) \ / \ \begin{array}{l} a \leq x \leq b \\ g_1(x) \leq y \leq g_2(x) \end{array} \right\} \text{ où } g_1 \text{ et } g_2 \text{ sont continues.}$$

$$\text{Type 2 : } D = \left\{ (x, y) \ / \ \begin{array}{l} c \leq y \leq d \\ h_1(y) \leq x \leq h_2(y) \end{array} \right\}.$$

Théorème 8.9. (Fubini)

a) Si f est continue sur D de type 1, alors f est intégrable et on a :

$$\iint_D f(x, y) dx dy = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx$$

b) Si f est continue sur D de type 2, alors f est intégrable et on a :

$$\iint_D f(x, y) dx dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy$$

Exemples

(1) $\iint_D (x + 2y) dx dy$: D est la région entre les deux paraboles $y = 2x^2$ et $y = 1 + x^2$.

(2) $\iint_D e^{x^2} dx dy$ sur le triangle $D = \left\{ (x, y) \ / \ \begin{array}{l} 0 \leq x \leq 1 \\ 0 \leq y \leq x \end{array} \right\}$.

(3) Le choix d'intégrer d'abord par rapport à x ou y peut amener des calculs plus ou moins long. Par exemple avec $\iint_D xy dx dy$ où D est le trapèze délimité par $y = 0$, $y = 1$ et les droites d'équation $y = 2 - x$ et $y = 1 + x/2$.

Définition 8.10. Si D est un domaine borné, on appelle **aire de D** : $\text{aire}(D) = \iint_D 1 dx dy$.

8.6 Intégrale double et changement de variables

Rappel à une variable :

$$\int_a^b f(x) dx = \int_c^d f(g(t)) g'(t) dt$$

où g est bijection de $[c, d]$ sur $[a, b]$.

Démonstration : Soit F une primitive de f .

On a d'une part

$$F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} F'(t) dt = \int_{g(a)}^{g(b)} f(t) dt,$$

d'autre part

$$F(g(b)) - F(g(a)) = \int_a^b (F \circ g)'(s) ds = \int_a^b F'(g(s)) g'(s) ds = \int_a^b f(g(s)) g'(s) ds.$$

□

Théorème 8.11. Si $G(u, v) = (x(u, v), y(u, v))$:

$$\iint_{G(S)} f(x, y) dx dy = \iint_S f(x(u, v), y(u, v)) |\det \text{Jac}(G(u, v))| du dv$$

$$\text{avec } \text{Jac}(G(u, v)) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

Remarque 8.12. Ce n'est pas très étonnant de trouver là le déterminant. Par exemple, la valeur absolue du déterminant d'une matrice 2×2 calcule l'aire du parallélogramme engendré par les vecteurs colonnes (par exemple).

Cas des coordonnées polaires

$$x = x(r, \theta) = r \cos \theta$$

$$y = y(r, \theta) = r \sin \theta$$

$$J(G) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

$$\text{d'où } \iint_{R=G(S)} f(x, y) dx dy = \iint_S f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Exemples

(1) Calcul de l'aire d'un disque.

(2) Calcul de l'aire à l'intérieur d'une ellipse.

$$(3) \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

(4) $\iint_D (x-y)^2 dx dy$ où D est le morceau du disque unité compris entre l'axe des x et la demi-droite $y = x$.

8.7 Intégrales triples

On définit et calcule de manière totalement similaire les intégrales pour les fonctions de trois variables (et plus...). On va voir quelques exemples...

$$\iiint_R f(x, y, z) dx dy dz$$

Exemple : $\iiint_D (x + y + z)^2 dx dy dz = 1/10$ où D est le domaine délimité par les plans d'équations $x = 0, y = 0, z = 0$ et $x + y + z = 1$.

Changement de variables en dimension 3. Cas des coordonnées sphériques.

$$\iiint_{R=G(S)} f(x, y, z) dx dy dz = \iiint_S f(r \cos \theta \sin \phi, r \sin \theta \sin \phi, r \cos \phi) r^2 \sin \phi dr d\theta d\phi.$$

Exemple : $\iiint_D (x^2 + y^2 + z^2) dx dy dz = 4\pi/5$ où D est la boule centrée en l'origine et de rayon 1.

8.8 Quelques calculs classiques

8.8.1 L'aire d'un disque

On peut trouver l'aire d'un disque grâce au calcul intégral. En intégrant par tranche :

$$\begin{aligned} \text{Aire} &= 4 \int_0^R \sqrt{R^2 - x^2} dx \\ &= 4 \int_0^{\pi/2} R \cos(t) R \cos(t) dt \\ &= 4R^2 \int_0^{\pi/2} \frac{1 + \cos(2t)}{2} dt \\ &= \pi R^2, \end{aligned}$$

calculée grâce au changement de variable $x = R \sin(t)$. On peut aussi (et c'est plus simple quand on connaît l'expression du jacobien du passage en polaires) utiliser les coordonnées polaires

$$\begin{aligned} \text{Aire} &= \int_0^R \int_0^{2\pi} r dr d\theta \\ &= 2\pi \cdot R^2/2 \\ &= \pi R^2. \end{aligned}$$

8.8.2 Le volume de la boule

On peut calculer le volume de la boule par un changement de variables en coordonnées sphériques.

$$\iiint_{B(0,R)} dx dy dz = \iiint_{[0,R] \times [0,2\pi] \times [0,\pi]} r^2 \sin \phi dr d\theta d\phi = \int_0^R r^2 dr \int_0^{2\pi} d\theta \int_0^\pi \sin \phi d\phi$$

et donc le volume vaut $4\pi R^3/3$.

8.8.3 Le volume d'une pyramide

Cas d'une pyramide P de base carrée de côté de longueur a et de hauteur h . On pose la pyramide sur le plan $z = 0$, centrée sur l'axe des z . La longueur du côté du carré C_z à la hauteur z est donc $a(1 - z/h)$. Ainsi par le théorème de Fubini

$$\iiint_P dx dy dz = \int_0^h \left(\iint_{C_z} dx dy \right) dz = \int_0^h (a^2(1 - z/h)^2) dz = a^2 h/3 = Sh/3$$

où S est l'aire de la base.

8.8.4 Solides de révolution

Soit f une fonction positive ou nulle définie sur un intervalle $[a, b]$. Considérons la partie de l'espace définie de la façon suivante :

$$V = \{(x, y, z) / x \in [a, b], \sqrt{y^2 + z^2} \leq f(x)\}.$$

C'est le solide obtenu en faisant tourner le graphe de f autour de l'axe des x . Le volume de V est donné par l'intégrale triple

$$\iiint_V dx dy dz.$$

En intégrant par tranche (d'abord en y, z , puis en x) on obtient :

$$\iiint_V dx dy dz = \int_a^b \left(\iint_{\{(y,z) / \sqrt{y^2+z^2} \leq f(x)\}} dy dz \right) dx = \int_a^b \pi f(x)^2 dx.$$

Le calcul de l'intégrale triple se ramène donc à un calcul d'intégrale simple.

Exemples

- (1) Cas d'un cône : on prend pour f la fonction définie sur $[0, h]$ par $f(x) = ax$ pour $a > 0$. Ici la base du cône a pour aire $S = \pi(ah)^2$, et donc le volume du cône est égal à

$$\int_0^h \pi(ax)^2 dx = \pi a^2 h^3/3 = Sh/3.$$

(2) Cas d'un cylindre de rayon R et de hauteur h :

$$\int_0^h \pi R^2 dx = \pi R^2 h = Sh$$

où S est l'aire de la base.