Fonctions de plusieurs variables

De omni quarta analysi et quibusdam aliis

Περιήγησις

Pourquoi étudier les fonctions de plusieurs variables ? 1

1 La norme euclidienne 3

2 Ensembles remarquables 7

3 Continuité. Limites. 9

4 La différentielle 13

5 Calcul différentiel d'ordre supérieur 21

6 Étude locale de fonctions 25

7 Convexité 27

8 Modélisation. Optimisation sous contraintes 29

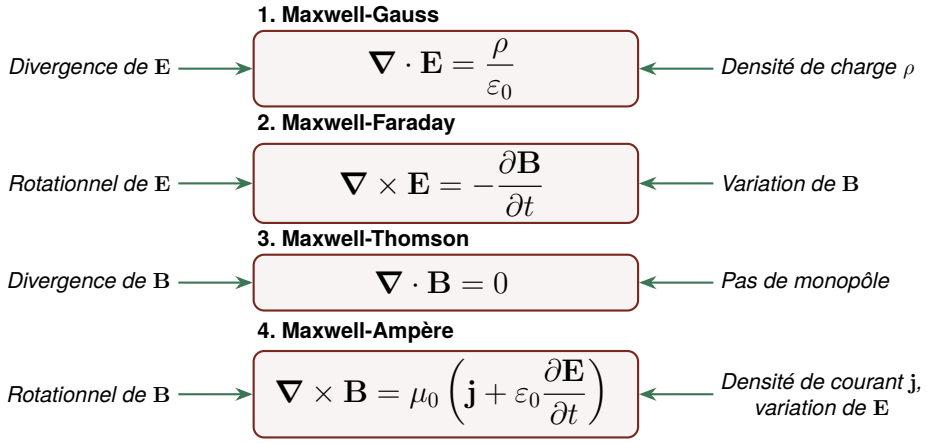
9 Arcs paramétrés. Courbes 33

Complément. Preuve du théorème des multiplicateurs de Lagrange 38

Pourquoi étudier les fonctions de plusieurs variables ?



Combien de fonctions de plusieurs variables dans cette image ? Tobias Bjorkli 2019



« [Il libro dell'universo] è scritto in lingua matematica » G. Galilei, Il Saggiatore (1623)
$\boldsymbol{\nabla}\cdot$: divergence $\boldsymbol{\nabla}\times$: rotationnel $\mathbf{E}$: champ électrique $\mathbf{B}$: champ magnétique
ρ : densité de charge $\mathbf{j}$: densité de courant ε_0 : permittivité du vide μ_0 : perméabilité du vide

Grandes lignes du cours

— Les trois premières leçons sont très proches de ce qui a été vu dans $\mathbb{R}$: les notions de distance, suite convergente, limite, fonction continue, seront transposées, sans difficulté notable, à $\mathbb{R}^n$. La principale nouveauté dans cette partie vient de la plus grande variété d'ensembles d'intérêt. Dans $\mathbb{R}$, les fonctions sont souvent définies sur un intervalle. Dans $\mathbb{R}^n$, il y a beaucoup plus de choix
— Le saut qualitatif intervient dans la quatrième leçon. Il s'agit de l'analogue de la dérivée pour les fonctions de plusieurs variables. Alors que pour les fonctions réelles $f'(a)$ est un nombre, son analogue pour des fonctions de plusieurs variables est une application linéaire. S'approprier ce nouveau paradigme sera le défi majeur de l'UE
— La cinquième leçon porte sur l'analogue de la dérivée seconde. De nouveaux phénomènes apparaissent, comme le théorème « de symétrie de Schwarz »
— Les trois leçons suivantes sont dédiées à la transposition de thèmes vus pour les fonctions réelles : étude de fonctions, modélisation, recherche de points d'extrémum, convexité. La principale nouveauté viendra d'un type de problème qui n'est pas pertinent pour les fonctions réelles : l'optimisation sous contraintes, traitée grâce au très beau théorème « des multiplicateurs de Lagrange »
— La dernière leçon est une introduction à l'étude des courbes

Apport de l'IA générative aux supports pédagogiques

— Une IA générative a fourni, à partir d'un cahier de charges précis, le modèle LaTeX du support de cours (mise en page, choix des polices et des couleurs) et le code LaTeX des illustrations du support de cours
— Des statistiques sur les usages, effectuées par l'IA, ont servi à choisir certaines notations ou noms de théorèmes parmi plusieurs variantes utilisées dans la littérature scientifique
— Une IA générative a été sollicitée pour fournir des corrigés à certains exercices. Le but était de pouvoir juger la qualité des réponses auxquelles ont accès les étudiants. Quelques-unes de ces réponses, souvent remaniées après des requêtes supplémentaires et des fois éditées pour les rendre plus proches de notations et résultats de cette UE, apparaissent sur les feuilles de TD. Elles ont vocation à servir de base de discussion entre les étudiants et l'équipe pédagogique, notamment au sujet des bénéfices et points de vigilance liés à l'apprentissage avec l'IA
— L'IA générative a également été sollicitée pour fournir des « mises en situation » pour deux exercices de modélisation et de nouveaux exemples de fonctions à partir de fonctions modèles
— Le texte comporte des diagrammes et des « patatoïdes » qui ont été encodés par une IA générative à partir de requêtes précises. Par ailleurs, une IA générative a contribué à l'intégration de vraies images dans le texte

Lexique, notations

- (i) ⚠ attire l'attention ; (ii) $:=$ signifie « est par définition égal à » ; (iii) ↯ marque une contradiction
- Une fonction $f : I \to \mathbb{R}$, avec $I \subset \mathbb{R}$, est une fonction « réelle »
- Une fonction $f : A \to \mathbb{R}$, avec $A \subset \mathbb{R}^n$, est une fonction « scalaire »
- Une fonction $f : A \to \mathbb{R}^n$, avec $A \subset \mathbb{R}^n$, est un « champ de vecteurs »
- ⚠ Si f est une fonction, on peut écrire $f = f(x)$ (qui est un abus de notation) pour signifier que l'on désigne par x l'argument (la variable) de f
- ⚠ La plupart du temps, x ne sera plus un réel, mais bien un « point » de $\mathbb{R}^n$, le plus souvent identifié avec le vecteur de ses coordonnées dans la base canonique
- Une fonction est continue si elle est continue en tout point, respectivement dérivable (différentiable) si elle est dérivable (différentiable) en tout point
- Le transposé d'un vecteur $v \in \mathbb{R}^n$ est noté $v^\intercal$. La transposée d'une matrice $A \in M_{m,n}(\mathbb{R})$ est notée $A^\intercal$
- ⚠ Les quantificateurs ne seront pas systématiquement à gauche. Quand la compréhension n'est pas tributaire de leur place, la propriété sera mise en premier, les quantificateurs en dernier. Exemple : $\sqrt{x^2} = |x|$, $\forall\, x \in \mathbb{R}$, plutôt que $\forall\, x \in \mathbb{R}$, $\sqrt{x^2} = |x|$

À réviser pour la première leçon

- Utilisation du symbole somme $\sum$ et développement de $(a_1 + a_2 + \cdots + a_n)^2$ en utilisant $\sum$
- L'inégalité de Cauchy–Schwarz

$$\forall\, a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R},\ \left|\sum_{j=1}^{n} a_j b_j\right| \leq \left(\sum_{j=1}^{n} (a_j)^2\right)^{1/2} \left(\sum_{j=1}^{n} (b_j)^2\right)^{1/2},$$

  et sa preuve (voir aussi l'exercice TD 1.5)
- Si (x_k) est une suite réelle et $x \in \mathbb{R}$, la définition de $x_k \to x$
- Si $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$, $a \in \mathbb{R}$ et $\ell \in \mathbb{R}$, la définition de $\lim\limits_{x \to a} f(x) = \ell$
- Si $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$ et $a \in \mathbb{R}$, la définition de $\lim\limits_{x \to a} f(x) = \infty$
- Si $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$ et $a \in A$:
  - Définition « avec ε et δ » de la continuité de f en a
  - Caractérisation avec des suites de la continuité de f en a, et la preuve de cette caractérisation

Leçon 1. La norme euclidienne

Objectifs

Établir un parallèle entre $(\mathbb{R}, |\ |)$ et $(\mathbb{R}^n, \|\ \|_2)$ en ce qui concerne : (i) les distances entre les points ; (ii) les suites ; (iii) les limites ; (iv) la continuité. Réviser : (i) les limites de suites ; (ii) les limites et la continuité de fonctions.

Notations

- Un point de $\mathbb{R}^n$ est noté x ou $(x_1, x_2, \ldots, x_n)$
- $x_1, \ldots, x_n$ sont les coordonnées de x dans la base canonique $(e^1, \ldots, e^n)$ de $\mathbb{R}^n$
- Si $n = 1$, x est identifié à un scalaire ; si $n = 2$, on note un point (x, y) plutôt que (x_1, x_2) ; si $n = 3$, (x, y, z) plutôt que (x_1, x_2, x_3)
- Pour garder les indices « d'en bas » pour les coordonnées, les suites sont notées (x^k)
- ⚠ Ci-dessus, k n'est pas un puissance, mais le rang de x^k dans la suite
- $(x^k) \subset A$ signifie que tous les termes de la suite (x^k) sont dans l'ensemble A
- ⚠ Si $(x^k) \subset \mathbb{R}^n$, $x^k = (x_1^k, x_2^k, \ldots, x_n^k)$, avec x_j^k la j^{e} coordonnée de x^k
- ⚠ Par défaut, le glyphe k désigne un entier ≥ 1, j un entier dans $[\![1, n]\!]$

La norme euclidienne

Si $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\|x\|_2 := \sqrt{(x_1)^2 + (x_2)^2 + \cdots + (x_n)^2} = \left(\sum_{j=1}^{n} (x_j)^2 \right)^{1/2}$.

Cas particuliers

— Si $n = 1$, $\|x\|_2 = \sqrt{|x|^2} = |x|$ (la valeur absolue de $x \in \mathbb{R}$)
— Si $n = 2$, $\|(x, y)\|_2 = \sqrt{x^2 + y^2} = |z|$, avec $z := x + \imath y \in \mathbb{C}$ (le module de $z \in \mathbb{C}$)

Propriétés fondamentales de la norme euclidienne

Proposition 1.1 La norme euclidienne a les propriétés suivantes :
(i) positivité : $\forall\, x \in \mathbb{R}^n$, $\|x\|_2 \geq 0$; (ii) séparation : $\|x\|_2 = 0 \Rightarrow x = 0$; (iii) homogénéité : $\forall\, x \in \mathbb{R}^n$, $\forall\, \lambda \in \mathbb{R}$, $\|\lambda x\|_2 = |\lambda|\, \|x\|_2$; (iv) inégalité triangulaire : $\forall\, x, y \in \mathbb{R}^n$, $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$.

Remarques

— ⚠ Le glyphe 0 a plusieurs significations. Il est conseillé de comprendre de quel 0 il s'agit dans une formule. Exemples dans les propriétés (i) et (ii) ci-dessus : $\|x\|_2 \geq 0_{\mathbb{R}}$; $\|x\|_2 = 0_{\mathbb{R}} \Rightarrow x = 0_{\mathbb{R}^n}$
— Les propriétés (i)–(iv) ci-dessus définissent une norme sur un espace vectoriel réel
— Des espaces plus généraux que $\mathbb{R}^n$ et des normes plus générales que $\| \ \|_2$ seront étudiés en L3

Esquisse de preuve de la proposition 1.1 (iv)

$$\|x + y\|_2 \leq \|x\|_2 + \|y\|_2 \iff (\|x + y\|_2)^2 \leq (\|x\|_2 + \|y\|_2)^2$$
$$\iff \sum_{j=1}^{n} x_j y_j \leq \|x\|_2 \|y\|_2, \text{ qui suit de l'inégalité de Cauchy–Schwarz :}$$

$$\forall\, a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}, \quad \left| \sum_{j=1}^{n} a_j b_j \right| \leq \left(\sum_{j=1}^{n} (a_j)^2 \right)^{1/2} \left(\sum_{j=1}^{n} (b_j)^2 \right)^{1/2}. \qquad \textit{CQFD} \quad \text{(CS)}$$

De $|\ |$ vers $\| \ \|_2$: guide pratique

— Pour des notions comme limites de suites, limites de fonctions, continuité, les définitions et la plupart des preuves vues pour les suites et les fonctions réelles se transfèrent aux suites de vecteurs et aux fonctions de plusieurs variables, en remplaçant $|\ |$ par $\| \ \|_2$
— Un exemple : si $(x^k) \subset \mathbb{R}^n$ et $x \in \mathbb{R}^n$, alors $x^k \to x \overset{\text{déf.}}{\iff} \big[\forall\, \varepsilon > 0,\ \exists\, k_0 = k_0(\varepsilon)$ tel que $[k \geq k_0 \implies \|x^k - x\|_2 < \varepsilon]\big]$ ⚠ Définition équivalente : $x^k \to x \overset{\text{déf.}}{\iff} \big[\forall\, \varepsilon > 0,\ \exists\, k_0 = k_0(\varepsilon)$ tel que $[k \geq k_0 \implies \|x^k - x\|_2 \leq \varepsilon]\big]$ ⚠ Voir aussi le principe $r(\varepsilon)$ dans la Leçon 3

Quiz : limites

Préciser le sens de :
— $\lim\limits_{x \to a} f(x) = \ell$, où $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}$, $a \in \mathbb{R}^n$, $\ell \in \mathbb{R}$
— $\lim\limits_{x \to a} f(x) = \ell$, où $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, $a \in \mathbb{R}^n$, $\ell \in \mathbb{R}^m$
— $\lim\limits_{x \to a} f(x) = \infty$, où $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}$, $a \in \mathbb{R}^n$

Remarque

La notation $x^k \to x$ sera privilégiée par rapport à $\lim_{k\to\infty} x^k = x$.

Exemple de passage de $|\ |$ à $\|\ \|_2$

Proposition 1.2 (Unicité de la limite) Si $(x^k) \subset \mathbb{R}^n$, $x, y \in \mathbb{R}^n$, $x^k \to x$, $x^k \to y$, alors $x = y$.

Esquisse de preuve de la proposition 1.2

Soit $\varepsilon > 0$. Soient k_1 tel que $\|x^k - x\|_2 < \varepsilon, \forall\, k \geq k_1$, respectivement k_2 tel que $\|x^k - y\|_2 < \varepsilon, \forall\, k \geq k_2$. Si $k \geq \max(k_1, k_2)$, l'inégalité triangulaire donne $\|y - x\|_2 < 2\varepsilon$. En faisant $\varepsilon \to 0$, nous trouvons $\|y - x\|_2 \leq 0$. Par séparation, $y = x$. *CQFD*

Exercice travaillé

Exercice CM1.3 Si $(x^k) \subset \mathbb{R}^n$, $(y^k) \subset \mathbb{R}^n$, $x, y \in \mathbb{R}^n$, $x^k \to x$, $y^k \to y$, montrer que $x^k + y^k \to x + y$.

Solution de l'exercice CM1.3

Soit $\varepsilon > 0$. Soient k_1 tel que $\|x^k - x\|_2 < \varepsilon/2, \forall\, k \geq k_1$, respectivement k_2 tel que $\|y^k - y\|_2 < \varepsilon/2$, $\forall\, k \geq k_2$. Si $k \geq \max(k_1, k_2)$, l'inégalité triangulaire donne

$$\|(x^k + y^k) - (x + y)\|_2 = \|(x^k - x) + (y^k - y)\|_2 \leq \|x^k - x\|_2 + \|y^k - y\|_2 < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

$\varepsilon > 0$ étant arbitraire, nous obtenons que $x^k + y^k \to x + y$. *CQFD*

Travail individuel

Exercice CM1.4 Soient $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}^m$ et $a \in A$.

- Écrire la définition « avec ε et δ » de la continuité de f en a
- Que devient cette définition si $n = 1$? Et si $m = 1$?
- Montrer que f est continue en a si et seulement si : $\forall\, (x^k) \subset A,\ x^k \to a \implies f(x^k) \to f(a)$

Exercice CM1.5 Si $(x^k) \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$, $a \in \mathbb{R}$, et $x^k \to x$, montrer que $ax^k \to ax$.

Distance euclidienne

Si $x, y \in \mathbb{R}^n$, $d(x, y) \coloneqq \|x - y\|_2$.

Propriétés fondamentales de la distance euclidienne

Proposition 1.6 La distance euclidienne a les propriétés suivantes :
(i) positivité : $\forall\, x, y \in \mathbb{R}^n$, $d(x, y) \geq 0$; (ii) séparation : $d(x, y) = 0 \Rightarrow x = y$; (iii) symétrie : $\forall\, x, y \in \mathbb{R}^n$, $d(x, y) = d(y, x)$; (iv) inégalité triangulaire : $\forall\, x, y, z \in \mathbb{R}^n$, $d(x, z) \leq d(x, y) + d(y, z)$; (v) inégalité triangulaire inversée : $\forall\, x, y, z \in \mathbb{R}^n$, $d(x, z) \geq d(x, y) - d(y, z)$.

Preuve de la proposition 1.6

Soient $x, y, z \in \mathbb{R}^n$. (i) $d(x, y) = \|x - y\|_2 \geq 0$, par positivité de $\|\ \|_2$. (ii) $d(x, y) = 0 \Rightarrow \|x - y\|_2 = 0 \Rightarrow x - y = 0 \Rightarrow x = y$, par séparation de $\|\ \|_2$. (iii) $d(y, x) = \|y - x\|_2 = \|(-1)(x - y)\|_2 = |-1|\|x - y\|_2 = \|x - y\|_2 = d(x, y)$, par homogénéité de $\|\ \|_2$. (iv) $d(x, z) = \|x - z\|_2 = \|(x - y) + (y - z)\|_2 \leq \|x - y\|_2 + \|y - z\|_2 = d(x, y) + d(y, z)$, par inégalité triangulaire pour $\|\ \|_2$. (v) suit de (iv) et (iii), car : $d(x, y) \leq d(x, z) + d(z, y) = d(x, z) + d(y, z)$. *CQFD*

Compléments

— Soit E un espace vectoriel réel. Soient $E \times E \ni (x, y) \mapsto x \cdot y$ un produit scalaire sur E, et $\|x\| := \sqrt{x \cdot x}$, $\forall\, x \in E$. $\|\ \|$ est « la norme associée au produit scalaire $\cdot$ »
— Les propriétés suivantes seront vues en Algèbre IV :
  — $\forall\, x, y \in E$, $|x \cdot y| \leq \|x\|\|y\|$ (inégalité de Cauchy–Schwarz)
  — $E \ni x \mapsto \|x\|$ a les propriétés (i)–(iv) de la proposition 1.1
  — Il y a égalité dans l'inégalité de Cauchy–Schwarz $|x \cdot y| \leq \|x\|\|y\|$ si et seulement si : soit (i) $x = 0$; soit (ii) il existe un scalaire $\lambda \in \mathbb{R}$ tel que $y = \lambda x$. (Et symétriquement en y.)
  — Il y a égalité dans l'inégalité triangulaire $\|x + y\| \leq \|x\| + \|y\|$ si et seulement si : soit (i) $x = 0$; soit (ii) il existe un scalaire $\lambda > 0$ tel que $y = \lambda x$ (Et symétriquement en y.)
— Dans le cas particulier du produit scalaire

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto x \cdot y := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{j=1}^{n} x_j y_j,$$

nous avons $\|x\| = \|x\|_2$, et les deux premières propriétés ci-dessus correspondent à (CS) et à la proposition 1.1

Des inégalités simples, mais utiles

$$\forall\, x \in \mathbb{R}^n,\ \forall\, 1 \leq j \leq n,\ |x_j| \leq \|x\|_2 \leq \sum_{\ell=1}^{n} |x_\ell|. \qquad \text{(IS)}$$

Esquisse de preuve de (IS)

La seconde inégalité suit de l'identité $\left(\sum_{\ell=1}^{n} |x_\ell|\right)^2 = (\|x\|_2)^2 + 2 \sum_{1 \leq \ell < m \leq n} |x_\ell||x_m|$. *CQFD*

Passage en coordonnées

Proposition 1.7 Si $(x^k) \subset \mathbb{R}^n$ et $x \in \mathbb{R}^n$, nous avons $x^k \to x \Leftrightarrow x_j^k \to x_j$, $\forall\, 1 \leq j \leq n$.

Proposition 1.8 Si $A \subset \mathbb{R}^n$, $f = (f_1, \ldots, f_m) : A \to \mathbb{R}^m$ et $a \in A$, alors f est continue en $a \Leftrightarrow f_\ell$ est continue en a, $\forall\, 1 \leq \ell \leq m$.

Esquisses de preuves des propositions 1.7 et 1.8

Proposition 1.7. « $\Rightarrow$ » Nous avons $\left|x_j^k - x_j\right| \leq \left\|x^k - x\right\|_2 \to 0$ (quand $k \to \infty$).

« $\Leftarrow$ » De (IS), nous avons $\left\|x^k - x\right\|_2 \leq \sum_{\ell=1}^{n} \left|x_\ell^k - x_\ell\right| \to 0$. *CQFD*

Proposition 1.8. De la proposition 1.7, nous avons $\left[[(x^k) \subset A, x^k \to x] \Rightarrow [f(x^k) \to f(a)]\right] \Leftrightarrow \left[[(x^k) \subset A, x^k \to x] \Rightarrow [(f(x^k))_\ell \to (f(a))_\ell,\ \forall\, 1 \leq \ell \leq m]\right]$. *CQFD*

À réviser pour la deuxième leçon

— La définition d'une partie bornée de $\mathbb{R}$. La caractérisation avec des suites d'une partie non-bornée de $\mathbb{R}$ (avec la preuve)
— Une preuve du fait que, dans $\mathbb{R}$, une suite convergente est bornée
— Les notions d'extraction et de suite extraite. Le théorème de Bolzano-Weierstrass
— Les notions d'image et image réciproque d'une fonction

— La définition et les propriétés des bornes inférieure et supérieure d'une partie non-vide A de $\mathbb{R}$

Leçon 2. Ensembles remarquables

Objectifs

S'approprier les définitions et les propriétés des ensembles ouverts, fermés, bornés ou compacts, et savoir les reconnaître. S'approprier les théorèmes de Bolzano-Weierstrass et des bornes atteintes.

Ensembles remarquables

— La boule ouverte de centre $a \in \mathbb{R}^n$ et rayon $r > 0$: $B(a,r) \coloneqq \{x \in \mathbb{R}^n \,;\, \|x-a\|_2 < r\}$
— La boule fermée de centre $a \in \mathbb{R}^n$ et rayon $r > 0$: $\overline{B}(a,r) \coloneqq \{x \in \mathbb{R}^n \,;\, \|x-a\|_2 \leq r\}$
— La sphère de centre $a \in \mathbb{R}^n$ et rayon $r > 0$: $S(a,r) \coloneqq \{x \in \mathbb{R}^n \,;\, \|x-a\|_2 = r\}$
— Un ensemble $U \subset \mathbb{R}^n$ est ouvert si : $\forall\, a \in U$, $\exists\, r > 0$ tel que $B(a,r) \subset U$
— Un ensemble $F \subset \mathbb{R}^n$ est fermé si son complémentaire F^c est ouvert
— Un ensemble $A \subset \mathbb{R}^n$ est borné si : $\exists\, M \in \mathbb{R}$ tel que $\forall\, x \in A$, $\|x\|_2 \leq M$
— Un ensemble $K \subset \mathbb{R}^n$ est compact si : K est fermé et borné
— Si $x, y \in \mathbb{R}^n$, le segment d'extrémités x et y est $[x,y] \coloneqq \{(1-t)x + ty \,;\, t \in [0,1]\}$
— Un ensemble $C \subset \mathbb{R}^n$ est convexe si : $\forall\, x, y \in C$, $[x,y] \subset C$

Remarques

— Dans la définition d'un ouvert U, r dépend de a (sauf si $U = \emptyset$ ou $U = \mathbb{R}^n$)
— Dans la définition d'un ensemble borné A, M dépend de A
— ⚠ La notation $[x,y]$ peut prêter à confusion si $x, y \in \mathbb{R}$. Si $x \leq y$, le segment $[x,y]$ coïncide avec l'intervalle $[x,y]$. Si $x > y$, le segment $[x,y]$ coïncide avec l'intervalle $[y,x]$ (Exercice TD1.13)

Exemples fondamentaux

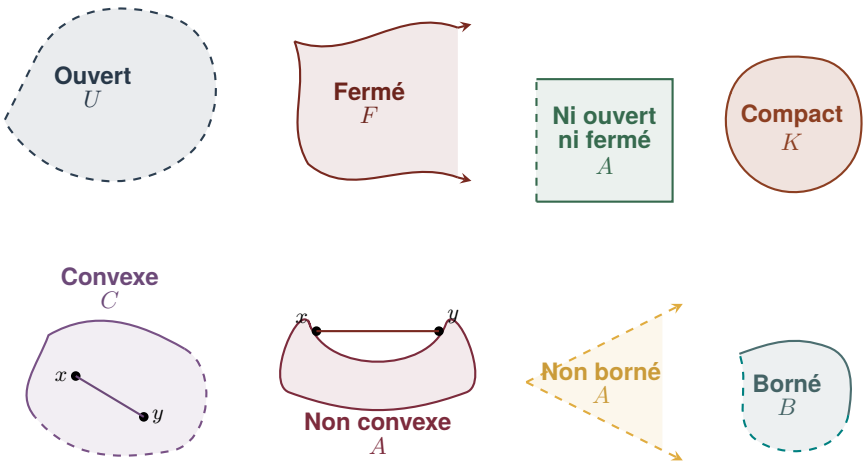
Proposition 2.1 Une boule ouverte est un ouvert. Une boule fermée est un fermé.

Esquisse de preuve de la proposition 2.1

Si $x \in B(a,r)$ et $\rho \coloneqq r - \|x-a\|_2 > 0$, alors $B(x,\rho) \subset B(a,r)$. En effet, $y \in B(x,\rho) \Rightarrow \|y-x\|_2 < \rho \Rightarrow$

$$\|y-a\|_2 = \|(y-x)+(x-a)\|_2 \leq \|y-x\|_2 + \|x-a\|_2 < \rho + \|x-a\|_2 = r \Rightarrow y \in B(a,r).$$

De manière similaire, si $x \in (\overline{B}(a,r))^c$ et $\rho \coloneqq \|x-a\|_2 - r > 0$, alors $B(x,\rho) \subset (\overline{B}(a,r))^c$. *CQFD*



Note : - - - Frontière exclue | —— Frontière incluse

Exemples graphiques d'ensembles remarquables

Propriétés fondamentales des ouverts et fermés

Proposition 2.2 (i) $\emptyset$ et $\mathbb{R}^n$ sont ouverts et fermés. (ii) Une union arbitraire d'ouverts est ouverte. (iii) Une intersection finie d'ouverts est ouverte. (iv) Une union finie de fermés est fermée. (v) Une intersection arbitraire de fermés est fermée. (vi) $F \subset \mathbb{R}^n$ est fermé si et seulement si : $\forall\, (x^k) \subset F$, $\forall\, x \in \mathbb{R}^n$, $x^k \to x \Rightarrow x \in F$. (vii) Si $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ sont ouverts et $f : U \to \mathbb{R}^m$ est continue, $f^{-1}(V)$ est ouvert. (viii) Si $F \subset \mathbb{R}^n$, $G \subset \mathbb{R}^m$ sont fermés et $f : F \to \mathbb{R}^m$ est continue, $f^{-1}(G)$ est fermé.

Remarques

- ⚠ Union arbitraire ou intersection arbitraire : $\cup_{i\in I} U_i$ ou $\cap_{i\in I} F_i$, où I est un ensemble qui n'est pas supposé fini
- ⚠ Union finie ou intersection finie : $\cup_{i=1}^N F_i$ ou $\cap_{i=1}^N U_i$, pour un certain entier N
- Exemples d'utilisation de la proposition 2.2 (vii) ou (viii) : si $f : \mathbb{R}^n \to \mathbb{R}$ est continue et $\lambda \in \mathbb{R}$, $\{x \in \mathbb{R}^n\,;\, f(x) < \lambda\} = f^{-1}(]-\infty, \lambda[)$ est ouvert, $\{x \in \mathbb{R}^n\,;\, f(x) \leq \lambda\} = f^{-1}(]-\infty, \lambda])$ est fermé
- ⚠ Le slogan (à manier avec précaution) qui va avec l'item précédent est « les inégalités strictes définissent des ouverts, les inégalités larges définissent des fermés »

Esquisse de preuve de la proposition 2.2

(i) Si $x \in \mathbb{R}^n$, $B(x,1) \subset \mathbb{R}^n$. $\mathbb{R}^n$ est donc ouvert. Comme $\emptyset^c = \mathbb{R}^n$, $\emptyset$ est fermé. $\emptyset$ est ouvert (par l'absurde). Comme $(\mathbb{R}^n)^c = \emptyset$, $\mathbb{R}^n$ est fermé. (ii) Si $x \in \cup_{i\in U_i}$, $\exists\, i_0 \in I$ tel que $x \in U_{i_0}$. $\exists\, r > 0$ tel que $B(x,r) \subset U_{i_0}$. Finalement, $B(x,r) \subset \cup_{i\in I} U_i$. (iii) Soit $x \in \cap_{i=1}^N U_i$. Soient $r_i > 0$ tels que $B(x,r_i) \subset U_i$, $\forall\, i$. Si $r := \min_{1\leq i\leq N} r_i$, alors $B(x,r) \subset \cap_{i=1}^N U_i$. (iv) $(\cup_{i=1}^N F_i)^c = \cap_{i=1}^N (F_i)^c$, qui est un ouvert, par (iii). (v) $(\cap_{i\in I} F_i)^c = \cup_{i\in I} (F_i)^c$, qui est un ouvert, par (ii). (vi) « $\Rightarrow$ » Par l'absurde. Sinon, $\exists\, (x^k) \subset F$, $x \in F^c$ tels que $x^k \to x$. Soit $r > 0$ tel que $B(x,r) \subset F^c$. Il existe k_0 tel que $\left\|x^k - x\right\|_2 < r$, $\forall\, k \geq k_0$. ↯ « $\Leftarrow$ » Par l'absurde. Soit $x \in F^c$ tel que $\forall\, r > 0$, $B(x,r) \not\subset F^c$. Pour $r = 1/k$, $k \geq 1$, $\exists\, x^k \in B(x,r) \setminus F^c$. Donc $x^k \in F$ et $\left\|x^k - x\right\|_2 < 1/k$. D'où $(x^k) \subset F$ et $x^k \to x \in F^c$. ↯ (vii) Soit $x \in f^{-1}(V)$. Soit $\varepsilon > 0$ tel que $B(f(x),\varepsilon) \subset V$. Soit $\delta > 0$ tel que $[y \in U, \|y - x\|_2 < \delta] \Rightarrow \|f(y) - f(x)\|_2 < \varepsilon$. Soit $\rho > 0$ tel que $B(x,\rho) \subset U$. Si $r := \min(\delta, \rho)$, alors $B(x,r) \subset f^{-1}(V)$. (viii) Soient $(x^k) \subset f^{-1}(G)$ et $x \in \mathbb{R}^n$ tels que $x^k \to x$. Comme $f(x^k) \in G$ et $f(x^k) \to f(x)$, nous avons $f(x) \in G$, d'où $x \in f^{-1}(G)$. *CQFD*

Le théorème de B. Bolzano (1781–1848) et K. T. W. Weierstrass (1815–1897)

Théorème 2.3 (Théorème de Bolzano-Weierstrass) (i) Une suite bornée de $\mathbb{R}^n$ contient une sous-suite convergente : $(x^k) \subset \mathbb{R}^n$ bornée $\Rightarrow \exists\, \varphi : \mathbb{N} \to \mathbb{N}$ extraction et $\exists\, x \in \mathbb{R}^n$ tels que $x^{\varphi(k)} \to x$. (ii) Si $K \subset \mathbb{R}^n$ est compact et $(x^k) \subset K$, $\exists\, \varphi : \mathbb{N} \to \mathbb{N}$ extraction et $\exists\, x \in K$ tels que $x^{\varphi(k)} \to x$.

Esquisse de preuve du théorème 2.3

(i) Soit $M \in \mathbb{R}$ tel que $\left\|x^k\right\|_2 \leq M$, $\forall\, k$. Par (IS), $\left|x_j^k\right| \leq M$, $\forall\, k$, $\forall\, 1 \leq j \leq n$. En particulier, la suite (x_1^k) est bornée. Il existe donc $\varphi_1 : \mathbb{N} \to \mathbb{N}$ une extraction et $x_1 \in \mathbb{R}$ tels que $x_1^{\varphi_1(k)} \to x_1$. La suite $(x_2^{\varphi_1(k)})$ étant bornée, il existe $\varphi_2 : \mathbb{N} \to \mathbb{N}$ extraction et $x_2 \in \mathbb{R}$ tels que $x_2^{\varphi_1(\varphi_2(k))} \to x_2$. Par récurrence sur j, il existe $\varphi_j : \mathbb{N} \to \mathbb{N}$ une extraction et $x_j \in \mathbb{R}$ tels que $x_j^{\varphi_1(\varphi_2(\ldots(\varphi_j(k))\ldots))} \to x_j$. Si nous posons $\varphi := \varphi_1 \circ \varphi_2 \circ \cdots \circ \varphi_n : \mathbb{N} \to \mathbb{N}$ et $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$, alors φ est une extraction et $x^{\varphi(k)} \to x$.
(ii) K est borné : $\exists\, M \in \mathbb{R}$ tel que $\|x\|_2 \leq M$, $\forall\, x \in K$. En particulier, $\left\|x^k\right\|_2 \leq M$, $\forall\, k$. De (i), il existe φ extraction et $x \in \mathbb{R}^n$ tels que $x^{\varphi(k)} \to x$. K étant fermé, nous avons $x \in K$. *CQFD*

Théorème des bornes atteintes

Théorème 2.4 Soit $K \subset \mathbb{R}^n$ un compact non vide. Si $f : K \to \mathbb{R}$ est continue, alors il existe $\underline{x}, \overline{x} \in K$ tels que $f(\underline{x}) \leq f(x) \leq f(\overline{x})$, $\forall\, x \in K$.

Remarque

Autrement énoncé, une fonction scalaire continue sur un compact atteint son minimum et son maximum.

Preuve du théorème 2.4

Soit $M := \sup_{x\in K} f(x) \in]-\infty, \infty]$. Il existe $(x^k) \subset K$ telle que $f(x^k) \to M$. Soient φ une extraction et $\overline{x} \in K$ tels que $x^{\varphi(k)} \to \overline{x}$. f étant continue, nous avons $f(\overline{x}) = M$. En particulier, M est fini et $\overline{x}$ est un point de maximum de f. L'existence de $\underline{x}$ est similaire. *CQFD*

Remarques

— Le théorème 2.4 a une « réciproque » : si $K \subset \mathbb{R}^n$ est tel que le théorème des bornes atteintes soit satisfait par toute fonction continue $f : K \to \mathbb{R}$, alors K est compact (exercice TD1.18)
— L'étude des ouverts, fermés, compacts et autres ensembles remarquables fait l'objet de la topologie générale, qui sera abordée en L3

Résultats admis (qui seront prouvés en L3)

Théorème 2.5 (Théorème de Heine) Si $K \subset \mathbb{R}^n$ est compact, une fonction continue $f : K \to \mathbb{R}^m$ est uniformément continue.

Proposition 2.6 Les parties de $\mathbb{R}^n$ à la fois ouvertes et fermées sont $\emptyset$ et $\mathbb{R}^n$. (Exercice TD2.19.)

Quiz

Soient $A \subset \mathbb{R}^n$ un ensemble non borné et $f : A \to \mathbb{R}$. Rappeler le sens de $\lim_{\|x\|_2\to\infty} f(x) = \infty$.

Fonction coercive

Soit $A \subset \mathbb{R}^n$ non borné. Une fonction $f : A \to \mathbb{R}$ est coercive si $\lim_{\|x\|_2\to\infty} f(x) = \infty$.

Théorème 2.7 Une fonction continue coercive sur un ensemble fermé non borné atteint son minimum.

Esquisse de preuve du théorème 2.7

Soit $a \in F$. Il existe $R > 0$ tel que $[x \in F, \|x\|_2 > R] \Rightarrow f(x) > f(a)$. Soit $K := F \cap \overline{B}(0, R)$, qui est fermé, borné et non vide ($a \in K$). Soit $\underline{x}$ un point de minimum de f sur K. Alors $\underline{x}$ est un point de minimum de f sur F. *CQFD*

À réviser pour la troisième leçon

Propriétés des fonctions continues réelles (continuité de la somme, du produit, des composées), exemples typiques de fonctions continues, et les propriétés des limites.

Leçon 3. Continuité. Limites

Objectifs

Savoir montrer la continuité de fonctions données par des expressions explicites. Utiliser la continuité des fonctions pour montrer des propriétés d'ensembles définis par des inégalités. S'approprier quelques techniques élémentaires d'étude de limites.

Exemple fondamental de fonction continue de plusieurs variables

Définition 3.1 Une fonction monomiale est une fonction de la forme $\mathbb{R}^n \ni x \mapsto x^\alpha := (x_1)^{\alpha_1}(x_2)^{\alpha_2}\cdots(x_n)^{\alpha_n}$, où $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$. Exemple : $(x, y, z)^{(1,2,4)} = xy^2z^4$.

⚠ Par abus de langage, la fonction $x \mapsto x^\alpha$ sera désignée comme un monôme, plutôt que comme une fonction monomiale.

Proposition 3.2 Un monôme est continu.

Esquisse de preuve de la proposition 3.2

Soient $(x^k) \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$ tels que $x^k \to x$. Alors $x_j^k \to x_j$, $\forall\, 1 \leq j \leq n$, d'où $(x_j^k)^{\alpha_j} \to (x_j)^{\alpha_j}$, $\forall\, 1 \leq j \leq n$, d'où $(x^k)^\alpha \to x^\alpha$. *CQFD*

Opérations qui préservent la continuité

Proposition 3.3 Avec $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$, f, g, h, Φ continues : (i) $[f, g : A \to \mathbb{R}^m] \Rightarrow f + g$ continue ; (ii) $[f : A \to \mathbb{R}, g : A \to \mathbb{R}^m] \Rightarrow fg$ continue ; (iii) $[h : A \to B, \Phi : B \to \mathbb{R}^p] \Rightarrow \Phi \circ h$ continue.

Quiz : opérations avec les fonctions

(i) $[f : A \to \mathbb{R}^*,\ f$ continue$] \Rightarrow 1/f$ continue ? (ii) $[f, g : A \to \mathbb{R}^m,\ f, g$ continues$] \Rightarrow fg$ continue ?

Esquisse de preuve de la proposition 3.3

(iii) est évident en utilisant la caractérisation de la continuité avec des suites. (i) Appliquer (iii) avec $p := 2m$, $\Phi(u, v) := u + v$, $\forall\, (u, v) \in \mathbb{R}^m$, $h(x) := (f(x), g(x))$, $\forall\, x \in A$. (ii) Appliquer (iii) avec $p := m + 1$, $\Phi(t, u) := tu$, $\forall\, (t, u) \in \mathbb{R} \times \mathbb{R}^m$, $h(x) := (f(x), g(x))$, $\forall\, x \in A$. *CQFD*

Travail individuel

Exercice CM3.4 Compléter la preuve de la proposition 3.3.

Exercice travaillé

Exercice CM3.5 Soient $A \subset \mathbb{R}^n$, $a \in \mathbb{R}^n$ point d'accumulation de A, $f : A \to \mathbb{R}^m$. Montrer que :

- Si $\ell \in \mathbb{R}^m$, $\lim_{x \to a} f(x) = \ell \Leftrightarrow [\forall\, (x^k) \subset A \setminus \{a\},\ x^k \to x \Rightarrow f(x^k) \to \ell]$
- Si $m = 1$, $\lim_{x \to a} f(x) = \infty \Leftrightarrow [\forall\, (x^k) \subset A \setminus \{a\},\ x^k \to x \Rightarrow f(x^k) \to \infty]$

Esquisse de solution de l'exercice CM3.5

- « $\Rightarrow$ » Soient $\varepsilon > 0$ et le δ correspondant. Soit k_0 tel que $\forall\, k \geq k_0$, $\left\|x^k - a\right\|_2 < \delta$. Pour $k \geq k_0$, nous avons $\left\|f(x^k) - \ell\right\|_2 < \varepsilon$. « $\Leftarrow$ » Par l'absurde. Soit $\varepsilon > 0$ tel que $\forall\, \delta > 0$, $\exists\, x \in A \setminus \{a\}$ tel que $\|x - a\|_2 < \delta$ et $\|f(x) - \ell\|_2 \geq \varepsilon$. En particulier, $\forall\, k \geq 1$, $\exists\, x^k \in A \setminus \{a\}$ tel que $\left\|x^k - a\right\|_2 < 1/k$ et $\left\|f(x^k) - \ell\right\|_2 \geq \varepsilon$. Comme $(x^k) \subset A \setminus \{a\}$ et $x^k \to x$, $\exists\, k_0$ tel que $k \geq k_0 \Rightarrow \left\|f(x^k) - \ell\right\|_2 < \varepsilon$. ↯
- « $\Rightarrow$ » Soient $M \in \mathbb{R}$ et le δ correspondant. Soit k_0 tel que $\forall\, k \geq k_0$, $\left\|x^k - a\right\|_2 < \delta$. Pour $k \geq k_0$, nous avons $f(x^k) > M$. « $\Leftarrow$ » Par l'absurde. Soit $M \in \mathbb{R}$ tel que $\forall\, \delta > 0$, $\exists\, x \in A \setminus \{a\}$ tel que $\|x - a\|_2 < \delta$ et $f(x) \leq M$. En particulier, $\forall\, k \geq 1$, $\exists\, x^k \in A \setminus \{a\}$ tel que $\left\|x^k - a\right\|_2 < 1/k$ et $f(x^k) \leq M$. Comme $(x^k) \subset A \setminus \{a\}$ et $x^k \to x$, $\exists\, k_0$ tel que $k \geq k_0 \Rightarrow f(x^k) > M$. ↯ *CQFD*

Remarque

L'implication « $\Leftarrow$ » est démontrée le plus souvent par l'absurde. Une preuve voisine repose sur la contraposée. Pour une approche différente, voir l'exercice TD2.18.

Quiz

- Pour quelle raison a est-il supposé point d'accumulation de A ?
- Que devient la condition « a est point d'accumulation de A » dans l'étude de $\lim_{\|x\|_2 \to \infty} f(x)$?

Quelques stratégies simples pour étudier des limites en 0

Proposition 3.6 Soit $A \subset \mathbb{R}^n$ tel que 0 soit un point d'accumulation de A. Soit $f : A \to \mathbb{R}^m$.

- S'il existe $R > 0$ et $g :]0, R] \to [0, \infty[$ tels que : (i) $[x \in A \setminus \{0\},\ \|x\|_2 \leq R] \Rightarrow \|f(x)\|_2 \leq g(\|x\|_2)$; (ii)

$\lim_{r\to 0} g(r) = 0$, alors $\lim_{x\to 0} f(x) = 0$
— S'il existe $\omega, \sigma \in \mathbb{R}^n \setminus \{0\}$ et $R > 0$ tels que : (i) $0 < t < R \Rightarrow t\omega, t\sigma \in A$; (ii) $\lim_{t\to 0} f(t\omega) \neq \lim_{t\to 0} f(t\sigma)$ (ou l'une des deux limites n'existe pas), alors $\lim_{x\to 0} f(x)$ n'existe pas

Esquisse de preuve de la proposition 3.6

— Soient $\varepsilon > 0$ et le δ correspondant (relatif à g). On peut supposer $\delta \leq R$. Si $x \in A \setminus \{0\}$ et $\|x\|_2 \leq \delta$, alors $\|f(x)\|_2 \leq g(\|x\|_2) < \varepsilon$
— Par l'absurde. Soit $\ell := \lim_{x\to 0} f(x)$. Si $\lim_{t\to 0} f(t\omega)$ n'existe pas ou est différente de ℓ, alors il existe une suite $(t^k) \subset]0, R[$ telle que $t^k \to 0$ et $f(t^k\omega) \not\to \ell$. Or, $(t^k\omega) \subset A \setminus \{0\}$ et $t^k\omega \to 0$, d'où $f(t^k\omega) \to \ell$ ↯ De même pour $\lim_{t\to 0} f(t\sigma)$ *CQFD*

Travail individuel

Exercice CM3.7 Soit $A \subset \mathbb{R}^n$ non borné. Soit $f : A \to \mathbb{R}^m$.
— S'il existe $R > 0$ et $g : [R, \infty[\to [0, \infty[$ tels que : (i) $[x \in A,\ \|x\|_2 \geq R] \Rightarrow \|f(x)\|_2 \leq g(\|x\|_2)$; (ii) $\lim_{r\to\infty} g(r) = 0$, alors $\lim_{\|x\|_2\to\infty} f(x) = 0$
— S'il existe $\omega, \sigma \in \mathbb{R}^n \setminus \{0\}$ et $R > 0$ tels que : (i) $t > R \Rightarrow t\omega, t\sigma \in A$; (ii) $\lim_{t\to\infty} f(t\omega) \neq \lim_{t\to\infty} f(t\sigma)$ (ou l'une des deux limites n'existe pas), alors $\lim_{\|x\|_2\to\infty} f(x)$ n'existe pas

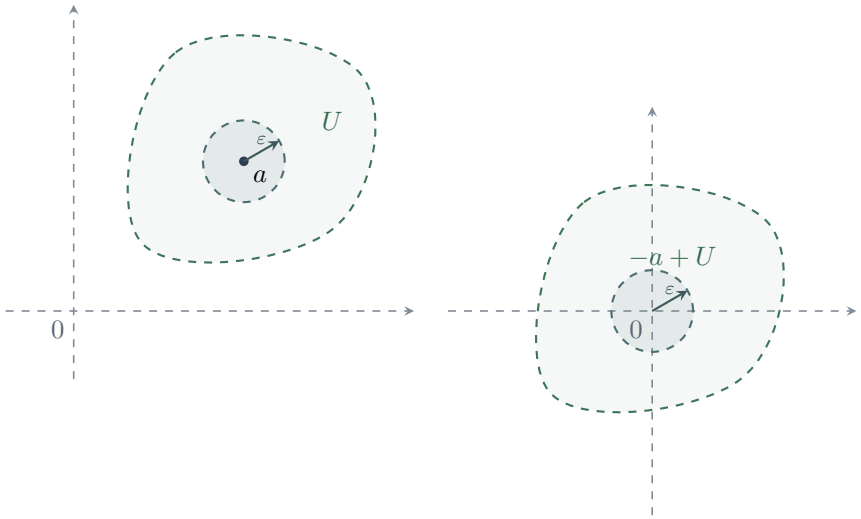
Cadre de travail pour le calcul différentiel

— $U \subset \mathbb{R}^n$ est un ouvert, $a \in U$, $g : U \setminus \{a\} \to \mathbb{R}^m$
— On fait le « changement de variable » $x = a + h$, avec $x \in U$, $h \in -a + U := \{y = -a + u\,;\ u \in U\}$

Proposition 3.8 Sous les hypothèses ci-dessus
— $-a + U$ est un ouvert contenant 0, et en particulier 0 est un point d'accumulation de $-a + U$
— $\lim_{x\to a} g(x) = \lim_{h\to 0} g(a + h)$

Esquisse de preuve de la proposition 3.8

— Soient $b \in U$ et $\varepsilon > 0$ tels que $B(b, \varepsilon) \subset U$. Alors $B(b - a, \varepsilon) \subset -a + U$. Pour la deuxième partie, voir l'exercice TD2.8
— Avec $h^k := -a + x^k$, nous avons $x^k \in U \setminus \{a\} \Leftrightarrow h^k \in (-a + U) \setminus \{0\}$, $x^k \to a \Leftrightarrow h^k \to 0$, et $g(x^k) = g(a + h^k)$. Donc, si $\ell \in \mathbb{R}^m$ (ou si $m = 1$ et $\ell = \infty$), $\big[\forall\, (x^k) \subset U \setminus \{a\}, [x^k \to a \Rightarrow g(x^k) \to \ell]\big] \Leftrightarrow \big[[\forall\, (h^k) \subset (-a + U) \setminus \{0\}, [h^k \to 0 \Rightarrow g(a + h^k) \to \ell]\big]$ *CQFD*



Si l'ouvert U contient a, son translaté $-a + U$ est un ouvert contenant 0

Notations de Landau (du nom de E. Landau (1877–1938))

— $f(h) = o(g(h))$ et $f(h) = O(g(h))$ s'entendent quand $h \to 0$
— Cadre : $A \subset \mathbb{R}^n$, 0 point d'accumulation de A, $f : A \setminus \{0\} \to \mathbb{R}^m$, $g : A \setminus \{0\} \to \mathbb{R}^p$ (souvent $p = 1$)
— $f(h) = o(g(h))$ si : $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ tel que $[h \in A \setminus \{0\},\ \|h\|_2 \leq \delta] \Rightarrow \|f(h)\|_2 \leq \varepsilon \|g(h)\|_2$
— $f(h) = O(g(h))$ si : $\exists\, R > 0$, $\exists\, C \in \mathbb{R}$ tels que $[h \in A \setminus \{0\},\ \|h\|_2 \leq R] \Rightarrow \|f(h)\|_2 \leq C\|g(h)\|_2$

En pratique (I)

— $0 \in A$, $f : A \to \mathbb{R}^m$, $g : A \to \mathbb{R}^p$, $f(0) = 0$, $g(0) = 0$. Dans ce cas, $f(h) = o(g(h)) \Leftrightarrow \forall\, \varepsilon > 0$, $\exists\, \delta > 0$ tel que $[h \in A,\ \|h\|_2 \leq \delta] \Rightarrow \|f(h)\|_2 \leq \varepsilon \|g(h)\|_2$
— $g(h) \neq 0$ si $h \neq 0$. Dans ce cas,

$$f(h) = o(g(h)) \Leftrightarrow \lim_{h \to 0} \frac{\|f(h)\|_2}{\|g(h)\|_2} = 0 \Leftrightarrow \lim_{h \to 0} \frac{f(h)}{\|g(h)\|_2} = 0$$

— Si, de plus, $p = 1$, $f(h) = o(g(h)) \Leftrightarrow \lim_{h \to 0} \frac{f(h)}{g(h)} = 0$

En pratique (II)

— $f(h) = g(h) + o(r(h))$ signifie que $(f - g)(h) = o(r(h))$
— « Principe $r(\varepsilon)$ ». Dans la preuve de $f(h) = o(g(h))$, ce sont plutôt des majorations de la forme $\|f(h)\|_2 \leq r(\varepsilon)\|g(h)\|_2$ qui ressortent naturellement. Si $\lim_{\varepsilon \to 0} r(\varepsilon) = 0$, cela suffit. En effet : $\exists\, \delta_1 > 0$ tel que $0 < \varepsilon_1 \leq \delta_1 \Rightarrow |r(\varepsilon_1)| \leq \varepsilon$, $\exists\, \delta > 0$ tel que $[h \in A \setminus \{0\},\ \|h\|_2 < \delta] \Rightarrow \|f(h)\|_2 \leq r(\delta_1)\|g(h)\|_2$. D'où $[h \in A \setminus \{0\},\ \|h\|_2 < \delta] \Rightarrow \|f(h)\|_2 \leq \varepsilon \|g(h)\|_2$

Les applications linéaires sont $O(\|h\|_2)$

Proposition 3.9 Si $T : \mathbb{R}^n \to \mathbb{R}^m$ est linéaire, il existe $M \in \mathbb{R}$ tel que $\|T(h)\|_2 \leq M\|h\|_2$, $\forall\, h \in \mathbb{R}^n$.

Esquisse de preuve de la proposition 3.9

Avec $(e^1, \ldots, e^n)$ la base canonique de $\mathbb{R}^n$ et $M_1 := \max\left\{\left\|T(e^j)\right\|_2 ;\ 1 \leq j \leq n\right\}$, (IS) donne

$$\|T(h)\|_2 = \left\| \sum_{j=1}^{n} h_j T(e^j) \right\|_2 \leq \sum_{j=1}^{n} |h_j| \left\|T(e^j)\right\|_2 \leq n M_1 \|h\|_2. \qquad \textit{CQFD}$$

Intuition derrière o et O, et quelques règles de calcul

— $f(h) = o(g(h))$ quantifie le fait que « f est négligeable devant g quand $h \to 0$ »
— $f(h) = O(g(h))$ quantifie le fait que « f est au plus de l'ordre de grandeur de g quand $h \to 0$ »

Quelques règles de calcul, qui ne seront ni utilisées, ni montrées dans ce cours, mais qui servent ailleurs

— $o + o = o$: $[f_1(h) = o(g(h)),\ f_2(h) = o(g(h))] \Rightarrow (f_1 + f_2)(h) = o(g(h))$
— $o = O$: $f(h) = o(g(h)) \Rightarrow f(h) = O(g(h))$
— $O + O = O$; $o + O = O$; $O(1) \cdot o(g(h)) = o(g(h))$; $o(O(h)) = o(h)$; $O(o(h)) = o(h)$

À réviser pour la quatrième leçon

La dérivée des fonction réelles, et les principaux résultats : interprétation géométrique de la dérivée, théorème des accroissements finis, inégalité des accroissements finis, formules de Taylor à l'ordre 1 (avec reste « en o », avec point intermédiaire, avec reste intégral), la règle de la chaîne. Preuve de « f dérivable$\Rightarrow$ f continue ». Fonctions de classe C^1. Le théorème fondamental du calcul intégral/de l'analyse (formule de Newton-Leibniz).

Leçon 4. La différentielle

La difficulté majeure de cette UE...

...consiste à appréhender la notion de différentielle et à manier cet objet.

Objectifs

Comprendre la notion de différentielle. Avoir une table de correspondances entre les propriétés de la dérivée et celles de la différentielle. Comprendre les limites de l'analogie entre les deux objets. Maîtriser le schéma de transfert dérivée $\rightsquigarrow$ différentielle. ⚠ Maîtriser la règle de la chaîne Comprendre la nature des objets d'encodage (gradient, matrice jacobienne) et savoir les calculer.

La formule de la dérivée qui « passe à plusieurs dimensions de l'espace »

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \Leftrightarrow \lim_{h\to 0} \frac{f(a+h)-f(a)-f'(a)h}{h} = 0$$

$$\Leftrightarrow f(a+h)-f(a)-f'(a)h = o(h) \Leftrightarrow \boxed{f(a+h) = f(a)+f'(a)h+o(|h|)}$$ Taylor à l'ordre 1 !

Clés pour passer à plusieurs dimensions

— $\mathbb{R} \ni h \mapsto L(h) \coloneqq f'(a)h \in \mathbb{R}$ est une application linéaire
— Si f est dérivable en a : $f(a+h) = f(a)+L(h)+o(|h|)$
— $f'(a)$ est le nombre qui « encode » $L(h)$: $L(h) = \alpha h$, avec $\alpha = f'(a)$

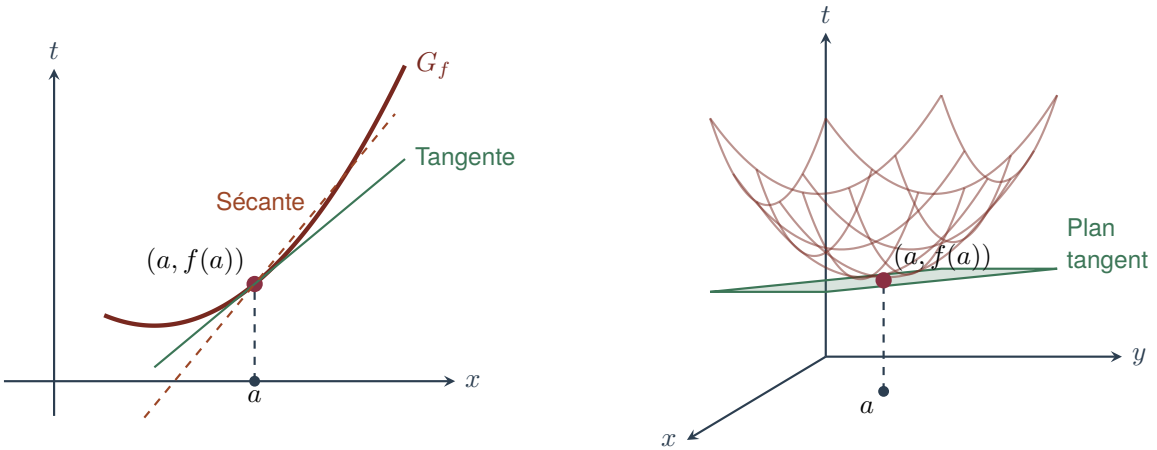
Cadre pour le « calcul différentiel » dans $\mathbb{R}^n$

(i) $U \subset \mathbb{R}^n$ ouvert ; (ii) $a \in U$; (iii) $f : U \to \mathbb{R}^m$ (ou $f : U \to V$, avec $V \subset \mathbb{R}^m$ ouvert) ; (iv) $h \in -a+U$ (ouvert contenant 0) : pour donner un sens à $f(a+h)$; (v) $h \in \mathbb{R}^n$: pour donner un sens à $L(h)$

⚠ **Ce cadre est implicite dans les définitions et résultats de cette leçon**

Définition 4.1 La différentielle de f en a (si elle existe) est l'application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ telle que $\boxed{f(a+h) = f(a)+L(h)+o(\|h\|_2)}$ ⚠ Si la différentielle existe, f est diférentiable en a

— ⚠ La différentielle est donc une application linéaire (pas un nombre !)
— Elle est notée $\boxed{df_a}$, d_af, $df(a)$, $Df(a)$,..., selon les disciplines



Interprétation géométrique de la différentielle. La droite tangente $t = f(a)+f'(a)h$ est la droite qui approche le mieux le graphe de f lorsque $h \coloneqq x-a \approx 0$. Il en va de même pour le plan tangent $t = f(a)+df_a(h)$

Interprétation analytique de la différentielle

$\boxed{f(a+h) \approx f(a) + df_a(h)}$ – « si h est petit, $df_a(h)$ est une bonne approximation de l'accroissement de f entre a et $a+h$ »

Premières propriétés de la différentielle

Proposition 4.2 — La différentielle, si elle existe, est unique

— Si $n = 1$, f différentiable en $a \Leftrightarrow f$ dérivable en a. De plus, $\boxed{df_a(h) = f'(a)h}$ et $\boxed{f'(a) = df_a(1)}$

Esquisse de preuve de la proposition 4.2

— Soient $\omega \in \mathbb{R}^n \setminus \{0\}$, L_1, L_2 vérifiant la définition de la différentielle en a, et $L \coloneqq L_1 - L_2$. Alors $L(h) = o(\|h\|_2)$. En testant cette propriété avec $h \coloneqq t\omega$, avec $t > 0$, nous obtenons $L(\omega) = 0$

— « $\Rightarrow$ » Si $L(h) = \alpha h$, alors $\alpha = f'(a)$; en particulier f est dérivable. « $\Leftarrow$ » Si f est dérivable, $L(h) \coloneqq f'(a)h$ satisfait la définition de la différentielle *CQFD*

Vecteurs lignes ou colonnes

— En analyse $\rightsquigarrow$ vecteurs lignes : $f(1,2) = (2,4,5)$, $f(2,4,1) = 5\ldots$

— En algèbre linéaire $\rightsquigarrow$ matrices, vecteurs colonnes : $\begin{pmatrix}5\\7\end{pmatrix} = \begin{pmatrix}1 & 2\\-1 & 4\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\ldots$

— En Analyse IV : (i) h est ligne dans l'argument d'une fonction : $f((0,1) + (1,2))$; (ii) dans les formules algébriques, la valeur d'une fonction est colonne : $f((0,1) + (1,2)) = \begin{pmatrix}1\\4\end{pmatrix}$; (iii) dans les formules algébriques, h est colonne : $\begin{pmatrix}-1\\4\end{pmatrix} \cdot \begin{pmatrix}1\\2\end{pmatrix}$ ou $\begin{pmatrix}5\\7\end{pmatrix} = \begin{pmatrix}1 & 2\\-1 & 4\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}$

— ⚠ Si $L : \mathbb{R}^n \to \mathbb{R}$, $\exists\, v$ colonne tel que $L(h) = v \cdot h = Ah$, $\forall\, h \in \mathbb{R}^n$ (h colonne), où $A \coloneqq v^\intercal$ ligne

Gradient, (matrice) jacobienne (du nom de J. S./C. G. J. Jacobi (1804–1851))

— Si $f : U \to \mathbb{R}$ est différentiable en a, le vecteur colonne $v \in \mathbb{R}^n$ tel que $df_a(h) = v \cdot h$, $\forall\, h \in \mathbb{R}^n$, est le gradient de f en a, noté $\nabla f(a)$ (∇ = « nabla »)

— Si $f : U \to \mathbb{R}^m$ est différentiable en a, la matrice $A \in M_{m,n}(\mathbb{R})$ telle que $df_a(h) = Ah$, $\forall\, h \in \mathbb{R}^n$, est la matrice jacobienne de f en a, notée $J_f(a)$

Le calcul différentiel en Analyse IV se réduit à un calcul de nombres

— ⚠ Même si conceptuellement dérivée $\neq$ différentielle (nombre vs application linéaire), pour déterminer la différentielle il suffit de calculer les n éléments de $\nabla f(a)$ ou les $m \times n$ éléments de $J_f(a)$, donc un nombre fini de nombres

— En L3 et M1, $\mathbb{R}^n$ sera remplacé par un espace possiblement de dimension infinie : dans ce cadre, le calcul de la différentielle ne se résume pas à un calcul de nombres

Fonctions partielles et dérivées partielles

— Si on « fige » les coordonnées $a_2, \ldots, a_n$ de a, l'ensemble $I_1 \coloneqq \{x_1 \in \mathbb{R}\,;\, (x_1, a_2, \ldots, a_n) \in U\}$ est un ouvert (image réciproque d'ouvert par une fonction continue) qui contient a_1

— Si la « fonction partielle » $f(\cdot, a_2, \ldots, a_n)$, $I_1 \ni x_1 \mapsto f(x_1, a_2, \ldots, a_n) \in \mathbb{R}^m$ a une dérivée en x_1, celle-ci est notée $\boxed{\dfrac{\partial f}{\partial x_1}(x_1, a_2, \ldots, a_n)}$ $\boxed{\partial_{x_1} f(x_1, a_2, \ldots, a_n)}$ $\boxed{\partial_1 f(x_1, a_2, \ldots, a_n)}$ « la dérivée partielle de f par rapport à x_1 en $(x_1, a_2, \ldots, a_n)$ » ⚠ Dans la jacobienne $\rightsquigarrow$ vecteur colonne

— On définit de même les fonctions et les dérivées partielles par rapport aux coordonnées $x_2, \ldots, x_n$

Quiz : dérivées partielles

— Pour $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) \coloneqq xe^y$, $\forall\, (x,y) \in \mathbb{R}^2$: (i) $\partial_x f(1,0) =$? (ii) $\partial_y f(0,1) =$?

— Pour $f : \mathbb{R} \to \mathbb{R}^2$, $f(x) := (x, e^x) \sim \begin{pmatrix} x \\ e^x \end{pmatrix}$, $\forall\, x \in \mathbb{R} : \partial_x f(1) = ?$

Comment déterminer $\nabla f(a)$ et $J_f(a)$

Proposition 4.3 — Si $f : U \to \mathbb{R}$ est différentiable en a, $\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \partial_2 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix}$

— Si $f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix} : U \to \mathbb{R}^m$ est différentiable en a, $J_f(a) = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \cdots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \cdots & \partial_n f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \cdots & \partial_n f_m(a) \end{pmatrix}$

Remarques

— Pour les fonctions scalaires ($m = 1$), $J_f(a) = [\nabla f(a)]^\intercal$

— Dans le cas général, $J_f(a) = \begin{pmatrix} [\nabla f_1(a)]^\intercal \\ [\nabla f_2(a)]^\intercal \\ \vdots \\ [\nabla f_m(a)]^\intercal \end{pmatrix} = \begin{pmatrix} \partial_1 f(a) & \partial_2 f(a) & \dots & \partial_n f(a) \end{pmatrix}$ (f colonne)

Esquisse de preuve de la proposition 4.3

La première formule est un cas particulier de la seconde. Pour la seconde : soit $(t^k) \subset \mathbb{R}^*$ telle que $a + t^k e^1 \in I_1$ et $t^k \to 0$. Alors $\lim\limits_{k\to\infty} \dfrac{f(a + t^k e^1) - f(a) - df_a(t^k e^1)}{|t^k|} = 0$, d'où $\exists\, \partial_1 f(a)$ et $\partial_1 f(a) = df_a(e^1)$.
De même, $\partial_j f(a) = df_a(e^j)$, $j = 2, \dots, n$. *CQFD*

Une condition nécessaire pour qu'une fonction soit différentiable

Corollaire 4.4 Une fonction différentiable admet des dérivées partielles.

Il suffit de s'intéresser aux coordonnées de f

Proposition 4.5 $f : U \to \mathbb{R}^m$ est différentiable en $a \Leftrightarrow f_\ell$ est différentiable en a, $\forall\, 1 \leq \ell \leq m$.

Esquisse de preuve de la proposition 4.5

« $\Rightarrow$ » $|f_\ell(a + h) - f(a) - [\nabla f_\ell(a)] \cdot h| \leq \|f(a + h) - f(a) - [J_f(a)]h\|_2 = o(\|h\|_2)$.
« $\Leftarrow$ » $\|f(a + h) - f(a) - [J_f(a)]h\|_2 \leq \sum_{\ell=1}^m |f_\ell(a + h) - f(a) - [\nabla f_\ell(a)] \cdot h| = o(\|h\|_2)$. *CQFD*

Une condition suffisante pour qu'une fonction soit différentiable

Proposition 4.6 Si f a des dérivées partielles **continues** en a, alors f est différentiable en a.

Définition 4.7 $f : U \to \mathbb{R}^m$ est de classe C^1 si f admet des dérivées partielles continues (en tout point).
⚠ $\boxed{f \in C^1}$ ou $\boxed{f \in C^1(U)}$ ou $\boxed{f \in C^1(U, \mathbb{R}^m)}$

Corollaire 4.8 Une fonction de classe C^1 est différentiable.

⚠ La façon « usuelle » de montrer qu'une fonction concrète est différentiable est de vérifier que ses dérivées partielles existent et sont continues.

Esquisse de preuve de la proposition 4.6 si $n = 2$

Il suffit de prendre $m = 1$. Soit $r > 0$ tel que $B(a, r) \subset U$. Soit $\varepsilon > 0$. Soit $\lambda > 0$ tel que $[x \in U,\ \|x - a\|_2 < \lambda] \Rightarrow [|\partial_j f(x) - \partial_j f(a)| < \varepsilon,\ j = 1, 2]$. Soit $\delta := \min(r, \lambda)$. Si $\|h\|_2 < \delta$, alors le TAF donne $\xi_1 \in [0, h_1]$ et $\xi_2 \in [0, h_2]$ tels que

$$\begin{aligned} f(a+h) - f(a) =& f(a+h) - f(a_1, a_2 + h_2) + f(a_1, a_2 + h_2) - f(a) \\ =& h_1 \partial_1 f(a_1 + \xi_1, a_2 + h_2) + h_2 \partial_2 f(a_1, a_2 + \xi_2), \end{aligned}$$

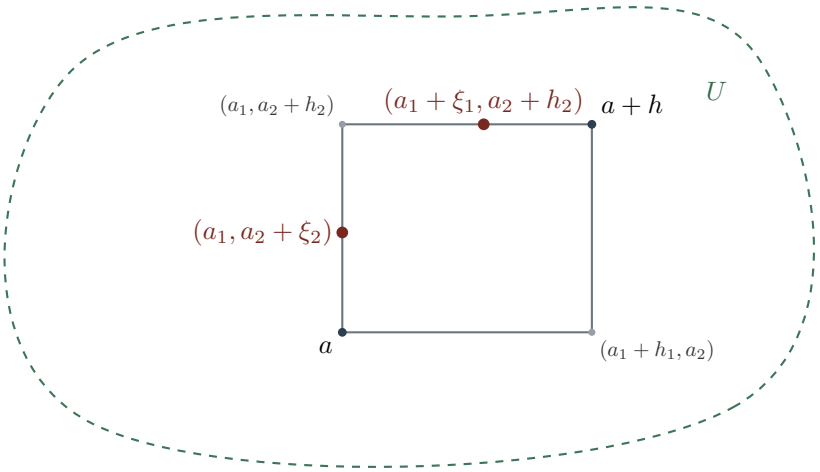
d'où

$$\begin{aligned} |f(a+h) - f(a) - [\nabla f(a)] \cdot h| \leq& |h_1| |\partial_1 f(a_1 + \xi_1, a_2 + h_2) - \partial_1 f(a)| \\ &+ |h_2| |\partial_2 f(a_1, a_2 + \xi_2) - \partial_2 f(a)| \leq \varepsilon(|h_1| + |h_2|) \leq 2\varepsilon \|h\|_2. \quad \textit{CQFD} \end{aligned}$$

Travail individuel

Exercice CM4.9 Prouver la proposition 4.6 pour $n \geq 3$.

Exercice CM4.10 ⚠ En considérant par exemple $f : \mathbb{R} \to \mathbb{R}^2$, $f(x) := (x^2, x^3)$, $\forall\, x \in \mathbb{R}$, montrer que le théorème des accroissements finis usuel n'est pas valide pour les fonctions vectorielles.



$(a_1 + \xi_1, a_2 + h_2)$ et $(a_1, a_2 + \xi_2)$ sont donnés par le théorème des accroissements finis

La règle de la chaîne

Théorème 4.11 (Règle de la chaîne) $[f : U \to V$ différentiable en $a \in U$, $g : V \to \mathbb{R}^p$ différentiable en $b := f(a) \in V] \Rightarrow g \circ f$ différentiable en a et $\boxed{d(g \circ f)_a(h) = dg_b(df_a(h)),\ \forall\, h \in \mathbb{R}^n}$

$$\begin{array}{ccccc} \mathbb{R}^n & & \mathbb{R}^m & & \\ \cup & & \cup & & \\ U & \xrightarrow{f} & V & \xrightarrow{g} & \mathbb{R}^p \\ \cup & & \cup & & \cup \\ a & \xrightarrow{f} & b := f(a) & \xrightarrow{g} & c := g(b) \end{array} \qquad \begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{df_a} & \mathbb{R}^m & \xrightarrow{dg_b} & \mathbb{R}^p \\ & & & & \\ \cup & & \cup & & \cup \\ h & \xrightarrow{df_a} & df_a(h) & \xrightarrow{dg_b} & dg_b(df_a(h)) \end{array}$$

À gauche, le diagramme des fonctions f et g. À droite, le diagramme de leurs différentielles

L'ubiquité de la règle de la chaîne

La plupart des règles de calcul (dérivées, dérivées partielles ou différentielles de $f + g$, fg, f/g, λf, f^a, e^f, etc.) sont des cas particuliers de la règle de la chaîne. Voir par exemple les exercices CM4.12, TD3.7, TD3.8 et TD3.10.

Calculs à la physicienne : essentiels pour comprendre, mais pas une preuve

« [...] Le physicien considère comme nulle (et non avenue) toute “perturbation” dont les effets se situent en-deçà de ses capacités observationnelles [...] » C. Aslangul, Mathématiques pour les physiciens
La différentielle et certaines de ses propriétés comme la règle de la chaîne peuvent se comprendre au travers des intuitions et calculs qui, tout en restant vagues, donnent les bons résultats. Des exemples :

- Les fonctions réelles usuelles se développent en série : $f(a+h) = f(a) + \alpha h + \beta h^2 + \cdots$: on trouve $f'(a)$ en négligeant les termes d'ordre ≥ 2. Les fonctions usuelles de plusieurs variables se développent sous la forme $f(a+h) = f(a)+$ termes du premier ordre + termes du second ordre $+\cdots$ Les termes du premier ordre sont de la forme $\alpha_j h_j$. Ceux du second ordre de la forme $\beta_{j,\ell} h_j h_\ell$. Etc. La différentielle $\alpha_1 h_1 + \cdots + \alpha_n h_n$ s'obtient en négligeant les termes d'ordre ≥ 2
- Les termes de second ordre et au-delà sont petits devant les termes du premier ordre quand h est petit. Pour identifier la dérivée ou la différentielle, on néglige les termes petits par rapport à h, et on identifie la partie linéaire du développement
- Exemple. Avec $f(x,y) := e^{x+\sin y}$, $f(h,k) = 1 + (h+\sin k) + (1/2)(h+\sin k)^2 + \cdots \approx 1 + h + k + (1/2)(h+k)^2 + \cdots \approx 1 + h + k$. D'où $df_{(0,0)}(h,k) = h + k$
- Règle de la chaîne : $g(f(a+h)) = g(f(a) + df_a(h) + o(h)) \approx g(f(a) + df_a(h)) = g(f(a)) + dg_{f(a)}(df_a(h)) + o(df_a(h)) \approx g(f(a)) + dg_{f(a)}(df_a(h))$. D'où $d_a(g \circ f)(h) = dg_{f(a)}(df_a(h))$

Un premier exemple d'application de la règle de la chaîne

Exercice CM4.12 $f, g : U \to \mathbb{R}$ différentiables en $a \Rightarrow fg$ différentiable en a et (règle du produit) $\boxed{d(fg)_a(h) = df_a(h)g(a) + f(a)dg_a(h), \forall\, h \in \mathbb{R}^n}$ De même si $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}^m$.

Esquisse de solution de l'exercice CM4.12

Soient $\Phi : \mathbb{R}^2 \to \mathbb{R}$, $\Phi(s,t) := st$, $\forall\,(s,t) \in \mathbb{R}^2$, et $F : U \to \mathbb{R}^2$, $F(x) := (f(x), g(x))$, $\forall\, x \in U$. Alors : (i) $\Phi \in C^1$ et $d\Phi_{(s,t)}(\sigma,\tau) = t\sigma + s\tau$, $\forall\,(s,t),(\sigma,\tau) \in \mathbb{R}^2$; (ii) F est différentiable et $dF_a(h) = (df_a(h), dg_a(h))$, $\forall\, a \in U$, $\forall\, h \in \mathbb{R}^n$; (iii) $fg = \Phi \circ F$. En appliquant la règle de la chaîne,

$$d(fg)_a(h) = d\Phi_{(f(a),g(a))}(df_a(h), dg_a(h)) = g(a)df_a(h) + f(a)dg_a(h). \qquad \textit{CQFD}$$

Règle de la chaîne pour les dérivées partielles

Proposition 4.13 Sous les hypothèses de la règle de la chaîne, $\boxed{\dfrac{\partial(g\circ f)}{\partial x_j}(a) = \sum_{k=1}^{m} \dfrac{\partial f_k}{\partial x_j}(a)\dfrac{\partial g}{\partial y_k}(b)}$

$\forall\, 1 \leq j \leq n$ (avec $b = f(a)$). ⚠ À la physicienne : $\boxed{\partial_{x_j} g(y) = \partial_{x_j} y_k\, \partial_{y_k} g}$, où $y := f(x)$

Esquisse de preuve de la proposition 4.13

$$\frac{\partial(g\circ f)}{\partial x_j}(a) = d(g\circ f)_a(e^j) = dg_b(df_a(e^j)) = dg_b\left(\frac{\partial f}{\partial x_j}(a)\right) = \sum_{k=1}^{m} \frac{\partial f_k}{\partial x_j}(a)\frac{\partial g}{\partial y_k}(b). \qquad \textit{CQFD}$$

Règle de la chaîne pour les dérivées

Proposition 4.14 Si $I \subset \mathbb{R}$ ouvert, $\gamma : I \to U$ dérivable en $t_0 \in I$ et $f : U \to \mathbb{R}^m$ différentiable en $\gamma(t_0)$, alors $f \circ \gamma$ est dérivable en t_0 et $\boxed{(f\circ\gamma)'(t_0) = \sum_{j=1}^{n} \gamma_j'(t_0)\dfrac{\partial f}{\partial x_j}(\gamma(t_0))}$

$$\begin{array}{ccccc} \mathbb{R} & & \mathbb{R}^n & & \\ \cup & & \cup & & \\ I & \xrightarrow{\gamma} & U & \xrightarrow{f} & \mathbb{R}^m \\ \cup & & \cup & & \cup \\ t_0 & \xrightarrow{\gamma} & b := \gamma(t_0) & \xrightarrow{f} & f(b) \end{array} \qquad \begin{array}{ccc} \mathbb{R}^n & \xrightarrow{df_b} & \mathbb{R}^m \\ \cup & & \cup \\ \gamma'(t_0) & \xrightarrow{df_b} & df_b(\gamma'(t_0)) = \sum_{j=1}^{n} \gamma_j'(t_0)\partial_{x_j} f(b) \end{array}$$

Schémas pour la règle de la chaîne pour les dérivées

Esquisse de preuve de la proposition 4.14

$$(f\circ\gamma)'(t_0)=d(f\circ\gamma)_{t_0}(1)=df_{\gamma(t_0)}(d\gamma_{t_0}(1))=df_{\gamma(t_0)}(\gamma'(t_0))=\sum_{j=1}^{n}\gamma_j'(t_0)\frac{\partial f}{\partial x_j}(\gamma(t_0)).$$ *CQFD*

Dérivées directionnelles

⚠ Si $v\in\mathbb{R}^n$, l'ensemble $J_{a,v}:=\{t\in\mathbb{R}\,;\,a+tv\in U\}$ est un ouvert qui contient 0. (Exercice TD3.15.)

Définition 4.15 Sous réserve d'existence, $\frac{\partial f}{\partial v}(a)\overset{\text{déf.}}{=}\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t}$ « dérivée (directionnelle) de f dans la direction du vecteur v au point a ».

Proposition 4.16 Si f est différentiable en a, alors, pour tout $v\in\mathbb{R}^n$, $\frac{\partial f}{\partial v}(a)$ existe et est donnée par

$$\boxed{\frac{\partial f}{\partial v}(a)=df_a(v)=\sum_{j=1}^{n}v_j\frac{\partial f}{\partial x_j}(a)}$$

Corollaire 4.17 Les conditions suivantes sont nécessaires pour que f soit différentiable en a : (i) les dérivées directionnelles $\frac{\partial f}{\partial v}(a)$ existent pour tout $v\in\mathbb{R}^n$; (ii) l'application $\mathbb{R}^n\ni v\mapsto\frac{\partial f}{\partial v}(a)$ est linéaire.

Esquisse de preuve de la proposition 4.16

Soit $\gamma(t):=a+tv$, $\forall\, t\in J_{a,v}$. Alors $\frac{\partial f}{\partial v}(a)=(f\circ\gamma)'(0)=df_a(\gamma'(0))=df_a(v)=\sum_{j=1}^{n}v_j\frac{\partial f}{\partial x_j}(a)$. *CQFD*

Règle de la chaîne pour les fonctions scalaires

Proposition 4.18 Sous les hypothèses de la règle de la chaîne, si f et g sont scalaires ($m=p=1$), alors $\boxed{d(g\circ f)_a=g'(f(a))df_a}$

Preuve de la proposition 4.18

Pour $h\in\mathbb{R}^n$, $d(g\circ f)_a(h)=dg_{f(a)}(df_a(h))=df_a(h)dg_{f(a)}(1)=g'(f(a))df_a(h)$. *CQFD*

Différentiable ⇒ continue

Proposition 4.19 Si f est différentiable en a : (i)$f(a+h)=f(a)+O(h)$; (ii) f est continue en a.

Esquisse de preuve de la proposition 4.19

(i) Soient $M\in\mathbb{R}$ et $\delta>0$ tels que $\|df_a(h)\|_2\le M\|h\|_2$, $\forall\, h\in\mathbb{R}^n$, et $[h\in -a+U,\ \|h\|_2\le\delta]\Rightarrow\|f(a+h)-f(a)-df_a(h)\|_2\le\|h\|_2$. L'inégalité triangulaire donne $[h\in -a+U,\ \|h\|_2\le\delta]\Rightarrow\|f(a+h)-f(a)\|_2\le(M+1)\|h\|_2$. (ii) Soit $(x^k)\subset U$ telle que $x^k\to a$. Soit k_0 tel que $\|x^k-a\|_2\le\delta$, $\forall\, k\ge k_0$. Pour $k\ge k_0$, $\|f(x^k)-f(a)\|_2\le(M+1)\|x^k-a\|_2\to 0$. *CQFD*

Esquisse de preuve de la règle de la chaîne (théorème 4.11)

Soient $\varepsilon>0$ et $\xi:=f(a+h)-f(a)$. Nous cherchons un $\delta>0$ tel que $[h\in -a+U,\ \|h\|_2\le\delta]\Rightarrow\|g(b+\xi)-g(b)-dg_b(df_a(h))\|_2\le r(\varepsilon)\|h\|_2$ (principe $r(\varepsilon)$).
Soient $\delta_1,\delta_2,\delta_3,N_1,N_2>0$ tels que : (i) $[h\in -a+U,\ \|h\|_2\le\delta_1]\Rightarrow\|f(a+h)-f(a)-df_a(h)\|_2\le\varepsilon\|h\|_2$; (ii) $[\xi\in -b+V,\ \|\xi\|_2\le\delta_2]\Rightarrow\|g(b+\xi)-g(b)-dg_b(\xi)\|_2\le\varepsilon\|\xi\|_2$; (iii) $[h\in -a+U,\ \|h\|_2\le\delta_3]\Rightarrow\|f(a+h)-f(a)\|_2\le N_1\|h\|_2$; (iv) $\|dg_b(\xi)\|_2\le N_2\|\xi\|_2$, $\forall\,\xi\in\mathbb{R}^m$.
Soit $\delta:=\min\{\delta_1,\delta_2/N_1,\delta_3\}$. Si $h\in U$ et $\|h\|_2\le\delta$:

$$\begin{aligned}\|g(b+\xi)-g(b)-dg_b(df_a(h))\|_2&\le\|g(b+\xi)-g(b)-dg_b(\xi)\|_2+\|dg_b(\xi)-dg_b(df_a(h))\|_2\\&\le\varepsilon\|\xi\|_2+N_2\|\xi-df_a(h)\|_2\le N_1\varepsilon\|h\|_2+N_2\varepsilon\|h\|_2=(N_1+N_2)\varepsilon\|h\|_2.\end{aligned}$$ *CQFD*

Transfert de $\mathbb{R}$ à $\mathbb{R}^n$

Nous allons apprendre un nouveau mécanisme de preuve : comment transférer à des fonctions scalaires des propriétés de fonctions réelles.

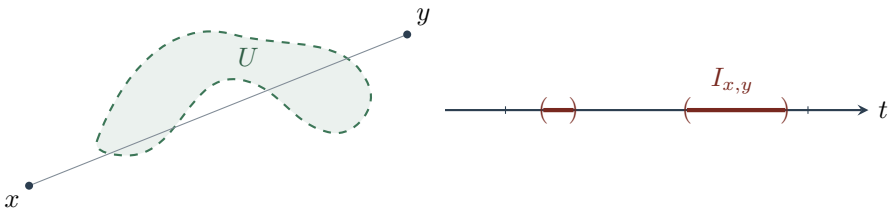
Définition 4.20 Soient $x, y \in \mathbb{R}^n$, $I_{x,y} \coloneqq \{t \in \mathbb{R}\,;\, (1-t)x + ty \in U\}$. La fonction de transfert est $\gamma = \gamma_{x,y} : I_{x,y} \to U$, $\gamma_{x,y}(t) \coloneqq (1-t)x + ty$.

- ⚠ $I_{x,y} \subset \mathbb{R}$ est un ouvert (éventuellement vide)
- $\gamma \in C^1$ et $\gamma'(t) = y - x$
- Si $f : U \to \mathbb{R}^m$ est différentiable, la règle de la chaîne donne $[f \circ \gamma]'(t) = df_{\gamma(t)}(y-x)$
- Dans le cas particulier où $f : U \to \mathbb{R}$, cette formule devient $[f \circ \gamma]'(t) = [(\nabla f)(\gamma(t))] \cdot (y - x)$

TAF et IAF

Proposition 4.21 (TAF et IAF) Si f est différentiable, scalaire, et $[x, y] \subset U$:

- $\exists \xi \in [x, y]$ tel que $f(y) - f(x) = [\nabla f(\xi)] \cdot (y - x)$ (TAF : théorème des accroissements finis)
- $|f(y) - f(x)| \leq \sup_{z \in [x,y]} \|\nabla f(z)\|_2 \|y - x\|_2$ (IAF : inégalité des accroissements finis)



À droite, l'ensemble $I_{x,y}$. Le segment $[x, y]$ correspond à $t \in [0, 1]$. Si $[x, y] \subset U$, $[0, 1] \subset I_{x,y}$

Esquisse de preuve de la proposition 4.21

- Nous avons $[0, 1] \subset I_{x,y}$. Par le TAF (version fonctions réelles), $\exists\, c \in]0, 1[$ tel que $f \circ \gamma(1) - f \circ \gamma(0) = [f \circ \gamma]'(c)$, ce qui revient au TAF avec $\xi \coloneqq (1-c)x + cy$
- Par (CS), $|f(y) - f(x)| \leq \|\nabla f(\xi)\|_2 \|y - x\|_2 \leq \sup_{z \in [x,y]} \|\nabla f(z)\|_2 \|y - x\|_2$ *CQFD*

Travail individuel

Exercice CM4.22 Si f est différentiable, scalaire, $x, y \in U$, et P est une ligne polygonale contenue dans U, d'extrémités x et y et longueur L, alors $|f(y) - f(x)| \leq \sup_{z \in P} \|\nabla f(z)\|_2 L$. ⚠ Faire un dessin !

Les formules de Taylor à l'ordre 1 (du nom de B. Taylor (1685–1731))

- « Avec reste en o » : Si f différentiable en a, $f(a+h) = f(a) + df_a(h) + o(\|h\|_2)$ ⚠ C'est la définition de la différentielle
- « Avec point intermédiaire » : si $f : U \to \mathbb{R}$ est différentiable et $[a, b] \subset U$, $\exists \xi \in [a, b]$ tel que $f(b) = f(a) + [\nabla f(\xi)] \cdot (b - a)$ ⚠ C'est le TAF
- « Avec reste intégral » : Si $f \in C^1(U, \mathbb{R})$ et $[a, b] \subset U$, $f(b) = f(a) + \displaystyle\int_0^1 df_{(1-t)a+tb}(b-a)\, dt$

Esquisse de preuve de la formule de Taylor avec reste intégral

$$f(b) - f(a) = \Big[(f \circ \gamma_{a,b})(t)\Big]_{t=0}^{t=1} = \int_0^1 [f \circ \gamma_{a,b}]'(t)\, dt = \int_0^1 df_{\gamma_{a,b}(t)}(b-a)\, dt. \qquad \textit{CQFD}$$

Travail individuel

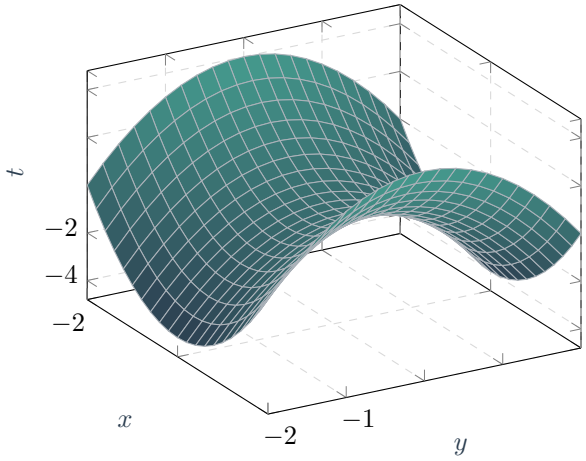
Exercice CM4.23 Énoncer et montrer la formule de Taylor avec reste intégral pour $f \in C^1(U, \mathbb{R}^m)$. Quel nouvel objet intervient dans cette formule ?

Visualisation des fonctions. Lien avec le calcul différentiel

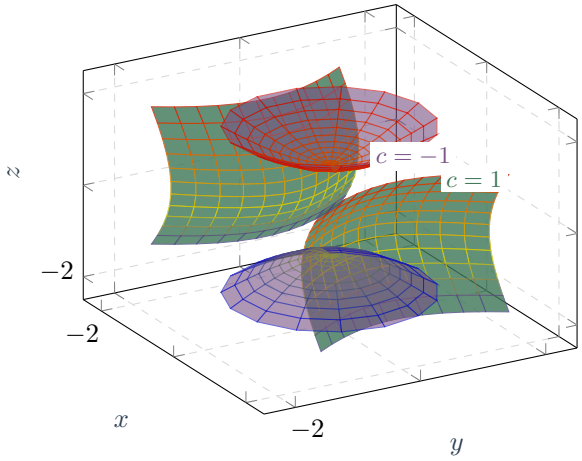
— Si pour les fonctions réelles le graphe donne une bonne indication visuelle de leur comportement, pour les fonctions de deux variables le graphe peut être moins instructif. D'autres méthodes sont utilisées : représentation des lignes de niveau $f = c$, éventuellement avec des codes couleurs (iso...), ou des champs de vecteurs avec un échantillon de flèches
— Pour les fonctions de trois variables, on peut visualiser des isosurfaces
— Et le calcul différentiel dans tout ça ? Les lignes, surfaces, etc. de niveau sont « orthogonales » au champ gradient $x \mapsto \nabla f(x)$. Par ailleurs, elles sont « souvent » de belles lignes ou surfaces
— Pour donner un sens rigoureux à ces propriétés et les montrer, rendez-vous en L3 (calcul différentiel) et M1 (géométrie différentielle) !

À réviser pour la cinquième leçon

Formules de Taylor à l'ordre 2 (avec reste « en o », avec point intermédiaire, avec reste intégral) pour les fonctions réelles. La méthode de transfert : la fonction auxiliaire $\gamma_{x,y}$ et le calcul de la dérivée $[f \circ \gamma_{x,y}]'$.

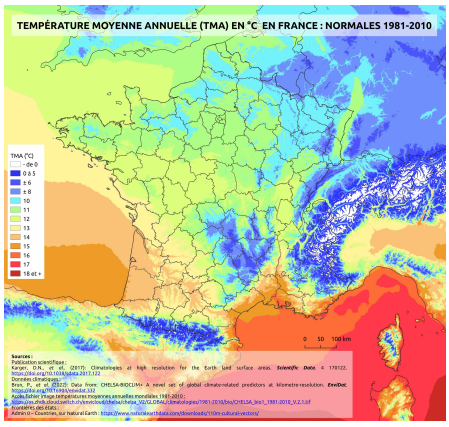


La surface qui représente le graphe de $(x, y) \mapsto x^2 - y^2$. L'origine est un point-selle

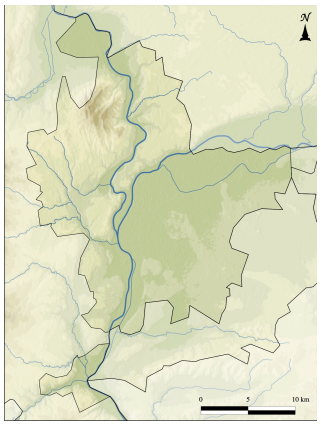


Isosurfaces de $(x, y, z) \mapsto x^2 - y^2 - z^2$

Surfaces visualisant des fonctions de deux et trois variables

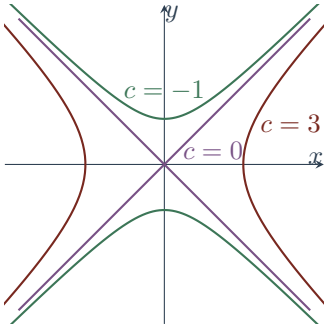


Moyenne des températures en France (1981–2010)

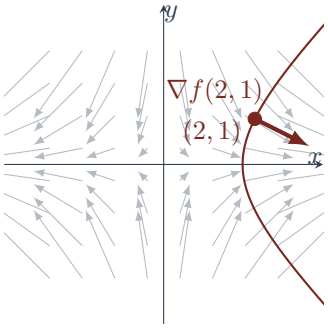


Relief de Lyon et de ses environs

Isothermes et isohypses matérialisées par des couleurs différentes.
© Wikimedia commons

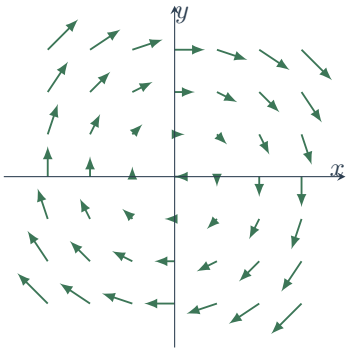


Lignes de niveau $x^2 - y^2 = c$ de la fonction $(x, y) \mapsto x^2 - y^2$

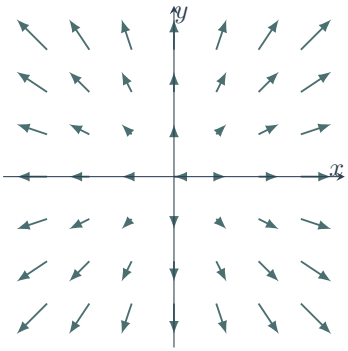


Le champ gradient de la fonction $(x, y) \mapsto x^2 - y^2$ (échelle 3 :10)

Les lignes de niveau sont orthogonales au champ gradient



Champ de rotation : $(x, y) \mapsto (y, -x)$ (échelle 1 : 4)



Champ divergent : $(x, y) \mapsto (x, y)$ (échelle 1 : 4)

Représentations de champs de vecteurs du plan

Leçon 5. Calcul différentiel d'ordre supérieur

Objectifs

Comprendre la difficulté conceptuelle du passage aux différentielles d'ordre supérieur et la notion de fonction deux fois différentiable. S'approprier les notations des dérivées partielles d'ordre supérieur. Comprendre le théorème de Schwarz (heuristique, preuve). Maîtriser les formules de Taylor à l'ordre 2.

Différentielles d'ordre supérieur

- Si $f : I \to \mathbb{R}$ est une fonction réelle, $f'' = (f')'$ (sous réserve d'existence)
- Et si $f : U \to \mathbb{R}^m$? La définition naturelle serait $d^2 f_a = d(df)_a$. Correct, mais ⚠ De quel espace vers quel espace va $a \mapsto df_a$?
- ⚠ $df : U \to \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m), d^2 f : U \to \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)), d^3 f : U \to \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)))$ 😱 Alors que dans le cas réel $f : I \to \mathbb{R}, f' : I \to \mathbb{R}, f'' : I \to \mathbb{R}$
- Perspective de simplification : df_a est encodée par un nombre fini de **fonctions** (les dérivées partielles $\partial_{x_j} f$)
- Nous allons nous rendre la vie plus simple en travaillant avec les différentielles des $\partial_{x_j} f$ au lieu des différentielles de df. L'approche « abstraite » $d^2 f = d(df)$ sera utilisée en L3. Son intérêt est de s'appliquer à des espaces plus généraux que $\mathbb{R}^n$. Pour des fonctions $f : U \to \mathbb{R}^m$, les deux approches sont équivalentes, et celle basée sur les $\partial_{x_j} f$ plus facile à implémenter

Fonction deux fois différentiable

Définition 5.1 Une fonction $f : U \to \mathbb{R}^m$ est deux fois différentiable en a si elle différentiable (partout) et ses dérivées partielles $\partial_{x_j} f$, $j = 1, \dots, n$, sont différentiables en a.

Dérivées partielles « générales » (d'ordre supérieur)

Sous réserve d'existence : $\dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) = \dfrac{\partial^2 f}{\partial x \partial y}$, $\dfrac{\partial}{\partial x}\left(\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial z}\right)\right) = \dfrac{\partial^3 f}{\partial x \partial y \partial z}$,… L'exposant $2, 3, \dots$, est l'**ordre** (de dérivation) : $\dfrac{\partial f}{\partial x}$ est une dérivée partielle du premier ordre, $\dfrac{\partial^2 f}{\partial x \partial y}$ du second ordre, …

Quiz

— Si f est deux fois différentiable, montrer que : (i) $f \in C^1$; (ii) f admet des dérivées partielles d'ordre deux ; (iii) $\dfrac{\partial f}{\partial v}$ différentiable, $\forall\, v \in \mathbb{R}^n$; (iv) $\exists\, \dfrac{\partial}{\partial w}\left(\dfrac{\partial f}{\partial v}\right)$, $\forall\, v, w \in \mathbb{R}^n$; (v) proposer un nom pour la propriété (iv)
— (vi) Proposer une définition des fonctions trois fois différentiables et énoncer quelques propriétés plausibles de telles fonctions ; (vii) que serait une fonction C^∞ ; (viii) un exemple de fonction C^∞ ?

Le cadre de cette leçon

— $f : U \to \mathbb{R}$ (fonctions scalaires), $a \in U$ ⚠ Le passage à $f : U \to \mathbb{R}^m$ est aisé
— df est encodée par $U \ni a \mapsto \nabla f(a) \in \mathbb{R}^n$
— La différentielle de ∇f sera encodée par $U \ni a \mapsto J_{\nabla f}(a) \in M_n(\mathbb{R})$ (matrice carrée). La différentielle seconde en tant qu'application $d^2 f : U \mapsto \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R}))$ ne sera pas utilisée ⚠ Attention aux corrigés disponibles en ligne ou proposés par les IA
— Nous n'aborderons pas les différentielles d'ordre ≥ 3. Les principaux résultats de cette leçon se généralisent sans difficulté à ces différentielles

La matrice hessienne (hessienne) (du nom de L. O. Hesse (1811–1874))

$$Hf(a) = H(f)(a) = \boxed{H_f(a)} := J_{\nabla f}(a) = \left(\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(a)\right)_{1 \leq i,j \leq n} = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right)_{1 \leq i,j \leq n}.$$

Exercice travaillé

Exercice CM5.2 Soient $v \in \mathbb{R}^n$, $A \in M_n(\mathbb{R})$ et $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) := v \cdot x + (Ax) \cdot x$, $\forall\, x \in \mathbb{R}^n$. Calculer ∇f et H_f.

Esquisse de solution de l'exercice CM5.2

$$\partial_{x_i} f = \partial_{x_i}\left[x \mapsto \left(\sum_j v_j x_j + \sum_{j,k} a_{jk} x_j x_k\right)\right] = \sum_j v_j \delta^i_j + \sum_{j,k} a_{jk}\left(\delta^i_j x_k + x_j \delta^i_k\right)$$
$$= v_i + \sum_k a_{ik} x_k + \sum_j a_{ji} x_j \Rightarrow \nabla f(x) = v + (A + A^\intercal)x \text{ (avec } x \text{ colonne)}$$

$$\partial_{x_j}\partial_{x_i} f = \partial_{x_j}\left(v_i + \sum_k a_{ik} x_k + \sum_\ell a_{\ell i} x_\ell\right) = \sum_k a_{ik}\delta^j_k + \sum_\ell a_{\ell i}\delta^j_\ell = a_{ij} + a_{ji} \Rightarrow H_f(x) = A + A^\intercal$$

Théorème de symétrie (du nom de H. A. Schwarz (1843–1921))

Théorème 5.3 (Théorème de Schwarz) Si f est deux fois différentiable en a, $\dfrac{\partial}{\partial v}\left(\dfrac{\partial f}{\partial w}\right)(a) = \dfrac{\partial}{\partial w}\left(\dfrac{\partial f}{\partial v}\right)(a)$, $\forall\, v, w \in \mathbb{R}^n$. De manière équivalente, $H_f(a)$ est symétrique.

Heuristique du théorème de symétrie et de sa preuve

Il suffit d'intuiter le cas d'une fonction g de deux variables lorsque $a = (0,0)$, $v = e^1$ et $w = e^2$. Si $u : \mathbb{R} \to \mathbb{R}$ est dérivable, nous avons intuitivement $u'(x) \approx [u(x+\varepsilon) - u(x)]/\varepsilon$. Lisons cette égalité approchée dans l'autre sens : $u(x+\varepsilon) - u(x) \approx \varepsilon u'(x)$. En appliquant deux fois cette heuristique à l'expression $G(\varepsilon)$ ci-dessous, nous obtenons

$$G(\varepsilon) \coloneqq [g(\varepsilon,\varepsilon) - g(\varepsilon,0)] - [g(0,\varepsilon) - g(0,0)] \approx \varepsilon \partial_y g(\varepsilon,0) - \varepsilon \partial_y g(0,0) \approx \varepsilon^2 \partial_x \partial_y g(0,0).$$

Par ailleurs, nous avons l'identité $G(\varepsilon) = [g(\varepsilon,\varepsilon) - g(0,\varepsilon)] - [g(\varepsilon,0) - g(0,0)]$. En reprenant l'heuristique ci-dessus,

$$G(\varepsilon) = [g(\varepsilon,\varepsilon) - g(0,\varepsilon)] - [g(\varepsilon,0) - g(0,0)] \approx \varepsilon \partial_x g(0,\varepsilon) - \varepsilon \partial_x g(0,0) \approx \varepsilon^2 \partial_y \partial_x g(0,0).$$

Ceci nous amène à la compréhension du résultat et à la stratégie de la preuve : (i) si $n = 2$, $g : V \to \mathbb{R}$, avec $(0,0) \in V$, et g est deux fois différentiable en $(0,0)$, montrer que $\lim\limits_{\varepsilon \to 0} \dfrac{G(\varepsilon)}{\varepsilon^2} = \partial_x \partial_y g(0,0) = \partial_y \partial_x g(0,0)$; (ii) réduire le cas général au cas particulier du (i).

Grandes lignes de la preuve du théorème 5.3

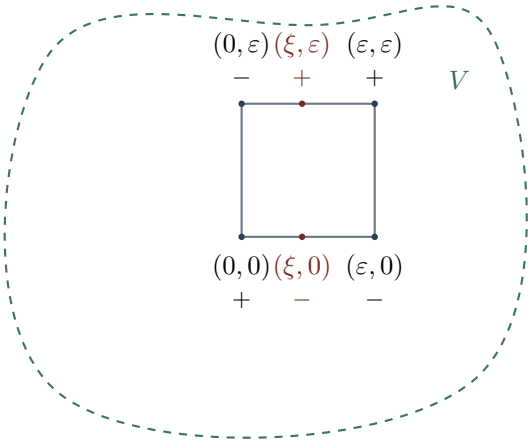
— Soit $V \coloneqq \{(x,y) \in \mathbb{R}^2 \,;\, a + xv + yw \in U\}$, ouvert contenant $(0,0)$. Soit $g : V \to \mathbb{R}$, $g(x,y) \coloneqq f(a + xv + yw)$. Alors : (i) g est deux fois différentiable en $(0,0)$; (ii) $\dfrac{\partial}{\partial v}\left(\dfrac{\partial f}{\partial w}\right)(a) = \partial_x \partial_y g(0,0)$; (iii) $\dfrac{\partial}{\partial w}\left(\dfrac{\partial f}{\partial v}\right)(a) = \partial_y \partial_x g(0,0)$. Nous avons ramené le théorème à $\partial_x \partial_y g(0,0) = \partial_y \partial_x g(0,0)$

— Soit $\lambda > 0$ tel que $B((0,0), \sqrt{2}\lambda) \subset V$. Soient $0 < \varepsilon < \lambda$ et $u(t) \coloneqq g(t,\varepsilon) - g(t,0)$, $\forall\, t \in]-\lambda, \lambda[$. Alors $G(\varepsilon) = u(\varepsilon) - u(0)$. Le TAF appliqué à u donne l'existence d'un $\xi = \xi(\varepsilon) \in]0, \varepsilon[$ tel que $G(\varepsilon) = \varepsilon u'(\xi) = \varepsilon[\partial_x g(\xi,\varepsilon) - \partial_x g(\xi,0)]$. En appliquant à la fonction différentiable $\partial_x g$ la formule de Taylor à l'ordre 1, nous obtenons

$$\begin{aligned} G(\varepsilon) &= \varepsilon[\partial_x \partial_x g(0,0)\xi + \partial_y \partial_x g(0,0)\varepsilon + o(\|(\xi,\varepsilon)\|_2) - \partial_x \partial_x g(0,0)\xi - o(\|(\xi,0)\|_2)] \\ &= \varepsilon^2 \partial_y \partial_x g(0,0) + \varepsilon[o(\|(\xi,\varepsilon)\|_2) - o(\|(\xi,0)\|_2)]. \end{aligned}$$

En explicitant o, nous obtenons $\lim_{\varepsilon \to 0} G(\varepsilon)/\varepsilon^2 = \partial_y \partial_x g(0,0)$. En utilisant la deuxième formule pour $G(\varepsilon)$ et la démarche ci-dessus, nous obtenons $\lim_{\varepsilon \to 0} G(\varepsilon)/\varepsilon^2 = \partial_x \partial_y g(0,0)$

— L'équivalence entre le théorème de Schwarz et la symétrie de la matrice hessienne suit de la formule des dérivées directionnelles (voir l'exercice TD4.1) *CQFD*



Avec G et u comme dans la preuve du théorème 5.3, les signes $+$ et $-$ correspondent aux formules de G, respectivement u'. La coordonnée ξ est donnée par le TAF

Rappels : formules de Taylor à l'ordre 2 (cas réel)

Soient $f : I \to \mathbb{R}$ deux fois dérivable, $a, b \in I$. Nous avons les formules de Taylor à l'ordre 2 :

— « Avec reste en o » : $f(a+h) = f(a) + f'(a)h + (1/2)f''(a)h^2 + o(h^2)$

— « Avec point intermédiaire » : $\exists\, c \in [a,b]$ tel que $f(b) = f(a) + f'(a)(b-a) + (1/2)f''(c)(b-a)^2$

— « Avec reste intégral » ($f \in C^2$) : $f(b) = f(a) + f'(a)(b-a) + \int_0^1 (1-t)f''((1-t)a+tb)(b-a)^2\,dt$

Formules de Taylor à l'ordre 2

Proposition 5.4 — « Avec reste en o » : f deux fois différentiable en $a \Rightarrow f(a+h) = f(a) + [\nabla f(a)] \cdot h + \frac{1}{2}\big[[H_f(a)]h\big] \cdot h + o((\|h\|_2)^2)$

— « Avec point intermédiaire » : f deux fois différentiable, $[a,b] \subset U \Rightarrow \exists\, \xi \in [a,b]$ tel que $f(b) = f(a) + [\nabla f(a)] \cdot h + \frac{1}{2}\big[[H_f(\xi)]h\big] \cdot h$

— « Avec reste intégral » : $f \in C^2$, $[a,b] \subset U \Rightarrow f(b) = f(a) + [\nabla f(a)] \cdot (b-a) + \displaystyle\int_0^1 (1-t)\big[[H_f((1-t)a+tb)](b-a)\big] \cdot (b-a)\,dt$

Exercice travaillé

Exercice CM5.5 Calculer $[f \circ \gamma]''$, avec γ la fonction de transfert et f deux fois différentiable.

Esquisse de solution de l'exercice CM5.5

$$[f \circ \gamma]''(t) = \frac{d}{dt}\left(t \mapsto \sum_{j=1}^{n}(y_j - x_j)(\partial_{x_j} f)(\gamma(t))\right) = \sum_{j=1}^{n}(y_j - x_j)\frac{d}{dt}\big(t \mapsto (\partial_{x_j} f)(\gamma(t))\big)$$

$$= \sum_{j=1}^{n}(y_j - x_j)\sum_{k=1}^{n}(y_k - x_k)(\partial_{x_k}\partial_{x_j} f)(\gamma(t)) = \big[[H_f(\gamma(t))](y-x)\big] \cdot (y-x). \qquad \textit{CQFD}$$

Esquisses de preuves des formules de Taylor avec point intermédiaire et reste intégral

On prend $x = a$ et $y = b$ dans la fonction de transfert γ et on pose $g := f \circ \gamma$.

— Il existe $c \in]0,1[$ tel que $g(1) = g(0) + g'(0) + (1/2)g''(c)$, ce qui donne (via l'exercice CM5.5) la formule de Taylor avec point intermédiaire avec $\xi := (1-c)a + cb$

— Si $f \in C^2$, alors $g \in C^2([0,1])$. La formule réelle de Taylor avec reste intégral (sur $[0,1]$) donne la formule de Taylor avec reste intégral (sur $[a,b]$) *CQFD*

Grandes lignes de la preuve de la formule de Taylor avec reste en o

— Soit $g(h) := f(a+h) - [\nabla f(a)] \cdot h - (1/2)\big[[H_f(a)]h\big] \cdot h$, $\forall\, h \in -a + U$. Nous devons montrer que $g(h) = o((\|h\|_2)^2)$

— En utilisant l'exercice CM5.2 et le théorème 5.3, nous avons $g(0) = 0$, $\partial_{x_j} g(0) = 0$, $d\partial_{x_j} g_0 = 0$, $\forall\, 1 \le j \le n$

— Si $r > 0$ est tel que $B(0,r) \subset -a + U$ et si $\|h\|_2 < r$, le TAF donne l'existence de $\xi = \xi(h) \in [0,h]$ tel que $g(h) = g(0) + \sum_{j=1}^n h_j \partial_{x_j} g(\xi) = \sum_{j=1}^n h_j \partial_{x_j} g(\xi)$

— Comme $\partial_{x_j} g(0) = 0$ et $d\partial_{x_j} g_0 = 0$, nous avons $\partial_{x_j} g(\xi) = o(\|\xi\|_2)$, $\forall\, 1 \le j \le n$

— En explicitant $o(\|\xi\|_2)$ et en utilisant $\|\xi\|_2 \le \|h\|_2$, nous obtenons $g(h) = o((\|h\|_2)^2)$ *CQFD*

Potentiel d'un champ de vecteurs

— Si on se donne un champ de vecteurs $F : U \to \mathbb{R}^n$, f est un « potentiel » de F si $\nabla f = F$ (⚠ parfois, en physique, $\nabla f = -F$). Si f existe, F est un « champ gradient »

— En dimension $n = 1$, F est une fonction, et toute fonction continue a un potentiel (théorème fondamental du calcul intégral). Ce n'est plus le cas si $n \ge 2$!

Lemme de H. Poincaré (1854–1912)

Théorème 5.6 (Lemme de Poincaré) $F \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ est un champ gradient $\Leftrightarrow \partial_{x_i} F_j = \partial_{x_j} F_i$, $\forall\, 1 \leq i, j \leq n$.

Grandes ligne de la preuve du théorème 5.6

« $\Rightarrow$ » Si le potentiel f existe, alors $f \in C^2$, et donc $\partial_{x_i}\partial_{x_j} f = \partial_{x_j}\partial_{x_i} f$, d'où $\partial_{x_i} F_j = \partial_{x_j} F_i$.

« $\Leftarrow$ » Heuristique : si $\exists\, f$ et $f(0) = 0$, on a $f(x) = \Big[f(tx)\Big]_{t=0}^{t=1} = \displaystyle\int_0^1 \frac{d}{dt}[t \mapsto f(tx)]\, dt = \int_0^1 [\nabla f(tx)]\cdot x\, dt =$ $\displaystyle\int_0^1 [F(tx)] \cdot x\, dt$, ce qui suggère de « poser » (1) $f(x) := \displaystyle\sum_{j=1}^n x_j \int_0^1 F_j(tx)\, dt$. En dérivant (1),

$$\partial_{x_i} f(x) = \int_0^1 F_i(tx)\, dt + \sum_{j=1}^n x_j \int_0^1 (\partial_{x_i} F_j)(tx) t\, dt = \int_0^1 F_i(tx)\, dt + \int_0^1 t \sum_{j=1}^n (\partial_{x_j} F_i)(tx) x_j\, dt =$$

$$\int_0^1 F_i(tx)\, dt + \int_0^1 t \frac{d}{dt}[t \mapsto (F_i(tx))]\, dt = \int_0^1 F_i(tx)\, dt + \Big[tF_i(tx)\Big]_{t=0}^{t=1} - \int_0^1 F_i(tx)\, dt = F_i(x). \qquad \textit{CQFD}$$

Remarques

- Pour justifier la dérivation sous l'intégrale, voir Compléments de mathématiques
- La preuve ci-dessus du lemme de Poincaré et sa conclusion restent valides si $F \in C^1(U, \mathbb{R}^n)$ avec $U \subset \mathbb{R}^n$ convexe. ⚠ Le lemme de Poincaré n'est pas vrai si U est un ouvert quelconque
- Niveau L3–M1 : le lemme de Poincaré pour F continu dans un ouvert « approprié »
- ⚠ En pratique, on peut se passer de la formule (1) ; voir l'exercice TD 4.9

Traduction en langage des physiciens et géomètres

En physique, la variation élémentaire d'une fonction « d'état », mettons $U = U(V, T)$, est la « forme » $dU = \partial_T U dT + \partial_V U dV$, que l'on peut identifier à la différentielle dU des analystes. Si la forme $PdT + QdV$ s'écrit comme dU, elle doit satisfaire $\partial_V P = \partial_T Q$. On dit alors que la forme est « fermée ». Le lemme de Poincaré affirme qu'une forme fermée est « exacte » (=différentielle d'une fonction).

À réviser pour la sixième leçon

Extrémum global ou local. Point critique. Étude locale de fonctions réelles : points critiques, rôle de la dérivée seconde. Matrices symétriques : valeurs propres, diagonalisation, polynôme caractéristique.

Leçon 6. Étude locale de fonctions

Objectifs

Savoir mener l'étude locale d'une fonction scalaire : trouver les points critiques, étudier leur nature locale et, si nécessaire, effectuer une analyse globale.

Cadre de cette leçon

$f : U \to \mathbb{R}$ est une fonction scalaire, différentiable (pour la recherche de points critiques) ou deux fois différentiable (pour l'étude locale).

Quiz : points d'extrémum

- Rappeler les notions de maximum, minimum, extrémum local ou global dans $\mathbb{R}$, et proposer les définitions adaptées à $\mathbb{R}^n$
- Condition nécessaire pour qu'un point soit un point d'extrémum local ? Suggestion de résultat

analogue dans $\mathbb{R}^n$?
— Définition d'un point critique ? Suggestion de définition analogue dans $\mathbb{R}^n$?
— Rôle de la dérivée seconde dans l'étude de la nature d'un point critique ?

Condition nécessaire du premier ordre

Soit $f : U \to \mathbb{R}$ différentiable.

Définition 6.1 $a \in U$ est un point critique de f si $df_a = 0$.

⚠ De manière équivalente, $\nabla f(a) = 0$.

Théorème 6.2 (Théorème de Fermat (condition nécessaire du 1er ordre)) Si a est un extrémum local de f, alors a est un point critique de f : $\boxed{\nabla f(a) = 0}$

Esquisse de preuve du théorème 6.2

Supposons a maximum local. Soit $r > 0$ tel que $B(a, r) \subset U$. Soit $\varepsilon > 0$ tel que $[x \in U, \|x - a\|_2 < \varepsilon] \Rightarrow f(x) \leq f(a)$. Soit $\rho := \min(r, \varepsilon)$. La fonction partielle $]a - \rho, a + \rho[\ni x_1 \mapsto g(x_1) := f(x_1, a_2, \ldots, a_n)$ est dérivable et atteint son maximum en $x_1 = a_1$, d'où $0 = g'(a_1) = \partial_{x_1} f(a)$. De même, $\partial_{x_j} f(a) = 0$, $\forall\, 2 \leq j \leq n$. *CQFD*

Exercice travaillé

Exercice CM6.3 Trouver les points critiques de $\mathbb{R}^2 \ni (x, y) \mapsto f(x, y) := x^3 + y^3 - 3x^2y - 3xy^2$.

Esquisse de solution de l'exercice CM6.3

$\nabla f(x, y) = 0 \Rightarrow 2xy = x^2 - y^2 = y^2 - x^2 \Rightarrow x^2 = y^2 \,\&\, xy = 0 \Rightarrow x = y = 0$. Par ailleurs, $(0, 0)$ est un point critique de f, donc le seul point critique. *CQFD*

Quelques propriétés des matrices symétriques $A \in M_n(\mathbb{R})$ (cf Algèbre IV)

— Toutes les valeurs propres (v. p.) $\lambda_1, \ldots, \lambda_n$ de A sont réelles
— Il existe une base orthonormée $(f^1, \ldots, f^n)$ de $\mathbb{R}^n$ telle que $Af^j = \lambda_j f^j$, $\forall\, 1 \leq j \leq n$
— Si $n = 2$: (i) $P_A(X) = X^2 - (\mathrm{tr}\ A)X + (\det A)I_2$; (ii) on peut lire le signe des v. p. à partir des signes de $\mathrm{tr}\ A$ et $\det A$ (exercice TD4.17)
— On peut mener une étude similaire si $n = 3$ (exercice TD4.18)
— On peut déterminer les signes des valeurs propres d'une matrice par un procédé algorithmique (réduction de Gauss, exercice TD4.20). Le fait que ce procédé donne les signes des valeurs propres est la loi d'inertie de Sylvester
— Un autre résultat utile est le critère de Sylvester : toutes les valeurs propres de A sont > 0 si et seulement si $\det A_k > 0$, $\forall\, 1 \leq k \leq n$, où $A_k := (a_{ij})_{1 \leq i,j \leq k}$

Condition suffisante du second ordre

⚠ Si A est symétrique, $A > 0$ (respectivement $A \geq 0$, $A \leq 0$, $A < 0$) signifie que toutes les v. p. de A sont > 0 (respectivement ≥ 0, ≤ 0, < 0).

Théorème 6.4 (Condition suffisante du second ordre) Soient $f : U \to \mathbb{R}$ deux fois différentiable et a un point critique de f.
— Si $H_f(a) > 0$, a est point de minimum local strict
— Si $H_f(a) < 0$, a est point de maximum local strict
— Si λ est une v. p. > 0 (respectivement < 0) de $H_f(a)$ et v un vecteur propre associé, alors $t \mapsto f(a + tv)$ a un minimum local strict (respectivement maximum local strict) en $t = 0$
— Si $H_f(a)$ a au moins une v. p. $\lambda > 0$ de vecteur propre v et une v. p. $\mu < 0$ de vecteur propre w, alors a est un point de minimum local strict dans la direction de v et un point de maximum local strict dans la direction de w. En particulier, a n'est pas un point d'extrémum local. ⚠ Un tel point est un point col ou point-selle
— Dans les autres cas, on ne peut pas conclure. ⚠ Mais il peut y avoir des indices. Par exemple, si

$H_f(a) \geq 0$ et $H_f(a) \neq 0_n$, a ne peut pas être un point de maximum local

Esquisse de preuve du théorème 6.4

— Cas où $H_f(a) > 0$. Soit $C > 0$ tel que $\big[[H_f(a)]h\big] \cdot h \geq C(\|h\|_2)^2$, $\forall\, h \in \mathbb{R}^n$ (cf exercice TD 1.21). Soit $\delta > 0$ tel que $[h \in -a+U, \|h\|_2 < \delta] \Rightarrow \big|f(a+h) - f(a) - (1/2)\big[[H_f(a)]h\big] \cdot h\big| \leq (C/4)(\|h\|_2)^2$. Alors $[h \in -a+U, 0 < \|h\|_2 < \delta] \Rightarrow f(a+h) > f(a)$. Le cas où $H_f(a) < 0$ est similaire
— Soit v vecteur propre correspondant à la v. p. $\lambda > 0$. Soit $\delta > 0$ tel que $[h \in -a+U, \|h\|_2 < \delta]$ $\Rightarrow \big|f(a+h) - f(a) - (1/2)\big[[H_f(a)]h\big] \cdot h\big| \leq (\lambda/4)(\|h\|_2)^2$. Alors $[a+tv \in U, 0 < |t| < \delta/\|v\|_2] \Rightarrow$ $f(a+tv) > f(a)$. Le cas où $\lambda < 0$ est similaire
— La conclusion pour les points-selles découle de l'item précédent. De même, si $H_f(a)$ a une v. p. > 0 (respectivement < 0), a n'est pas un point de maximum local (respectivement minimum local)
— Pour voir que dans les cas restants il n'y a pas de conclusion a priori possible, voir l'exercice TD 4.21 *CQFD*

À réviser pour la septième leçon

Fonctions convexes réelles : exemples, inégalité(s) de Jensen. Caractérisation des fonctions convexes dérivables et deux fois dérivables. Fonctions monotones. La méthode de transfert : la fonction auxiliaire $\gamma_{x,y}$ et le calcul des dérivées de $f \circ \gamma_{x,y}$.

Leçon 7. Convexité

Objectifs

Maîtriser les caractérisations des fonctions convexes. Comprendre la notion de fonction monotone dans $\mathbb{R}^n$. Savoir montrer qu'une fonction concrète est convexe.

Quiz : fonctions réelles convexes

— Donner : (i) trois exemples de fonctions convexes ; (ii) un exemple de fonction convexe périodique ; (iii) un exemple de fonction convexe discontinue
— Une fonction dérivable est convexe si et seulement si... Une fonction deux fois dérivable est convexe si et seulement si...
— Proposer la définition d'une fonction scalaire convexe

Fonctions convexes

Définition 7.1 Une fonction $f : C \to \mathbb{R}$ est convexe si $C \subset \mathbb{R}^n$ est convexe et

$$\forall\, x, y \in C, \forall\, t \in [0,1],\ f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \text{ (inégalité de Jensen).} \qquad \text{(J)}$$

⚠ Du nom de J. L. W. V. Jensen (1859–1925)

Le cadre de cette leçon

$f : C \to \mathbb{R}$ une fonction scalaire, avec $C \subset \mathbb{R}^n$ convexe.

Mécanisme de transfert pour les fonctions convexes

Proposition 7.2 f est convexe si et seulement si $\forall\, x, y \in C$, $g_{x,y} \coloneqq f \circ \gamma_{x,y} : [0,1] \to \mathbb{R}$ est convexe.

Esquisse de preuve de la proposition 7.2

« $\Leftarrow$ » Si $x, y \in C$ et $t \in [0,1]$, l'inégalité $g_{x,y}(t) \leq (1-t)g_{x,y}(0) + tg_{x,y}(1)$ donne (J).
« $\Rightarrow$ » Si $x, y \in C$ et $s_1, s_2, t \in [0,1]$, l'inégalité $g_{x,y}((1-t)s_1 + ts_2) \leq (1-t)g_{x,y}(s_1) + tg_{x,y}(s_2)$ équivaut à $f((1-t)X + tY) \leq (1-t)f(X) + tf(Y)$, avec $X \coloneqq (1-s_1)x + s_1 y, Y \coloneqq (1-s_2)x + s_2 y \in C$. *CQFD*

Travail individuel (inégalité de Jensen généralisée)

Exercice CM7.3 Soit $f : C \to \mathbb{R}$ convexe. Montrer l'inégalité de Jensen généralisée

$$\left[\ell \geq 2, \lambda_1, \dots, \lambda_\ell \geq 0, \sum_{j=1}^{\ell} \lambda_j = 1, x^1, \dots, x^\ell \in C\right] \Rightarrow f\left(\sum_{j=1}^{\ell} \lambda_j x^j\right) \leq \sum_{j=1}^{\ell} \lambda_j f(x^j).$$

Caractérisations des fonctions convexes

Proposition 7.4 Si C est ouvert et f est différentiable, il y a équivalence entre : (a) f convexe ; (b) $[\nabla f(y) - \nabla f(x)] \cdot (y - x) \geq 0, \forall\, x, y \in C$; (c) $f(y) \geq f(x) + [\nabla f(x)] \cdot (y - x), \forall\, x, y \in C$.

⚠ Reformulation de (c) : le graphe G_f de f est au-dessus du plan tangent au graphe en $(x, f(x))$, $\forall\, x \in C$.

Définition 7.5 (Fonction monotone) Une fonction $F : A \to \mathbb{R}^n$ (avec $A \subset \mathbb{R}^n$) est monotone si $[F(y) - F(x)] \cdot (y - x) \geq 0, \forall\, x, y \in A$.

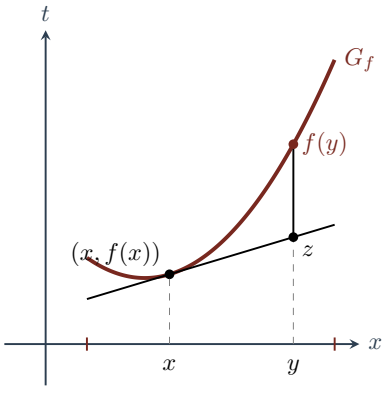
⚠ Reformulation de la proposition 7.4 : pour f différentiable, f convexe $\Leftrightarrow$ ∇f monotone $\Leftrightarrow$ le graphe G_f de f est au-dessus de tous les plans tangents à G_f.

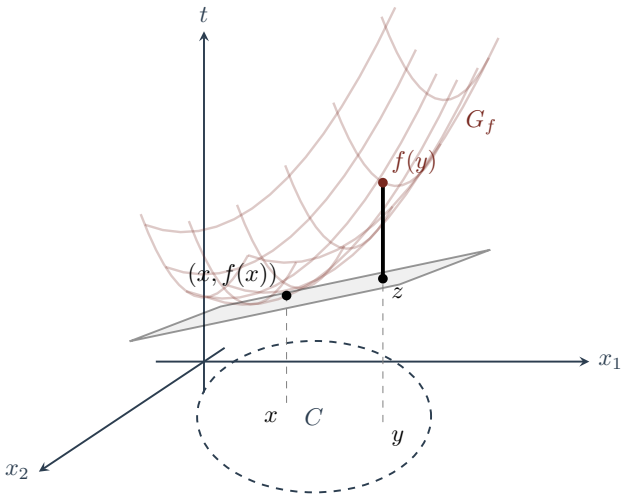
Proposition 7.6 Si C est ouvert et f est deux différentiable, f est convexe si et seulement si $H_f(a) \geq 0$, $\forall\, x, y \in C$.

Travail individuel

Exercice CM7.7 Soient $A \in M_n(\mathbb{R})$ symétrique et $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) := (Ax) \cdot x, \forall\, x \in \mathbb{R}^n$. Montrer que f est convexe si et seulement si $A \geq 0$.

Exercice CM7.8 Montrer que les fonctions $\mathbb{R}^n \ni x \mapsto \|x\|_2$ et $]1, \infty[^2 \ni (x, y) \mapsto x^4 - 4xy + y^2$ sont convexes.





Une fonction différentiable est convexe $\Leftrightarrow$ son graphe est au-dessus de chaque plan tangent

Ceci se traduit par $f(y) \geq z, \forall\, x, y \in C$

Esquisses de preuves des propositions 7.4 et 7.6

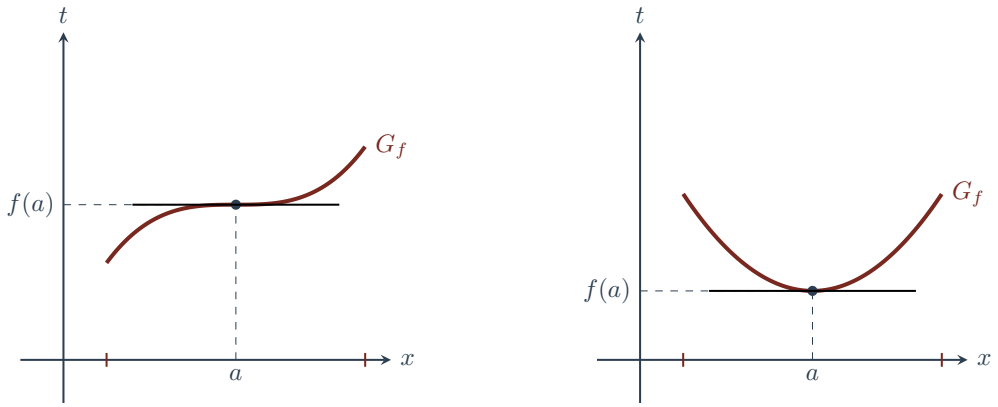
Proposition 7.4. « (a) $\Rightarrow$ (b) » $g_{x,y}$ est convexe et dérivable, d'où $(g_{x,y})'(1) - (g_{x,y})'(0) \geq 0$, ce qui revient à $[\nabla f(y) - \nabla f(x)] \cdot (y - x) \geq 0$. « (b) $\Rightarrow$ (c) » Soient $x, y \in C$. Il existe $t \in]0, 1[$ tel que, avec $z := (1-t)x + ty$, $f(y) = f(x) + [\nabla f(z)] \cdot (y - x)$. De (b), nous avons $[\nabla f(z)] \cdot (z - x) \geq [\nabla f(x)] \cdot (z - x)$, d'où $[\nabla f(z)] \cdot (y - x) \geq [\nabla f(x)] \cdot (y - x)$, ce qui implique (c). « (c) $\Rightarrow$ (a) » Soient $x, y \in C$, $t \in [0, 1]$ et $z := (1 - t)x + ty$. De (c), $f(x) \geq f(z) + [\nabla f(z)] \cdot (x - z)$, $f(y) \geq f(z) + [\nabla f(z)] \cdot (y - z)$; d'où $(1 - t)f(x) + tf(y) \geq f(z)$. *CQFD*

Proposition 7.6. « $\Rightarrow$ » Nous avons $\big[[H_f(a)]h\big] \cdot h = \dfrac{\partial}{\partial h}\left(\dfrac{\partial f}{\partial h}\right)(a) = \lim\limits_{t \to 0} \dfrac{[\nabla f(a + th)] \cdot h - [\nabla f(a)] \cdot h}{t}$. Or,

par la proposition 7.4 (b), ce quotient est ≥ 0. « $\Leftarrow$ » Soient $x, y \in C$, $t \in [0,1]$ et $z := (1-t)x + ty$. De la formule de Taylor à l'ordre 2 avec point intermédiaire et l'hypothèse $H_f(a) \geq 0$, $\forall\, a \in C$, $f(x) \geq f(z) + [\nabla f(z)] \cdot (x - z)$ et $f(y) \geq f(z) + [\nabla f(z)] \cdot (y - z)$, d'où $(1-t)f(x) + tf(y) \geq f(z)$. *CQFD*

Condition nécessaire et suffisante du premier ordre

Proposition 7.9 Si C est ouvert, f est convexe et différentiable et $a \in C$, il y a équivalence entre : (a) a est un point de minimum de f ; (b) a est un point de minimum local de f ; (c) a est un point critique de f.



La droite tangente en un point critique est horizontale. Propriété analogue pour le plan tangent
Pour une fonction convexe, plan tangent horizontal $\Rightarrow$ point de minimum

Preuve de la proposition 7.9

« (a) $\Rightarrow$ (b) » est évident. « (b) $\Rightarrow$ (c) » suit de la condition nécessaire du premier ordre. « (c) $\Rightarrow$ (a) » f étant convexe et différentiable, nous avons $f(x) \geq f(a) + [\nabla f(a)] \cdot (x - a) = f(a)$, $\forall\, x \in C$. *CQFD*

Quiz

(a) Une fonction convexe $f : \mathbb{R} \to \mathbb{R}$ est-elle dérivable ? (b) A-t-elle tout de même une propriété de dérivabilité ? (c) Quel serait l'analogue de cette propriété pour des fonctions de plusieurs variables ?

Autres propriétés des fonctions convexes (admises)

— Une fonction f convexe sur un ouvert $C \subset \mathbb{R}^n$ est continue. De plus, si $K \subset C$ est compact, la restriction de f à K est lipschitzienne. (Résultats de niveau M1.)
— Une fonction convexe sur un ouvert admet des dérivées directionnelles « latérales » en tout point. (Idée de preuve : transférer ce résultat depuis les fonctions réelles.)

À réviser pour la huitième leçon

La dépendance/indépendance linéaire des vecteurs de $\mathbb{R}^n$.

Leçon 8. Modélisation. Optimisation sous contraintes

Objectifs

Savoir modéliser des problèmes simples d'optimisation libre ou sous contraintes. Savoir implémenter la méthode des multiplicateurs de Lagrange.

Exercices travaillés

Exercice CM8.1 Modéliser les problèmes suivants, en identifiant : les données, les inconnues, les hypothèses simplificatrices et l'objectif.

- Parmi les parallélépipèdes rectangles de volume 1, en trouver un de surface latérale minimale
- Une personne peut courir en bord de mer à la vitesse de 3 m/s et nager à la vitesse de 0,9 m/s. Trouver le temps minimum dont elle a besoin pour relier un point donné en bord de mer à un point donné en mer
- Trouver l'emplacement idéal d'une mangeoire pour oiseaux nichant dans trois zones de forêt.
- Parmi les triangles rectangles de périmètre 1, en trouver un d'aire maximale
- Trouver le chemin optimal d'un facteur qui dessert n maisons à partir de son bureau

Esquisse de solution de l'exercice CM8.1

- Le parallélépipède est déterminé par les longueurs inconnues a, b, c de ses côtés. Il faut minimiser l'« objectif surface » $2(ab + bc + ca)$ sous les contraintes $a, b, c > 0$, $abc = 1$
- Hypothèses simplificatrices : (i) le bord de mer est rectiligne ; (ii) la démarche raisonnable consiste à courir d'abord jusqu'à un autre point du bord de mer, puis de nager de ce point au point en mer. Si on identifie le bord de mer à l'axe $0x$, et la mer est « en haut », les données sont un point $(X, 0)$ sur le bord et un point (Y, Z), avec $Z > 0$, en mer. Inconnue : $(x, 0)$ sur le bord. Il faut minimiser l'« objectif temps » $\|(x, 0) - (X, 0)\|_2/3 + \|(Y, Z) - (x, 0)\|_2/0,9$, sans contrainte sur x
- Données : les populations a, b, c d'oiseaux et les emplacements $A, B, C \in \mathbb{R}^2$ des zones de forêt. Hypothèse simplificatrice : les besoins nutritionnels sont proportionnels à l'« objectif coût » $a\|X - A\|_2 + b\|X - B\|_2 + c\|X - C\|_2$, qu'il faut minimiser sans contrainte sur X
- Inconnues : les longueurs a, b des cathètes du triangle. Il faut maximiser l'« objectif aire » $ab/2$ sous les contraintes $a, b > 0$, $a + b + \sqrt{a^2 + b^2} = 1$
- Hypothèses simplificatrices : on peut joindre le bureau à toute maison et toutes les maisons entre elles, et le terrain est plat. Données : les distances d_{ij} de la maison i à la maison j, avec $0 \leq i, j \leq n$. ($i = 0$ est pour le bureau.) Inconnue : l'ordre de passage, décrit par une permutation $\sigma \in S_n$. Il faut minimiser, sans contrainte, l'« objectif distance » $d_{0\sigma(1)} + d_{\sigma(1)\sigma(2)} + \cdots + d_{\sigma(n)0}$ *CQFD*

Cadre et terminologie pour l'optimisation sans contrainte ou sous contraintes

- $U \subset \mathbb{R}^n$ un ensemble ouvert (pour simplifier) : l'ensemble où vivent les inconnues ⚠ Exemple dans le problème course/nage : $U = \mathbb{R}$
- Une fonction objectif, à maximiser ou minimiser ⚠ Exemple dans le problème des parallélépipèdes rectangles : $f(a, b, c) = 2(ab + bc + ca)$, à minimiser
- Des contraintes, éventuellement inexistantes, d'égalité ou d'inégalité ⚠ Nous ne considérons que le cas d'un nombre fini de contraintes ⚠ Exemples dans le problème des triangles rectangles : contraintes d'inégalité stricte $a > 0$, $b > 0$, contrainte d'égalité $a + b + \sqrt{a^2 + b^2} = 1$
- De manière condensée, nous cherchons $\max_{x \in F} f(x)$ ou $\min_{x \in F} f(x)$, où F est l'ensemble des points de U respectant les contraintes
- Un point $x^* \in F$ qui réalise le maximum ou le minimum recherché est une solution optimale, et la valeur $f(x^*)$ est l'optimum du problème ⚠ Dans des problèmes de gestion ou finance, on minimise les coûts ou on maximise les bénéfices. Dans ce contexte, l'optimum est traditionnellement désigné comme coût minimal ou bénéfice maximal
- ⚠ Illustration sur le problème des triangles rectangles : $U = \mathbb{R}^2$, $F = \{(a, b) \in \mathbb{R}^2 \,;\, a > 0, b > 0, a + b + \sqrt{a^2 + b^2} = 1\}$, $f(a, b) = ab/2$, et le problème est $\max\{f(a, b) \,;\, (a, b) \in F\}$

Multiplicateurs de Lagrange (du nom de G. L. Lagrangia/J.-L. Lagrange (1736-1813))

⚠ On peut toujours mettre une contrainte sous la forme $g(x) \lozenge 0$, ou le « glyphe mystère » $\lozenge$ peut être $=$ (contrainte d'égalité), $\geq$ ou $\leq$ (contrainte d'inégalité), $>$ ou $<$ (contrainte d'inégalité stricte).

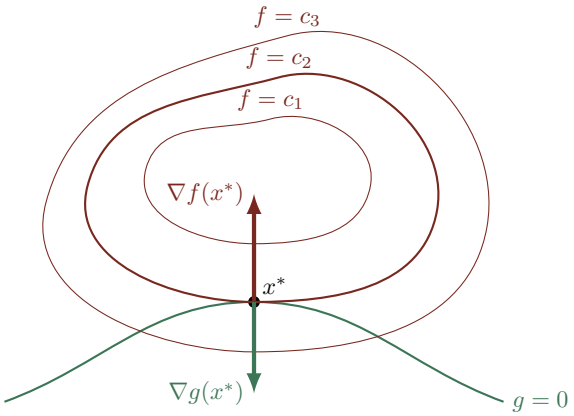
Définition 8.2 (Contrainte active) Si $x \in F$, une contrainte g est active en x si $g(x) = 0$.

Théorème 8.3 (Théorème des multiplicateurs de Lagrange) On suppose la fonction objectif et les contraintes de classe C^1. Si x^* est une solution optimale et les contraintes actives en x^* sont $g_1, \ldots, g_\ell$,

alors la famille $\{\nabla f(x^*), \nabla g_1(x^*), \dots, \nabla g_\ell(x^*)\}$ est liée. ⚠ Le théorème n'affirme pas que x^* existe.

Implémentation de la méthodes des multiplicateurs de Lagrange

- On élargit l'ensemble F en rendant larges les contraintes strictes. ⚠ But : rendre F fermé
- On établit l'existence d'une solution optimale. ⚠ Outils possibles : théorème des bornes atteintes, ou minimum d'une fonction coercive
- On résout tous les systèmes $[g_1(x^*) = 0, \dots, g_\ell(x^*) = 0, \{\nabla f(x^*), \nabla g_1(x^*), \dots, \nabla g_\ell(x^*)\}$ liée$]$. ⚠ On obtient tous les candidats x^*
- On compare les valeurs $f(x^*)$. ⚠ On obtient la solution optimale et l'optimum. ⚠ Au passage, on élimine tous les x^* qui ne respectent pas les contraintes strictes



Interprétation géométrique du théorème des multiplicateurs de Lagrange dans le cas d'une seule contrainte d'égalité $g = 0$. Tant que la ligne de niveau de f n'est pas tangente à la ligne $g = 0$, on peut augmenter ou diminuer f tout en respectant la contrainte $g = 0$. En la solution optimale x^*, les lignes de niveau sont tangentes. Les gradients étant orthogonaux aux lignes de niveau, ils sont colinéaires en x^*.

Autour du théorème des multiplicateurs de Lagrange

- Pourquoi « multiplicateurs de Lagrange » ? Si en x^* aucune contrainte n'est active, le théorème de Fermat donne $\nabla f(x^*) = 0$. Si les contraintes actives sont $g_1, \dots, g_\ell$, souvent les vecteurs $\nabla g_1(x^*), \dots, \nabla g_\ell(x^*)$ sont libres. Dans ce cas, le théorème affirme qu'il existe des scalaires $\lambda_1, \dots, \lambda_\ell$ (les multiplicateurs de Lagrange) tels que $\nabla f(x^*) = \lambda_1 \nabla g_1(x^*) + \dots + \lambda_\ell \nabla g_\ell(x^*)$
- Le théorème des multiplicateurs de Lagrange est admis. Une preuve basée sur des arguments géométriques sera vue en L3
- Une preuve hors programme, de niveau Analyse IV, est présentée en complément de ce cours. Elle permet de voir comment apparaissent les multiplicateurs
- ⚠ Le théorème reste vrai si l'hypothèse x^* solution optimale (c'est-à-dire point d'extrémum de $f_{|F}$) est remplacée par x^* point d'extrémum local de $f_{|F}$
- Les systèmes à résoudre, en général nonlinéaires, sont à :
  - $n + \ell + 1$ inconnues : les n coordonnées de x^* et les $\ell + 1$ scalaires inconnus $\lambda_0, \dots, \lambda_\ell$ qui apparaissent dans la relation de dépendance linéaire (1) $\lambda_0 \nabla f(x^*) + \sum_{j=1}^{\ell} \lambda_\ell \nabla g_j(x^*) = 0$
  - $n + \ell$ équations : n équations correspondant à (1), et les ℓ équations $g_j(x^*) = 0$, $j = 1, \dots, \ell$
- ⚠ On peut rendre les systèmes « carrés » en prenant l'un des scalaires λ_j égal à 1

Quiz

Parmi les cinq problèmes de l'exercice CM8.1, lesquels ne relèvent pas de l'Analyse IV ?

Exercice travaillé

Exercice CM8.4 Parmi les triangles rectangles de périmètre 1, en trouver un d'aire maximale.

Esquisse de solution de l'exercice CM8.4

Nous cherchons une solution optimale de $\max\{f(a,b)\,;\,(a,b)\in F\}$, où : $F := \{(a,b)\in\mathbb{R}^2\,;\,g_1(a,b)=0,\ g_2(a,b)>0, g_3(a,b)>0\}$, $f(a,b) := ab/2$, $g_1(a,b) := a+b+\sqrt{a^2+b^2}$, $g_2(a,b) := a$, $g_3(a,b) := b$.
Étape 1. Nous remplaçons F par $G := \{(a,b)\in\mathbb{R}^2\,;\,g_1(a,b)=0, g_2(a,b)\geq 0, g_3(a,b)\geq 0\}$, qui est fermé et borné ($(a,b)\in G \Rightarrow 0\leq a,b\leq 1$). Sur G, le problème a une solution optimale (x,y) (théorème des bornes atteintes).
Étape 2. En (x,y), nous avons $x,y>0$, sinon $f(x,y)=0$. Il s'ensuit que $\nabla f(x,y)$ et $\nabla g_1(x,y)$ sont colinéaires, d'où $\begin{vmatrix}1/2 & 1+x/\rho\\ 1/2 & 1+y/\rho\end{vmatrix}=0$ (avec $\rho := \sqrt{x^2+y^2}$), d'où $x=y>0$. La contrainte $g_1(x,y)=0$ donne $x=y=1-\sqrt{2}/2$, correspondant à l'unique triangle rectangle isocèle de périmètre 1. *CQFD*

Exercice travaillé

Exercice CM8.5 Parmi les parallélépipèdes rectangles de volume 1, en trouver un de surface latérale minimale.

Esquisse de solution de l'exercice CM8.5

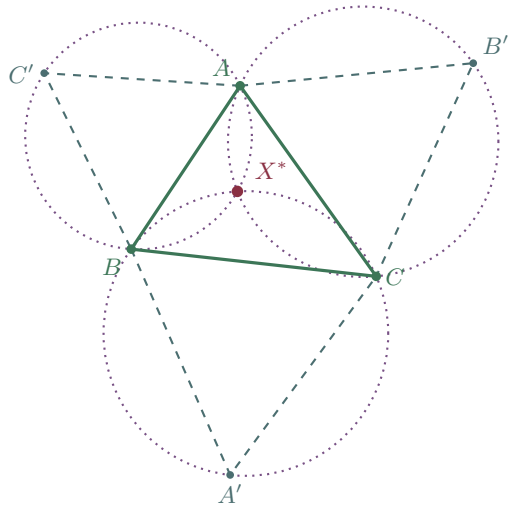
Nous cherchons une solution optimale de $\min\{f(a,b,c)\,;\,(a,b,c)\in F\}$, avec $F := \{(a,b,c)\in\mathbb{R}^3\,;\,g_1=0, g_j>0, j=2,3,4\}$, $f(a,b,c) := 2(ab+bc+ca)$, $g_1(a,b,c) := abc-1$, $g_2(a,b,c) := a$, $g_3(a,b,c) := b$, $g_4(a,b,c) := c$.
F est fermé, car la contrainte $g_1(a,b,c)=0$ fait que $F=\{(a,b,c)\in\mathbb{R}^3\,;\,g_1=0, g_j\geq 0, j=2,3,4\}$.
Étape 1. f est coercive sur F. En effet, si, par exemple, $a\geq b\geq c$ et $\|(a,b,c)\|_2 = R$, alors $a\geq R/\sqrt{3}$. La contrainte $abc=1$ donne $bc\leq \sqrt{3}/R$, d'où $c\leq \sqrt[4]{3}/\sqrt{R}$, d'où $ab\geq \sqrt{R}/\sqrt[4]{3}$, d'où $f(a,b,c)\geq 2\sqrt{R}/\sqrt[4]{3}$. f étant par ailleurs clairement continue, il existe une solution optimale (x,y,z).
Étape 2. En (x,y,z), $\nabla f(x,y,z)$ et $\nabla g_1(x,y,z)\neq 0_{\mathbb{R}^3}$ sont colinéaires. Il existe donc $\lambda\in\mathbb{R}$ tel que $\nabla f(x,y,z)=\lambda\nabla g_1(x,y,z)$. En combinant ce système avec l'égalité $g_1(x,y,z)=0$, nous obtenons $xy=yz=zx$. En utilisant à nouveau la contrainte $g_1(x,y,z)=0$, nous obtenons $x=y=z=1$. La seule solution optimale est le cube unité. *CQFD*

Exercice travaillé

Exercice CM8.6 Trouver l'emplacement idéal d'une mangeoire pour oiseaux nichant dans trois zones de forêt.

Grandes lignes de la solution de l'exercice CM8.6

Nous cherchons une solution optimale de $\min\{f(X)\,;\,X\in\mathbb{R}^2\}$, avec $f(X) := a\|X-A\|_2+b\|X-B\|_2+c\|X-C\|_2$. Les données sont $a,b,c>0$ et A,B,C trois points distincts de $\mathbb{R}^2$.
Étape 1. f est coercive, car $f(X)\geq (a+b+c)\|X\|_2 - a\|A\|_2 - b\|B\|_2 - c\|C\|_2$. f étant par ailleurs clairement continue, il existe une solution optimale X^*.
Étape 2. f est strictement convexe sur $\mathbb{R}^2$: si $X,Y\in\mathbb{R}^2$, $t\in]0,1[$ et $X\neq Y$, alors $f((1-t)X+tY)<(1-t)f(X)+tf(Y)$. Il s'ensuit que f a exactement un point de minimum X^*. La fonction f étant différentiable sur $\mathbb{R}^2\setminus\{A,B,C\}$, soit X^* est l'un des points A,B,C, soit X^* est un point critique de f : $\nabla f(X^*)=0$. Cette équation équivaut à (1) $a(X-A)/\|X-A\|_2+b(X-B)/\|X-B\|_2+c(X-C)/\|X-C\|_2=0_{\mathbb{R}^2}$ et a au plus une solution (par convexité stricte de f). Il y a donc quatre candidats possibles : A, B, C et l'éventuelle solution de (1).
Complément. Si $a=b=c$, (1) a l'interprétation géométrique suivante : $\widehat{AXB}=\widehat{BXC}=\widehat{CXA}=2\pi/3$. On peut construire de manière géométrique le point « de Fermat » X^* avec cette propriété. *CQFD*



Cas où $a = b = c$. La condition $\nabla f(X^*) = 0$ revient à (2) $\omega_A + \omega_B + \omega_C = 0$, où $\omega_A := (X^* - A)/\|X^* - A\|_2$, etc. Les vecteurs $\omega_A, \omega_B, \omega_C$ étant unitaires, (2) implique que les angles entre ces vecteurs valent $2\pi/3$. Comme $\widehat{AX^*B} = 2\pi/3$, X^* doit se trouver sur le petit arc $\widehat{AB}$ du cercle circonscrit au triangle équilatéral ABC'. De même pour les deux autres triangles équilatéraux du dessin. X^* est donc le point d'intersection des trois petits arcs. NB. Ce raisonnement « sur le dessin » n'est correct que si les angles du triangle ABC sont $< 2\pi/3$. Dans ce cas, X^* est bien la solution optimale. Pour les autres triangles, la solution optimale est l'un des sommets.

Construction du point de Fermat X^* pour un triangle ABC « pas trop obtus »

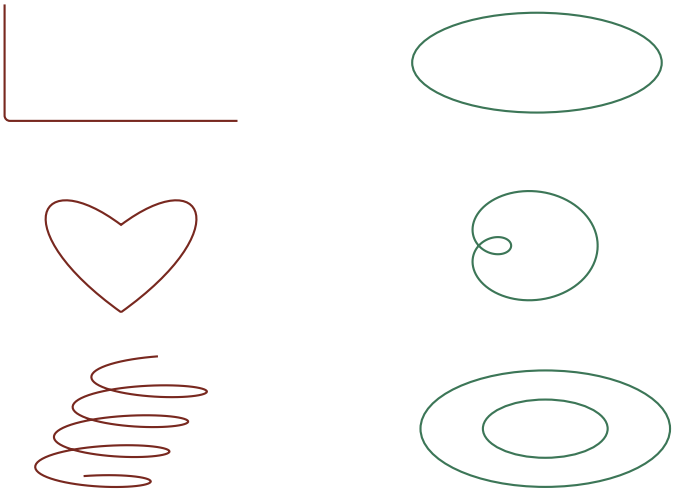
Leçon 9. Arcs paramétrés. Courbes

Objectifs

Comprendre les notions de droite tangente, abscisse curviligne et repère de Frenet pour une courbe paramétrée, et savoir expliciter ces objets. Comprendre la notion de dérivation le long d'un arc paramétré, et son heuristique. Savoir dériver le long d'un arc paramétré.

Le contexte

- En cinématique, un arc paramétré décrit le mouvement d'un point matériel : $t \mapsto M(t)$, avec $t \in I \subset \mathbb{R}$ intervalle (de temps) et $M(t) \in \mathbb{R}^3, \mathbb{R}^2$ ou $\mathbb{R}$
- En géométrie, une courbe est un objet géométrique « lisse » qui « ressemble », près de chacun de ces points, à une droite
- Liens entre les deux : (i) des fois, la « matérialisation » du mouvement, c'est-à-dire l'ensemble $\{M(t)\,;\, t \in I\}$ est une courbe lisse ; (ii) une courbe lisse « faite d'une seule pièce » est un arc paramétré
- Dans cette leçon, nous allons suivre plutôt le point de vue cinématique, en étudiant les arcs paramétrés, tout en nous rapprochant de l'étude géométrique des courbes, sans néanmoins en donner leur définition. (Pour la définition des courbes, attendre la L3 et le M1.)



Quiz. Parmi ces dessins, lesquels peuvent représenter des arcs paramétrés et/ou des courbes ?

Cadre pour l'étude des arcs paramétrés

— $I \subset \mathbb{R}$ intervalle, $\varphi : I \to \mathbb{R}^m$
— Exemple typique : I intervalle de temps, $m = 3$, $\varphi(t)$ la position d'un point matériel à l'instant t
— ⚠ L'intervalle I n'est pas nécessairement ouvert, mais n'est pas réduit à un point

Arc paramétré, courbe support

Définition 9.1 Un arc paramétré différentiable (ou de classe C^k) est une application $\varphi : I \to \mathbb{R}^m$, différentiable ou de classe C^k.

— ⚠ Différentiable = dérivable. L'hypothèse sur φ est que chacune de ses coordonnées est dérivable ou C^k
— Dans la « vraie vie », φ n'est pas nécessairement dérivable (on peut changer de direction !), mais est au moins continue et souvent dérivable par morceaux

Définition 9.2 La courbe (géométrique) support de φ est $\mathscr{C} := \varphi(I)$.

— ⚠ $\mathscr{C}$ porte mal son nom : si φ est constante, $\mathscr{C}$ est un singleton, ce qui ne correspond pas à l'intuition d'une courbe

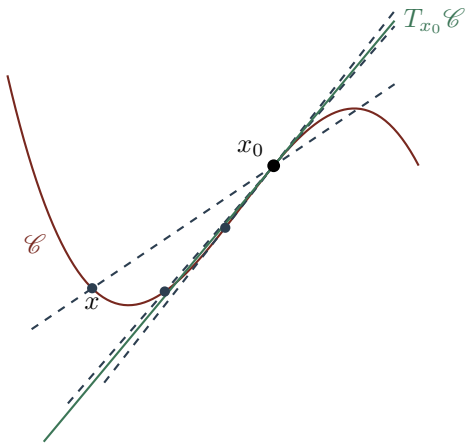
Quiz

Identifier les courbes supports des arcs paramétrés suivants.
— $I = \mathbb{R}$, $m = 2$, $\varphi(t) := (2t + 1, t - 1)$, $\forall\, t \in \mathbb{R}$
— $m = 2$, $\varphi(t) := (t, f(t))$, $\forall\, t \in I$, où $f : I \to \mathbb{R}$
— $I = \mathbb{R}$, $m = 2$, $\varphi(\theta) := (\cos\theta, \sin\theta)$, $\forall\, \theta \in \mathbb{R}$
— $I = \mathbb{R}$, $m = 3$, $\varphi(t) := (\cos t, \sin t, t)$, $\forall\, t \in \mathbb{R}$

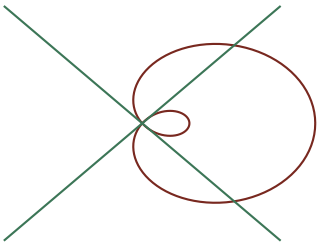
Arcs paramétrés injectifs. Point régulier, point singulier, droite tangente

Pour éviter des cas compliqués comme celui des limaçons, nous considérons dans un premier temps des arcs paramétrés injectifs : φ est injective.

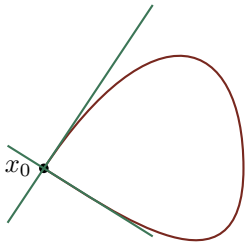
Définition 9.3 (Point régulier, point singulier, droite tangente) Soit $\varphi : I \to \mathbb{R}^m$ une fonction dérivable et injective. Pour $x_0 \in \mathscr{C}$, soit $t_0 := \varphi^{-1}(x_0)$.
— x_0 est un point régulier si $\varphi'(t_0) \neq 0$, singulier si $\varphi'(t_0) = 0$
— Si x_0 est un point régulier de $\mathscr{C}$, la tangente $T_{x_0}\mathscr{C}$ à $\mathscr{C}$ au point x_0 est la droite qui passe par x_0, de directeur $\varphi'(t_0)$ (càd $\{x_0 + s\varphi'(t_0)\,;\, s \in \mathbb{R}\}$)



Interprétation géométrique : la droite tangente est la limite des droites x_0x, avec $x \in \mathscr{C}$ et $x \to x_0$
Interprétation cinématique : $\varphi'(t_0)$ est la vitesse instantanée, $T_{x_0}\mathscr{C}$ est la trajectoire approchée



Au point double, la tangente n'est pas uniquement définie



Quiz. Quelle est la tangente en x_0 ?

Exercice travaillé

Exercice CM9.4 Soit $\mathscr{C}$ le cercle unité du plan, paramétré par $[0, 2\pi[\ni \theta \mapsto \varphi(\theta) \coloneqq (\cos\theta, \sin\theta)$. Déterminer la tangente $T_{(x,y)}\mathscr{C}$, où $(x, y) \in \mathscr{C}$.

Solution de l'exercice CM9.4

Avec $\theta \in [0, 2\pi[$ tel que $x = \cos\theta$, $y = \sin\theta$, nous avons $\varphi'(\theta) = (-\sin\theta, \cos\theta) = (-y, x)$. Sous forme paramétrée, $T_{(x,y)}\mathscr{C} = \{(x - sy, y + sx), s \in \mathbb{R}\}$. *CQFD*

Quiz

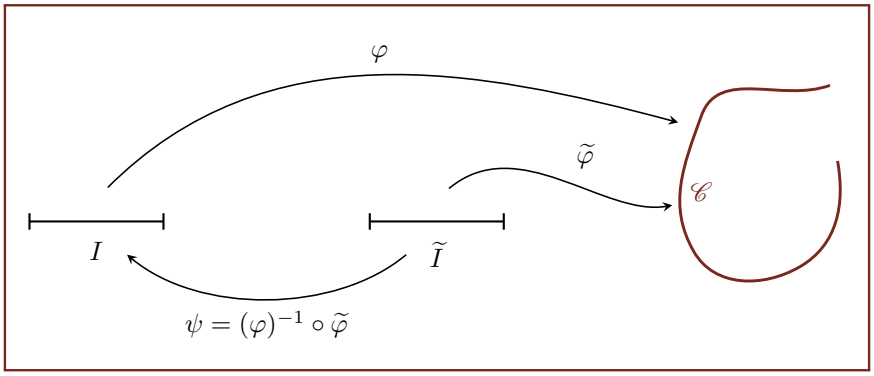
— Quelque chose d'anormal dans l'exercice précédent et sa solution ?
— Proposer une interprétation géométrique du vecteur directeur $\varphi'(t_0)$ de $T_{x_0}\mathscr{C}$

Critère de pertinence géométrique

— Un objet ou nombre « géométrique » lié à $\mathscr{C}$ ne doit pas dépendre du choix de la paramétrisation
— ⚠ Ce critère donne une condition **nécessaire** pour qu'une notion soit « géométrique »

Reparamétrisation

Définition 9.5 Une reparamétrisation est une application $\psi : \widetilde{I} \to I$, avec $\widetilde{I} \subset \mathbb{R}$ intervalle et ψ « difféomorphisme » (ψ bijective, ψ et ψ^{-1} dérivables, ou C^k)



Une reparamétrisation ψ correspond à un changement de paramétrisation :
$\widetilde{\varphi} \coloneqq \varphi \circ \psi : \widetilde{I} \to \mathscr{C}$ est une nouvelle paramétrisation de $\mathscr{C}$

Quiz

Existe-t-il toujours une reparamétrisation entre deux intervalles I et $\widetilde{I}$?

Reformulation du critère de pertinence géométrique

Une condition nécessaire pour la pertinence géométrique d'un nombre ou objet est son « invariance par reparamétrisation » ⚠ C'est-à-dire ?

Exemple d'invariance par reparamétrisation

Proposition 9.6 Soient $\varphi : I \to \mathbb{R}^m$ dérivable et injective et $\mathscr{C} := \varphi(I)$. En un point régulier x_0, $T_{x_0}\mathscr{C}$ est invariante par reparamétrisation.

Preuve de la proposition 9.6

Soient $t_0 := \varphi^{-1}(x_0) \in I$, $s_0 := \psi^{-1}(t_0) \in \widetilde{I}$. Par la règle de la chaîne pour les fonctions réelles, $(\varphi \circ \psi)'(s_0) = \psi'(s_0)\varphi'(t_0)$. Comme $\psi'(s_0) \neq 0$ (car ψ est un difféomorphisme ; voir l'exercice TD6.3), le vecteur $(\varphi \circ \psi)'(s_0)$ est non-nul, et colinéaire avec $\varphi'(t_0)$. Il s'ensuit que les droites qui passent par x_0, de directeur $\varphi'(t_0)$, respectivement $(\varphi \circ \psi)'(s_0)$, coïncident. *CQFD*

Vitesse (scalaire instantanée)

Définition 9.7 La vitesse (scalaire instantanée) à l'instant $t \in I$ est $\|\varphi'(t)\|_2$.

Abscisse curviligne

Cadre de cette partie : $\varphi : I \to \mathbb{R}^m$, $\varphi \in C^1$, φ injective.

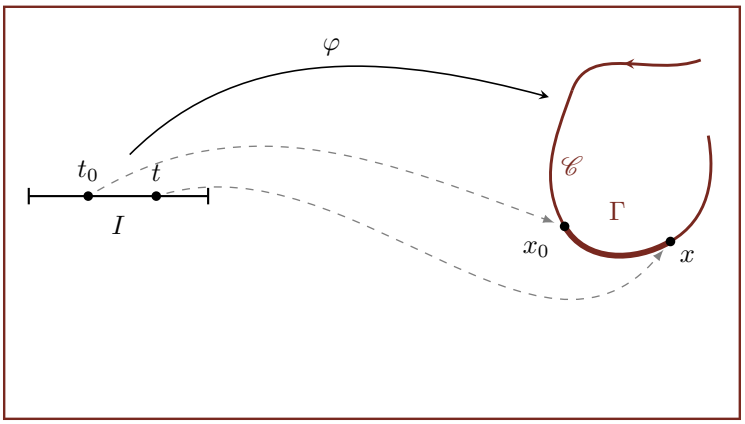
Définition 9.8 On fixe arbitrairement un point « origine » $x_0 = \varphi(t_0) \in \mathscr{C}$. L'abscisse curviligne en $x = \varphi(t)$ est $\ell(x) := \displaystyle\int_{t_0}^{t} \|\varphi'(s)\|_2 \, ds$. ⚠ L'abscisse peut être négative

- ⚠ La définition dépend du choix de l'origine x_0. Exemple : avec la paramétrisation $t \mapsto t$ de la droite réelle et x_0 comme origine, $\ell(x) = x - x_0$, $\forall\, x \in \mathbb{R}$
- ⚠ Même une fois l'origine fixée, l'abscisse dépend du sens du parcours : pour les paramétrisations $t \mapsto t$, respectivement $t \mapsto -t$ de $\mathbb{R}$ et l'origine 0, nous avons $\ell(x) = x$, respectivement $\ell(x) = -x$

Proposition 9.9 (Abscisse curviligne et reparamétrisation) Si $\psi \in C^1$ est une reparamétrisation, l'abscisse curviligne $\widetilde{\ell}$ qui correspond à $\widetilde{\varphi} := \varphi \circ \psi$ et à une origine $x_1 \in \mathscr{C}$ est (1) $\widetilde{\ell}(x) = \pm(\ell(x) - \ell(x_1))$, avec le choix + si ψ est croissante, respectivement - si ψ est décroissante. ⚠ ψ est strictement monotone

Corollaire 9.10 À origine fixée, l'abscisse curviligne est invariante par reparamétrisation C^1 croissante.

Corollaire 9.11 (Reparamétrisation par la longueur de l'arc) Si tous les points de $\mathscr{C}$ sont réguliers, il existe une paramétrisation $\widetilde{\varphi} \in C^1$ de $\mathscr{C}$ pour laquelle $x_1 = \widetilde{\varphi}(0)$, $\widetilde{\ell}(\widetilde{\varphi}(u)) = u$ et $\|\widetilde{\varphi}'(u)\|_2 = 1$, $\forall\, u \in \widetilde{I}$.



$\ell(x)$ est la longueur de l'arc Γ entre x_0 et x, avec le signe $+$ si le sens de parcours de $\mathscr{C}$ est de x_0 vers x, $-$ sinon. Pour la reparamétrisation par la longueur de l'arc, $\widetilde{\ell}(\widetilde{\varphi}(u)) = u$, et $\mathscr{C}$ est parcourue à vitesse constante 1

Interprétations géométriques

— L'abscisse curviligne $x \mapsto \ell(x)$ est $\pm$ la longueur de Γ, l'arc de $\mathscr{C}$ d'extrémités x_0 et x, avec : + si x est après x_0, - si x est x_0 dans le sens de parcours de $\mathscr{C}$ (indiqué par le sens des vecteurs $\varphi'(t)$)
— Par abus de langage, l'abscisse curviligne est désignée comme la « longueur de l'arc » ⚠ Cette longueur peut être négative !
— Reformulation intuitive du corollaire 9.11 : si une paramétrisation φ est C^1, injective, et tous les points de $\mathscr{C}$ sont réguliers, on peut parcourir $\mathscr{C}$ à vitesse scalaire constante 1

Esquisses de preuves de la proposition 9.9 et du corollaire 9.11

Proposition 9.9. Soient $\widetilde{\varphi} := \varphi \circ \psi$ et $\widetilde{\ell}$ l'abscisse curviligne correspondante. Si $x_0 = \widetilde{\varphi}(s_0)$, $x_1 = \widetilde{\varphi}(u_1) = \varphi(t_1)$, et $x = \widetilde{\varphi}(u) = \varphi(t)$, le changement de variable $s = \psi(v)$ et la règle de la chaîne donnent (1) $\ell(x) - \ell(x_1) = \int_{t_1}^{t} \|\varphi'(s)\|_2 \, ds = \int_{u_1}^{u} \|\varphi'(\psi(v))\|_2 \psi'(v) \, dv = \pm \int_{u_1}^{u} \|\widetilde{\varphi}'(v)\|_2 \, dv = \pm\widetilde{\ell}(x)$. *CQFD*

Corollaire 9.11. Soit $F(t) := \ell(\varphi(t)), \forall\, t \in I$. De $F'(t) = \|\varphi'(t)\|_2 > 0$, on obtient que $F : I \to F(I)$ est un difféomorphisme C^1. Soit $x_1 := \varphi(t_1) \in \mathscr{C}$. En utilisant (1), on obtient le corollaire pour $\widetilde{I} := -F(t_1) + F(I)$ et $\psi(u) := F^{-1}(u + F(t_1)), \forall\, u \in \widetilde{I}$. *CQFD*

Quiz

Lorsque l'on parcourt une route « très pentue » de montagne, l'altitude change-t-elle rapidement ?

Dérivation le long d'un arc

Définition 9.12 (Dérivation le long d'un arc) Si $\varphi : I \to V \subset \mathbb{R}^m$ et $f : V \to \mathbb{R}$ sont différentiables, la « dérivée de f le long de l'arc φ » est $\boxed{I \ni t \mapsto (f \circ \varphi)'(t) = [(\nabla f)(\varphi(t))] \cdot \varphi'(t)}$

Interprétation cinématique

— $(f \circ \varphi)'$ mesure la vitesse de changement de f sur $\mathscr{C} = \varphi(I)$ adaptée à φ ⚠ Pas géométrique !
— ⚠ Si $\varepsilon \ll 1$, $\mathbb{R} \ni x \mapsto f(x) := \varepsilon x$ est une fonction qui « varie peu ». Mais, si $\varphi(t) := t/\varepsilon^2$, $(f \circ \varphi)'(t) = 1/\varepsilon \gg 1$. ⚠ $(f \circ \varphi)'$ rend compte des variations de f et de la vitesse de parcours

Exercice travaillé

Exercice CM9.13 Comment un cycliste doit-il adapter son parcours pour monter en deux heures une route de montagne à effort constant ?

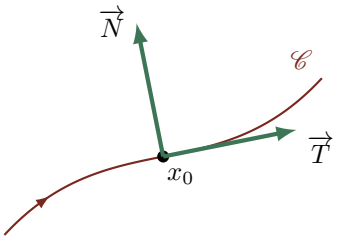
Esquisse de solution de l'exercice CM9.13

Données : un arc paramétré $\varphi : I = [a, b] \to \mathbb{R}^3$, $\varphi(t) = (x(t), y(t), z(t))$. Hypothèses simplificatrices : (i) $\varphi \in C^1$; (ii) $z'(t) > 0$, $\forall\, t \in I$; (iii) pas de frottement (la puissance sert uniquement à compenser l'action de la gravité). Inconnue : une reparamétrisation $\psi : [0, 7200] \to I$ telle que $\psi(0) = a$, $\psi(7200) = b$ et « effort constant ». Modélisation : l'effort est proportionnel à la vitesse de changement de l'altitude z. Nous voulons donc (1) $(z \circ \psi)' = c$. D'où $z(b) - z(a) = 7200\, c$, ce qui donne c. En intégrant (1), $(z \circ \psi)(t) = ct + z(a)$, d'où $\psi(t) = z^{-1}(ct + z(a))$. *CQFD*

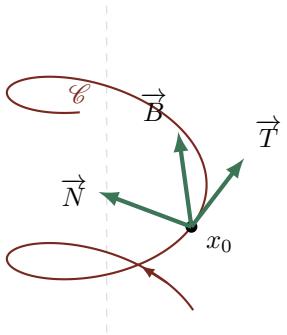
Repère de Frenet d'une courbe plane (du nom de J. F. Frenet (1816–1900))

Définition 9.14 (Repère de Frenet d'une courbe plane) Si $\varphi : I \to \mathbb{R}^2$ est un arc paramétré injectif et $x_0 = \varphi(t_0)$ est un point régulier, le « repère de Frenet » en x_0 est le repère orthonormé direct $(\vec{T}, \vec{N})$ de $\mathbb{R}^2$, où $\vec{T} := \varphi'(t_0)/\|\varphi'(t_0)\|_2$.

— ⚠ $\vec{T}$ est le vecteur vitesse après reparamétrisation pas la longueur de l'arc et indique le sens du parcours de $\mathscr{C}$
— ⚠ Par rapport au sens du parcours sur $\mathscr{C}$, $\vec{N}$ est « à gauche »



Repére de Frenet (courbe plane)



Repére de Frenet (courbe « gauche ») (M1 !)

Invariance du repère de Frenet

Proposition 9.15 Le repère de Frenet est invariant par reparamétrisation croissante.

Preuve de la proposition 9.15

Si $x = \varphi(t) = \widetilde{\varphi}(s)$, alors $\widetilde{\varphi}'(s) = \psi'(s)\varphi'(t)$. Les deux vecteurs sont donc colinéaires et du même sens, d'où $\vec{T}$ est le même pour les deux paramétrisations, et par conséquent $\vec{N}$ l'est également. *CQFD*

Courbe simple fermée

Définition 9.16 Une courbe simple fermée est la courbe support $\mathscr{C}$ de $\varphi : \mathbb{R} \to \mathbb{R}^m$ avec : (i) φ différentiable ou C^k ; (ii) $\varphi'(t) \neq 0, \forall\, t \in \mathbb{R}$; (iii) φ T-périodique pour un $T > 0$; (iv) φ injective sur $[0, T[$.

- ⚠ En M1, cette définition deviendra un théorème : c'est la caractérisation des courbes fermées et connexes (au sens de la géométrie différentielle)
- ⚠ La définition de la droite tangente reste la même : si $x = \varphi(t)$, $T_x\mathscr{C}$ est la droite qui passe par x, de directeur $\varphi'(t)$ ⚠ La définition ne dépend pas du choix de t (exercice TD6.8)

Travail individuel

Exercice CM9.17 (Courbe simple en coordonnées polaires) Soit $r : \mathbb{R} \to]0, \infty[$ une fonction dérivable et 2π-périodique. Montrer que $\{(r(\theta)\cos\theta, r(\theta)\sin\theta)\,;\, \theta \in \mathbb{R}\}$ est une courbe simple fermée.

Complément. Preuve du théorème des multiplicateurs de Lagrange

Extrémum local sous contraintes

Ce complément du cours est dédié à la généralisation suivante du théorème des multiplicateurs de Lagrange.

Théorème 10.1 (Théorème des multiplicateurs de Lagrange) Nous considérons un problème d'optimisation sous contraintes, dont la fonction objectif et les contraintes sont de classe C^1. Si x^* est une extrémum local pour ce problème et les contraintes actives en x^* sont $g_1, \ldots, g_\ell$, alors la famille $\{\nabla f(x^*), \nabla g_1(x^*), \ldots, \nabla g_\ell(x^*)\}$ est liée.

À propos de l'énoncé et de sa preuve

- Rappels : la fonction objectif f et les contraintes g_j sont définies sur un ouvert $U \subset \mathbb{R}^n$. F désigne l'ensemble des points de U qui satisfont les contraintes
- La preuve que nous présentons est due à F. John (1910–1994)
- L'idée principale est de se ramener à un problème sans contrainte, et d'appliquer à ce problème le théorème de Fermat. L'inconvénient de cette approche est que nous ne résolvons plus le

problème initial. La stratégie consiste alors à considérer des problèmes sans contrainte « de plus en plus proches » du problème initial, et d'obtenir le problème initial par passage à la limite
— Dans le jargon des mathématiciens, cela revient à remplacer un problème par un « problème approché ». Pour obtenir à la limite le problème initial, on introduit un paramètre que l'on fait tendre vers 0 ou ∞ (dans notre cas, ∞). Dans le jargon des mathématiciens, il s'agit d'une « méthode de pénalisation », et ce paramètre est un « paramètre de pénalisation ». De telles approches seront approfondies en L3 et master
— Pour faciliter la lecture, la preuve est découpée en étapes dont l'objectif est annoncé au début de chaque étape. L'idée principale de la preuve apparaît dans l'étape 4
— ⚠ Même si chaque étape est de niveau Analyse IV, la complexité de la preuve la rend beaucoup moins facile à suivre que les autres preuves vues dans ce cours

Esquisse de preuve du théorème 10.1

Schéma de la preuve
— Les trois premières étapes permettent de réduire le cas général au cas d'un minimum global sur une boule, avec des contraintes d'égalité et une fonction objectif qui « explose » au bord de la boule. ⚠ Dans ces étapes, le passage du cas général au cas particulier se fait au prix du changement, d'une étape sur l'autre, de U, f, g_j et x^*. Ainsi, à partir de l'étape 4 ces objets ne sont plus ceux du début
— La quatrième étape est au cœur de la preuve : nous y définissons les problèmes approchés
— Les quatre étapes finales permettent de passer à la limite dans le théorème de Fermat appliqué aux problèmes approchés et d'obtenir, comme résultat du passage à la limite, le théorème des multiplicateurs de Lagrange

Étape 1. Réduction au cas d'un minimum global en 0. Quitte à remplacer f par $-f$, nous pouvons supposer qu'il s'agit d'un problème de minimisation. Quitte à remplacer f et les g_j par $x \mapsto f(x+x^*)-f(x^*)$ et $x \mapsto g_j(x+x^*)$, nous pouvons supposer que $x^* = 0$ et $f(0) = 0$. Soit $r > 0$ tel que $B(0,r) \subset U$. Soit $\varepsilon > 0$ tel que $[x \in F, \|x\|_2 < \varepsilon] \Rightarrow f(x) \geq f(0)$. Si $\rho < \min(r, \varepsilon)$, alors, dans $B(0,\rho)$: (i) toutes les fonctions f et g_j sont définies et C^1 ; (ii) f est continue sur $\overline{B}(0,\rho)$; (iii) 0 est une solution optimale (minimum de f dans F) ; (iv) la valeur optimale est 0.

Par la suite, nous supposons $U = B(0,\rho)$ et (i)–(iv). Sous ces hypothèses, la conclusion désirée est : (C) la famille $\{\nabla f(0), \nabla g_1(0), \dots, \nabla g_\ell(0)\}$ est liée.

Étape 2. Réduction au cas d'un minimum global sous contraintes d'égalité. Les contraintes actives en 0 sont $g_1, \dots, g_\ell$. Soient $g_{\ell+1}, \dots, g_m$ les autres contraintes. (S'il n'y a pas ces autres contraintes, prendre $\delta = \rho$ dans ce qui suit.) Comme elles sont inactives en 0, il s'agit de contraintes d'inégalité, et ces inégalités sont strictes en 0. Supposons par exemple que $g_j(0) > 0$. Par continuité de g_j, il existe $\delta_j > 0$ tel que $[x \in B(0,\rho), \|x\|_2 < \delta_j] \Rightarrow g_j(x) > 0$. Si nous posons $\delta := \min(\rho, \delta_{\ell+1}, \dots, \delta_m)$, alors les contraintes $g_{\ell+1}, \dots, g_m$ sont inactives dans $B(0,\delta)$. Par ailleurs, si nous posons $G := \{x \in B(0,\delta)\,;\, g_j(x) = 0,\ j = 1, \dots, \ell\}$, alors $G \subset F$ et donc $\min_{x\in G} f(x) \geq \min_{x\in F} f(x)$. Comme $0 \in G$, nous avons $\min_{x\in F} f(x) = f(0) \leq \min_{x\in G} f(x)$. Nous obtenons que : (v) 0 est le minimum global de f dans $B(0,\delta)$ sous les contraintes $g_j(x) = 0$, $j = 1, \dots, \ell$.

Par la suite, nous supposons $U = B(0,\delta)$ et (i)–(v). La conclusion désirée est toujours (C).

Étape 3. Remplacement de f par une fonction qui « explose » au bord de $B(0,\delta)$. Soit $h : [0,\delta^2[\to [0,\infty[$ une fonction de classe C^1 telle que $h(0) = 0$, $\lim_{t\to\delta^2} h(t) = \infty$. (Un exemple de telle fonction est $h(t) := t/(\delta^2 - t)$, mais la formule exacte de h ne joue aucun rôle dans la preuve.) Considérons le problème de minimisation

$$\min\{u(x)\,;\, x \in B(0,\delta),\, g_j(x) = 0, j = 1, \dots, \ell\}, \text{ où } u(x) := f(x) + h\big((\|x\|_2)^2\big), \forall\, x \in B(0,\delta).$$

Comme $u(x) \geq f(x)$ et $u(0) = f(0)$, le point 0 est solution optimale de ce problème. Comme, par ailleurs, $\nabla u(0) = \nabla f(0)$, montrer (C) équivaut à monter que la famille $\{\nabla u(0), \nabla g_1(0), \dots, \nabla g_\ell(0)\}$ est liée.

Grâce à la propriété (ii), f est bornée sur le compact $\overline{B}(0,\delta)$, et donc, en utilisant le fait que $\lim_{t\to\delta^2} h(t) = \infty$, nous avons $\lim_{\|x\|_2\to\delta} u(x) = \infty$. Ainsi, quitte à remplacer f par u, nous pouvons supposer que (vi) $\lim_{\|x\|_2\to\delta} f(x) = \infty$.

Par la suite, nous supposons $U = B(0,\delta)$ et (i)–(vi). La conclusion désirée est toujours (C).

Étape 4. Construction d'un problème approché. Soit $k \geq 1$ un entier. (k va jouer le rôle du paramètre de pénalisation .) Nous allons définir une fonction F_k qui combine la fonction objectif f et les contraintes g_j,

$j = 1, \dots, \ell$, de sorte que, lorsque $k \to \infty$, un point de minimum de F_k soit « contraint » de tendre vers 0. Un choix possible est le suivant :

$$F_k(x) := f(x) + \frac{1}{2}(\|x\|_2)^2 + \frac{k}{2}\sum_{j=1}^{\ell}(g_j(x))^2, \forall\, x \in B(0,\delta).$$

Étape 5. f est minorée. Soit $0 < r < \delta$ tel que $r < \|x\|_2 < \delta \Rightarrow f(x) > f(0) = 0$. (L'existence de r découle de la propriété (vi) de f.) Pour un tel f, nous avons $\min_{x\in B(0,\delta)} f(x) = \min_{\overline{B}(0,r)} f(x) := m > -\infty$.

Étape 6. F_k a un point de minimum x^k. En effet, avec r comme dans l'étape précédente, nous avons $r < \|x\|_2 < \delta \Rightarrow F_k(x) \geq f(x) > 0 = F_k(0)$. Il s'ensuit que $\min_{x\in B(0,\delta)} F_k(x) = \min_{\overline{B}(0,r)} F_k(x)$, et ce dernier est atteint en (au moins) un point x^k.

Étape 7. $x^k \to 0$. Preuve par l'absurde. Sinon, il existe $t > 0$ et une suite extraite telle que $\left\|x^{\varphi(k)}\right\|_2 \geq t$, $\forall\, k$. Comme par ailleurs $\left\|x^k\right\|_2 \leq r$, $\forall\, k$, il existe une extraction ψ et un $x \in \mathbb{R}^n$ tels que $t \leq \|x\|_2 \leq r$ et $x^{\varphi(\psi(k))} \to x$. De

$$0 = F_k(0) \geq F_k(x^k) = f(x^k) + \frac{1}{2}\left(\|x^k\|_2\right)^2 + \frac{k}{2}\sum_{j=1}^{\ell}(g_j(x^k))^2 \geq m + \frac{k}{2}\sum_{j=1}^{\ell}(g_j(x^k))^2, \tag{1}$$

nous obtenons

$$\sum_{j=1}^{\ell}(g_j(x^k))^2 \leq -\frac{2m}{k},$$

d'où $g_j(x^k) \to 0$, $\forall\, j$. Il s'ensuit que $g_j(x) = \lim_k g_j(x^{\varphi(\psi(k))}) = 0$, $\forall\, j$. Ainsi, x appartient à F. Par conséquent, $f(x) \geq f(0) = 0$. En reprenant (1), nous avons

$$0 = F_k(0) \geq f(x^{\varphi(\psi(k))}) + \frac{1}{2}\left(\|x^{\varphi(\psi(k))}\|_2\right)^2 \to f(x) + \frac{1}{2}(\|x\|_2)^2 > f(0)\,\text{↯}$$

Étape 8. Application du théorème de Fermat et passage à la limite. Le théorème de Fermat donne

$$\nabla f(x^k) + x^k + \sum_{j=1}^{\ell}(kg_j(x^k))\nabla g_j(x^k) = 0. \tag{2}$$

Soit $\mu_k := \sqrt{1 + \sum_{j=1}^{\ell}(kg_j(x^k))^2} \geq 1$. En divisant (2) par μ_k, nous obtenons

$$\lambda_0^k \nabla f(x^k) + \sum_{j=1}^{\ell}\lambda_j^k \nabla g_j(x^k) = -\frac{1}{\mu_k}x^k, \tag{3}$$

avec

$$\lambda_0^k := 1/\mu_k,\ \lambda_j^k := kg_j(x^k)/\mu_k,\ j = 1, \dots, \ell,\ \sum_{j=0}^{\ell}(\lambda_j^k)^2 = 1.$$

Il existe donc une extraction ζ et des nombres $\lambda_0, \dots, \lambda_\ell$ tels que

$$\lambda_j^{\zeta(k)} \to \lambda_j,\ j = 0, \dots, \ell,\ \sum_{j=0}^{\ell}(\lambda_j)^2 = 1. \tag{4}$$

En utilisant (4), le fait que $x^k \to 0$ et l'inégalité $\mu_k \geq 1$, nous obtenons, par passage à la limite dans (3),

$$\lambda_0 \nabla f(0) + \sum_{j=1}^{\ell}\lambda_j \nabla g_j(0) = 0,\ \sum_{j=0}^{k}(\lambda_j)^2 = 1. \qquad \textit{CQFD}$$

Travail individuel

L'exercice suivant propose une preuve alternative du théorème 10.1 à partir de l'étape 3 : au lieu de se ramener au cas d'une fonction qui « explose » au bord de $B(0,\delta)$, on se ramène au cas d'une fonction qui est nulle près du bord de $B(0,\delta)$.

Exercice CM10.2 Nous commençons à la fin de l'étape 2 de la preuve du théorème 10.1.

- Montrer qu'il existe une fonction $\zeta : B(0,\delta) \to \mathbb{R}$ telle que : (i) $\zeta \in C^1$; (ii) $0 \leq \zeta(x) \leq 1$, $\forall\, x \in B(0,\delta)$; (iii) $\|x\|_2 \leq \delta/3 \Rightarrow \zeta(x) = 1$; (iv) $2\delta/3 \leq \|x\|_2 < \delta \Rightarrow \zeta(x) = 0$. (Indication : commencer par le cas $n = 1$, puis remplacer la variable $x \in]-\delta,\delta[$ par $\|x\|_2$)
- Montrer que 0 est solution optimale du problème de minimisation

$$\min\{u(x)\,;\, x \in B(0,\delta),\, g_j(x) = 0, j = 1,\dots,\ell\},\ \text{où } u(x) \coloneqq \zeta(x) f(x), \forall\, x \in B(0,\delta).$$

  Ainsi, quitte à remplacer f par ζf, nous pouvons supposer que (vi) $f(x) = 0$ si $\delta/3 \leq \|x\|_2 < \delta$.

  L'étape 4 reste la même
- Montrer l'étape 5. Par ailleurs, montrer que $\delta/3 \leq \|x\|_2 < \delta \Rightarrow F_k(x) \geq 2\delta^2/9 > F_k(0)$
- Compléter la preuve du théorème 10.1 en adaptant les étapes 6 et 7. (L'étape 8 reste la même.)

“ Some bookes are to be tasted, others to be swallowed, and some few to be chewed and digested.” F. Bacon, Essayes or Counsels, Civill and Morall (1625)
« On se persuade mieux pour l'ordinaire par les raisons qu'on a soi-même trouvées que par celles qui sont venues dans l'esprit des autres. » B. Pascal, Pensées (1670)
« [Le] travail inconscient [...] n'est [...] fécond que s'il est [...] précédé et [...] suivi d'une période de travail conscient. [Les] inspirations subites ne se produisent qu'après quelques jours d'efforts volontaires, qui ont paru absolument infructueux [...] Ces efforts [...] ont mis en branle la machine inconsciente et sans eux [...] elle n'aurait rien produit. » H. Poincaré, Science et méthode (1908)

Cette preuve est l'une des toutes premières preuves longues et complexes que vous rencontrez. Sa maîtrise demande de la patience. C'est le temps nécessaire aux « raisons que l'on a soi-même trouvées » pour germer. Dans un premier temps, vous pouvez vérifier sa validité. Puis, vous réaliserez qu'elle n'est pas « figée », et que d'autres étapes ou d'autres pénalisations sont possibles. À vous de trouver votre preuve, celle qui rend le résultat évident, ou du moins limpide, et permet peut-être d'entrapercevoir la beauté du résultat démontré !

Remerciements

En plus de l'IA générative, plusieurs IN ont contribué à améliorer les notes de cours et les feuilles d'exercices

- Antoine Detaille, Louis Dupaigne et Frédéric Lagoutière ont relu le cours et proposé des améliorations
- Vincent Borrelli et Philippe Malbos ont proposé plusieurs exercices

Index, notations

$A > 0$, 26
$A \geq 0$, 26
$A \leq 0$, 26
$A < 0$, 26
$A^\intercal$, 3
$B(a,r)$, 7
$\overline{B}(a,r)$, 7
$\mathscr{C}$, 34
C^1, 15
df_a, 13
$\frac{\partial f}{\partial x_j}$, 14
$\frac{\partial f}{\partial v}(a)$, 18
$\frac{\partial^2 f}{\partial x \partial y}$, 22
$\partial x_j f$, 14
$(e^1,\dots,e^n)$, 3
$\gamma_{x,y}$, 19
$H_f(a)$, 22
$I_{x,y}$, 19

(IS), 6
(J), 27
$J_f(a)$, 14
$\ell(x)$, 36
$\nabla f(a)$, 14
$O(g(h))$, 12
$o(g(h))$, 12
$S(a,r)$, 7
$T_{x_0}\mathscr{C}$, 34
$v^\intercal$, 3
(x^k), 3
$(x^k) \subset A$, 3
$[x,y]$, 7
$\|x\|_2$, 4
x^α, 9

abscisse curviligne, 36
arc paramétré, 34
    C^k, 34
    différentiable, 34
    droite tangente, 34
    injectif, 34

champ de vecteurs, 3
    potentiel, 24
condition
    nécessaire du premier ordre, 26
    nécessaire et suffisante du premier ordre, 29
    suffisante du second ordre, 26
contrainte, 30
    active, 30
    d'égalité, 30
    d'inégalité, 30
    d'inégalité stricte, 30
courbe
    simple fermée, 38
    support, 34

dérivée
    directionnnelle, 18
    le long d'un arc, 37
    partielle, 14
    partielle d'ordre supérieur, 22
difféomorphisme, 35
différentielle, 13
distance euclidienne, 5

ensemble
    borné, 7
    compact, 7
    convexe, 7
    fermé, 7
    ouvert, 7

fonction
    coercive, 9
    convexe, 27
    deux fois différentiable, 22
    différentiable, 13
    objectif, 30
    partielle, 14
    réelle, 3
    scalaire, 3
formule de Taylor
    à l'ordre 1 avec point intermédiaire, 19
    à l'ordre 1 avec reste en o, 19
    à l'ordre 1 avec reste intégral, 19
    à l'ordre 2 avec point intermédiaire, 24
    à l'ordre 2 avec reste en o, 24
    à l'ordre 2 avec reste intégral, 24

gradient, 14

hessienne, 22

IAF, 19
invariance par paramétrisation, 36
inégalité
    de Cauchy–Schwarz, 3, 4
    de Jensen, 27
    de Jensen généralisée, 28

jacobienne, 14

lemme de Poincaré, 25
longueur de l'arc, 37

matrice
    hessienne, 22
    jacobienne, 14
monôme, 9
multiplicateurs de Lagrange, 31

optimum, 30
ordre de dérivation, 22

point
    col, 26
    critique, 26
    régulier, 34
    -selle, 26
    singulier, 34
principe $r(\varepsilon)$, 12

reparamétrisation, 35
repère de Frenet, 37
règle de la chaîne, 16
    pour les dérivées, 17
    pour les dérivées partielles, 17
règle du produit, 17

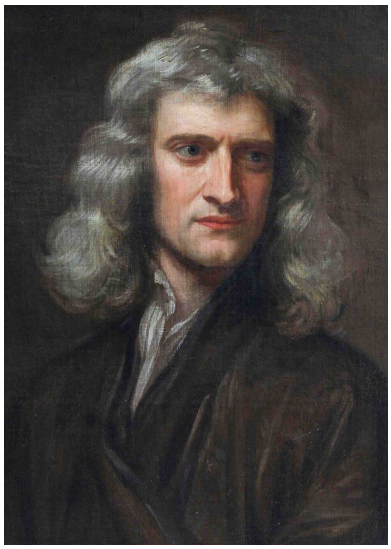
segment, 7
solution optimale, 30

TAF, 19
théorème
    de Bolzano-Weierstrass, 8
    de Fermat, 26
    de Schwarz, 22
    des bornes atteintes, 8
    des multiplicateurs de Lagrange, 30

vitesse
    instantanée, 36
    scalaire, 36



P. de Fermat
(1601–1665)



I. Newton
(1643–1727)



G. W. Leibniz
(1646–1716)



B. Taylor
(1685–1731)



É. du Châtelet
(1706–1749)



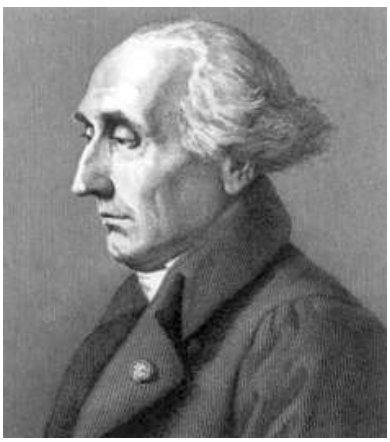
L. Euler
(1707–1783)



J. L. R. d'Alembert
(1717–1783)

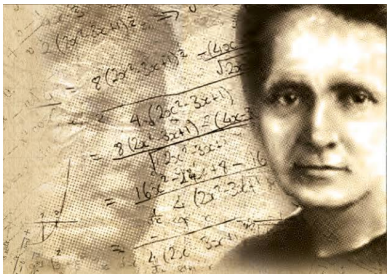


M. G. Agnesi
(1718–1799)



J.-L. Lagrange
(1736–1813)

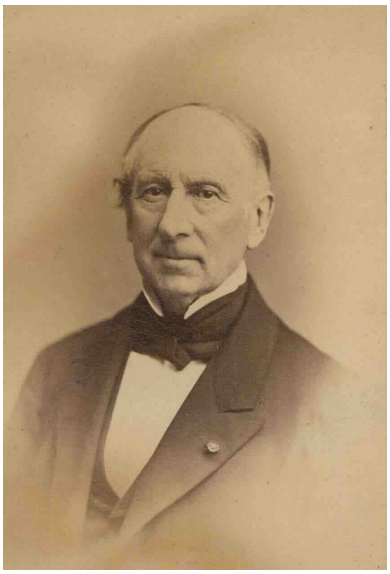
Quelques héros et héroïnes de notre περίπλους au pays du calcul différentiel. Connaissez-vous certaines de leurs contributions ?



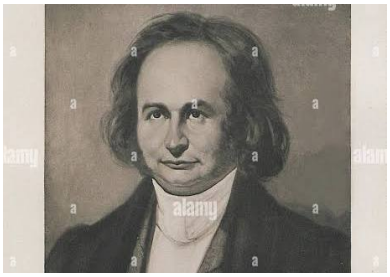
S. Germain
(1776–1831)



C. F. Gauss
(1777–1855)



A.-L. Cauchy
(1789–1857)



C. G. J. Jacobi
(1804–1851)



W. R. Hamilton
(1805–1865)



L. O. Hesse
(1811–1874)



K. Weierstrass
(1815–1897)



J. C. Maxwell
(1831–1879)



O. Stolz
(1842–1905)

Quelques héros et héroïnes de notre περίπλους au pays du calcul différentiel. Connaissez-vous certaines de leurs contributions ?



H. A. Schwarz
(1843–1921)



S. Kovalevskaïa
(1850–1891)



H. Poincaré
(1854–1912)



J. Jensen
(1859–1925)



W. H. Young
(1863–1942)



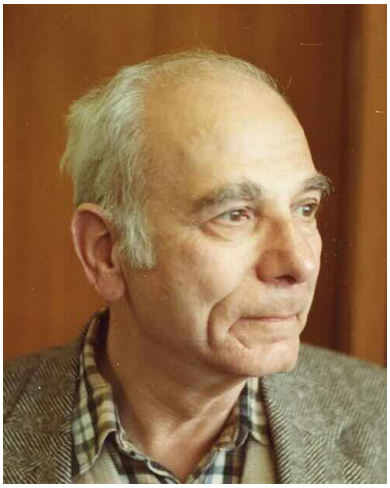
H. Minkowski
(1864–1909)



G. C. Young
(1868–1944)



H. Hopf
(1894–1971)



F. John
(1910–1994)

Quelques héros et héroïnes de notre περίπλους au pays du calcul différentiel. Connaissez-vous certaines de leurs contributions ?