

Calcul différentiel

Résumé de cours – Mathématiques 2026

Formes et Formules

Idée clé

Le calcul différentiel généralise la notion de dérivée aux fonctions de plusieurs variables, à valeurs dans des espaces vectoriels normés. Il permet d'approximer localement une fonction par une application linéaire (la différentielle), d'optimiser des fonctions sous contraintes, d'analyser la géométrie locale des ensembles (hyperplans tangents), d'étudier les lignes de niveaux et de réaliser des changements de variables dans les intégrales multiples.

Table des matières

1	Applications de classe C^1	3
1.1	Différentielle	3
1.2	Dérivation d'applications particulières	4
1.3	Règle de composition des différentielles	5
1.4	Applications : Produit scalaire et Norme	6
1.5	Dérivée selon un vecteur et dérivées partielles	7
1.6	Matrice Jacobienne	8
1.7	Règles de dérivation pratique	10
1.8	Gradient associé à une fonction numérique	11
1.9	Gradient et lignes de niveaux	11
1.10	Dérivées le long d'un arc et vecteurs tangents	13
2	Applications de classe C^k	14
2.1	Dérivées partielles d'ordres supérieurs	14
2.2	Hessienne	14
3	Optimisation	14
3.1	Condition d'ordre 1	14
3.2	Condition d'ordre 2	15
3.3	Optimisation sous contrainte	15
4	Intégrales multiples et changement de variables	16

4.1	Le théorème de changement de variables	16
4.2	Coordonnées cylindriques	16
4.3	Coordonnées sphériques	17

1 Applications de classe C^1

1.1 Différentielle

Définition : Différentielle en un point

Soit $f : U \subset E \rightarrow F$ où E et F désignent deux e.v.n. sur $\mathbb{R}$ de dimensions respectives p et n , et U un ouvert de E .

L'application f est dite **différentiable** en $a \in U$ s'il existe $\varphi \in \mathcal{L}(E, F)$ tel que :

$$\forall h \in E, \quad f(a+h) = f(a) + \varphi(h) + o(\|h\|)$$

L'application linéaire est alors unique, on l'appelle **différentielle de f au point a** . On la note df_a ou $df(a)$.

Formule

Notation des restes : $o(h) = \|h\|\varepsilon(h)$ où $\varepsilon : E \rightarrow F$ et $\lim_{h \rightarrow 0} \varepsilon(h) = 0_F$.

Ainsi, le développement s'écrit :

$$f(a+h) = f(a) + df_a(h) + o(\|h\|)$$

Lien / Remarque : Interprétation dans le cas $E = F = \mathbb{R}$

Lorsque $E = F = \mathbb{R}$, on retrouve la dérivée usuelle d'une fonction $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$. En effet, une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$ est nécessairement de la forme $h \mapsto \lambda \cdot h$ pour un certain $\lambda \in \mathbb{R}$. La différentielle df_a est donc l'application :

$$df_a : h \longmapsto f'(a) \cdot h$$

où $f'(a)$ est la dérivée usuelle de f en a . Le développement limité s'écrit alors :

$$f(a+h) = f(a) + f'(a) \cdot h + o(|h|)$$

ce qui est exactement la formule de Taylor-Young à l'ordre 1 en dimension 1.

Ainsi, la différentielle df_a **généralise** la dérivée $f'(a)$: en dimension 1, elle se réduit à la multiplication par le nombre $f'(a)$, tandis qu'en dimension supérieure, elle devient une application linéaire à part entière (représentée par une matrice jacobienne).

1.2 Dérivation d'applications particulières

Avant d'énoncer les propriétés générales de la différentielle, il est essentiel de savoir calculer la différentielle de quelques applications fondamentales. Ces résultats serviront de briques élémentaires pour tous les calculs ultérieurs.

Formule : Application linéaire

Soit $L \in \mathcal{L}(E, F)$ une application linéaire continue entre deux e.v.n.

Alors L est différentiable en tout point $a \in E$ et sa différentielle est **égale à L elle-même** :

$$\forall a \in E, \quad dL_a = L$$

Démonstration : Pour tout $h \in E$, on a exactement $L(a + h) = L(a) + L(h)$, donc le reste est nul (pas seulement un $o(\|h\|)$). L'application linéaire tangente est donc L elle-même.

Formule : Application bilinéaire

Soit $B : E \times F \rightarrow G$ une application bilinéaire continue entre e.v.n.

Alors B est différentiable en tout point $(a, b) \in E \times F$ et :

$$dB_{(a,b)}(h, k) = B(h, b) + B(a, k)$$

Démonstration : Par bilinéarité :

$$B(a + h, b + k) = B(a, b) + B(h, b) + B(a, k) + B(h, k)$$

Le terme $B(h, k)$ est le reste. Comme B est continue, il existe $C > 0$ tel que $\|B(h, k)\| \leq C\|h\|\|k\| = o(\|(h, k)\|)$, ce qui conclut.

Lien / Remarque : Détail sur le reste $B(h, k) = o(\|(h, k)\|)$

Pourquoi le terme $B(h, k)$ est-il un $o(\|(h, k)\|)$?

En dimension finie, toute application bilinéaire continue B admet une **norme d'opérateur** finie :

$$C = \|B\| = \sup_{\substack{\|a\| \leq 1 \\ \|b\| \leq 1}} \|B(a, b)\| < +\infty$$

(ce supremum est atteint car l'ensemble est compact et B continue).

Pour tout (h, k) non nul, on factorise :

$$B(h, k) = \|h\| \cdot \|k\| \cdot B\left(\frac{h}{\|h\|}, \frac{k}{\|k\|}\right)$$

Les vecteurs normalisés $\frac{h}{\|h\|}$ et $\frac{k}{\|k\|}$ étant de norme 1, on obtient :

$$\|B(h, k)\| \leq C\|h\|\|k\|$$

En munissant $E \times F$ de la norme $\|(h, k)\| = \max(\|h\|, \|k\|)$, on a $\|h\| \leq \|(h, k)\|$ et $\|k\| \leq \|(h, k)\|$, donc :

$$\|B(h, k)\| \leq C\|(h, k)\|^2 \implies \frac{\|B(h, k)\|}{\|(h, k)\|} \leq C\|(h, k)\| \xrightarrow{(h, k) \rightarrow 0} 0$$

Ce qui signifie précisément que $B(h, k) = o(\|(h, k)\|)$.

1.3 Règle de composition des différentielles

La composition est l'outil central qui permet de combiner les briques élémentaires ci-dessus.

Formule : Composition des différentielles

Soient $f : U \subset E \rightarrow F$ différentiable en $a \in U$ et $g : V \subset F \rightarrow G$ différentiable en $f(a) \in V$ (avec $f(U) \subset V$).

Alors $g \circ f$ est différentiable en a et :

$$d(g \circ f)_a = dg_{f(a)} \circ df_a$$

Lien / Remarque : Pourquoi cette formule ?

Cette règle découle directement de la nature de la différentielle comme **approximation linéaire locale**.

- Localement, $f(a + h) \approx f(a) + df_a(h)$.
- De même, $g(y + k) \approx g(y) + dg_y(k)$.
- En composant, on remplace k par la variation induite par f , soit $df_a(h)$:

$$g(f(a + h)) \approx g(f(a) + df_a(h)) \approx g(f(a)) + dg_{f(a)}(df_a(h))$$

Le terme linéaire dominant est bien la composée $dg_{f(a)} \circ df_a$. Cela explique aussi la formule matricielle $J_{g \circ f} = J_g \times J_f$: la matrice de la composée de deux applications linéaires est le produit de leurs matrices.

1.4 Applications : Produit scalaire et Norme

Nous pouvons maintenant appliquer la règle de composition aux structures euclidiennes.

Formule : Produit scalaire

Soient $f, g : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ deux applications différentiables en $a \in U$.

L'application $\varphi : x \mapsto \langle f(x), g(x) \rangle$ est différentiable en a et :

$$d\varphi_a(h) = \langle df_a(h), g(a) \rangle + \langle f(a), dg_a(h) \rangle$$

Démonstration : Le produit scalaire $B(u, v) = \langle u, v \rangle$ est une application bilinéaire continue. On a $\varphi(x) = B(f(x), g(x))$.

En appliquant la formule de la différentielle d'une bilinéaire composée avec $x \mapsto (f(x), g(x))$ (dont la différentielle est $h \mapsto (df_a(h), dg_a(h))$), on obtient directement le résultat.

Formule : Norme euclidienne

Soit E un espace euclidien. La fonction $N : x \mapsto \|x\|$ est différentiable en tout point $a \neq 0_E$ et :

$$dN_a(h) = \frac{\langle a, h \rangle}{\|a\|}$$

Démonstration : On écrit $\|x\| = \sqrt{\langle x, x \rangle}$. En posant $\varphi(x) = \langle x, x \rangle$ et $S(t) = \sqrt{t}$, on a $N = S \circ \varphi$.

Par la formule du produit scalaire ci-dessus : $d\varphi_a(h) = 2\langle a, h \rangle$.

Par la règle de composition : $dN_a(h) = S'(\|a\|^2) \cdot d\varphi_a(h) = \frac{1}{2\|a\|} \cdot 2\langle a, h \rangle = \frac{\langle a, h \rangle}{\|a\|}$.

Attention

La fonction $x \mapsto \|x\|$ n'est **pas différentiable** en 0.

En effet, si elle l'était, sa différentielle en 0 serait une application linéaire L telle que $\|h\| = L(h) + o(\|h\|)$. En remplaçant h par $-h$, on obtiendrait $\|h\| = -L(h) + o(\|h\|)$, donc $L(h) = o(\|h\|)$, ce qui implique $L = 0$. Mais alors $\|h\| = o(\|h\|)$, absurde pour $h \neq 0$.

Géométriquement, la norme présente une **pointe** en 0 (comme la fonction $|x|$ en dimension 1).

Lien / Remarque : Cas particulier en dimension 1

En dimension 1 ($E = \mathbb{R}$), la norme est la valeur absolue $x \mapsto |x|$. On retrouve le résultat classique :

- Pour $a > 0$: $d|x|_a(h) = h$ (car $|x| = x$ au voisinage de a).
- Pour $a < 0$: $d|x|_a(h) = -h$ (car $|x| = -x$ au voisinage de a).
- En $a = 0$: non différentiable (la célèbre « pointe » de la fonction valeur absolue).

La formule $d|x|_a(h) = \frac{a \cdot h}{|a|} = \text{signe}(a) \cdot h$ pour $a \neq 0$ est bien cohérente avec ces cas.

Autres propriétés algébriques :

- Si f est différentiable en a , alors f est continue en a .
- Si f et g sont différentiables en a , alors $f + \mu g$ aussi et : $d(f + \mu g)_a = df_a + \mu dg_a$.

Définition : Application de classe C^1

- f est dite **différentiable sur** U si f est différentiable en tout point de U . On appelle alors différentielle de f l'application :

$$df : U \rightarrow \mathcal{L}(E, F), \quad a \mapsto df_a$$

- L'application $f : U \subset E \rightarrow F$ est dite de **classe C^1** sur U si f est différentiable sur U et si sa différentielle df est continue sur U .

Si f est de classe C^1 sur U , on dira aussi que f est continûment différentiable sur U .

1.5 Dérivée selon un vecteur et dérivées partielles

Définition : Dérivée selon un vecteur

Soit $u \in E$. L'application f est dite **dérivable en a selon le vecteur u** si la fonction $t \mapsto f(a + tu)$ est dérivable en 0. On pose dans ce cas :

$$D_u(f)(a) = \lim_{t \rightarrow 0} \frac{f(a + tu) - f(a)}{t}$$

Lien / Remarque

Lorsqu'une fonction est différentiable, elle est dérivable dans toutes les directions.

Proposition : Si f est différentiable en a , alors f est dérivable en a selon u pour tout vecteur $u \in E$ et $D_u(f)(a) = df_a(u)$.

On munit désormais E d'une base orthonormée $(e_1, \dots, e_p)$.

Définition : Dérivées partielles

Pour $j \in \{1, \dots, p\}$, on appelle **dérivée partielle en a d'indice j** la dérivée de f en a suivant e_j , c'est-à-dire :

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_j) - f(a)}{t}$$

Si f est différentiable en a , alors les dérivées partielles existent et :

$$\forall j \in \{1, \dots, p\}, \quad \frac{\partial f}{\partial x_j}(a) = df_a(e_j)$$

Formule : Expression de la différentielle

Pour tout vecteur $h = \sum_{j=1}^p h_j e_j \in E$, la linéarité de df_a permet d'écrire :

$$df_a(h) = df_a\left(\sum_{j=1}^p h_j e_j\right) = \sum_{j=1}^p h_j df_a(e_j) = \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a)$$

Remarque importante : Ici, chaque $\frac{\partial f}{\partial x_j}(a)$ est un **vecteur de l'espace d'arrivée** F (de dimension n).

- Si $F = \mathbb{R}^n$, ce sont des vecteurs colonnes.
- Si $F = \mathbb{R}$, ce sont des scalaires.

En notant dx_j les formes linéaires coordonnées $h \mapsto h_j$, on a la notation différentielle classique :

$$df_a = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(a) dx_j$$

Théorème de caractérisation : Soit $f : U \subset E \rightarrow F$. f est de classe C^1 sur U si et seulement si les dérivées partielles de f existent et sont continues en tout point de U .

1.6 Matrice Jacobienne

On munit désormais F d'une base orthonormée $(e'_1, \dots, e'_n)$.

1. Définition et calcul

On appelle **jacobienne** de f au point $x = (x_1, \dots, x_p)$ la matrice représentative de l'application linéaire df_x dans les bases (e_j) de E et (e'_i) de F . Elle est de taille $n \times p$:

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_p}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \cdots & \frac{\partial f_n}{\partial x_p}(x) \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

Calcul / Exemple : Exemple : Application de $\mathbb{R}^2$ vers $\mathbb{R}^3$

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par $f(x, y) = (x^2y, x + y, e^{xy})$.

Ici $p = 2$ et $n = 3$. Les composantes sont $f_1 = x^2y$, $f_2 = x + y$, $f_3 = e^{xy}$.

La matrice jacobienne est de taille 3×2 :

$$J_f(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} \end{pmatrix} = \begin{pmatrix} 2xy & x^2 \\ 1 & 1 \\ ye^{xy} & xe^{xy} \end{pmatrix}$$

Au point $(1, 1)$, on a :

$$J_f(1, 1) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ e & e \end{pmatrix}$$

Cette matrice représente l'application linéaire $df_{(1,1)} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$.

2. La différentielle d'une composée (Théorème du rang)

Si f est de classe C^1 sur U et g de classe C^1 sur V (avec $f(U) \subset V$), alors la différentielle de la composée est la composée des différentielles :

$$d(g \circ f)_a = dg_{f(a)} \circ df_a$$

En termes de matrices, cela se traduit par le **produit des matrices jacobienes** :

Formule : Formule de la chaîne (Chain Rule)

$$J_{g \circ f}(x) = J_g(f(x)) \times J_f(x)$$

Attention à l'ordre : La matrice de g (qui arrive en dernier) est à gauche, celle de f (qui arrive en premier) est à droite. Les dimensions doivent être compatibles : si J_f est $n \times p$ et J_g est $m \times n$, le produit est bien défini et donne une matrice $m \times p$.

Calcul / Exemple : Exemple : Composition d'applications

Reprenons $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ de l'exemple précédent. Soit $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ définie par $g(u, v, w) = u + v + w$.

On veut calculer $J_{g \circ f}(1, 1)$.

Méthode 1 : Par le produit des jacobienes. La jacobienne de g est $J_g(u, v, w) = (1 \ 1 \ 1)$ (taille 1×3).

On a vu que $J_f(1, 1) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ e & e \end{pmatrix}$.

Le produit vaut :

$$J_{g \circ f}(1, 1) = J_g(f(1, 1)) \times J_f(1, 1) = (1 \ 1 \ 1) \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ e & e \end{pmatrix} = (3 + e \ 2 + e)$$

Méthode 2 : Par calcul direct. $(g \circ f)(x, y) = x^2y + (x + y) + e^{xy}$.

$\frac{\partial}{\partial x} = 2xy + 1 + ye^{xy}$ et $\frac{\partial}{\partial y} = x^2 + 1 + xe^{xy}$.

En $(1, 1)$: $\frac{\partial}{\partial x} = 2 + 1 + e = 3 + e$ et $\frac{\partial}{\partial y} = 1 + 1 + e = 2 + e$.

On retrouve bien $J_{g \circ f}(1, 1) = (3 + e \ 2 + e)$.

1.7 Règles de dérivation pratique

Formule

Dérivation le long d'un arc : Soient U un ouvert de $\mathbb{R}^2$ et I un intervalle de $\mathbb{R}$. On considère les applications de classe C^1 : $\varphi : I \rightarrow U, t \mapsto (x(t), y(t))$ et $f : U \rightarrow \mathbb{R}, (x, y) \mapsto f(x, y)$.

L'application $t \mapsto f(x(t), y(t))$ est de classe C^1 sur I et pour tout $t \in I$:

$$(f \circ \varphi)'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t))$$

Formule

Changement de variables : On considère les applications $f, \varphi, \psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 et $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $F(x, y) = f(\varphi(x, y), \psi(x, y))$.

Alors F est de classe C^1 sur $\mathbb{R}^2$ et pour tout $(x, y) \in \mathbb{R}^2$:

$$\begin{aligned} \frac{\partial F}{\partial x}(x, y) &= \frac{\partial \varphi}{\partial x} \frac{\partial f}{\partial x}(\varphi, \psi) + \frac{\partial \psi}{\partial x} \frac{\partial f}{\partial y}(\varphi, \psi) \\ \frac{\partial F}{\partial y}(x, y) &= \frac{\partial \varphi}{\partial y} \frac{\partial f}{\partial x}(\varphi, \psi) + \frac{\partial \psi}{\partial y} \frac{\partial f}{\partial y}(\varphi, \psi) \end{aligned}$$

Lien / Remarque : La dérivée de la composée dans $\mathbb{R}$: un cas particulier

En classe de terminale, on apprend la formule bien connue :

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$$

Cette formule n'est **qu'un cas particulier** de la règle générale de composition des différentielles vue plus haut. En effet, lorsque $E = F = \mathbb{R}$, une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$ est simplement une multiplication par un scalaire. Les différentielles df_x et dg_y sont donc les applications $h \mapsto f'(x) \cdot h$ et $k \mapsto g'(y) \cdot k$.

La composée $dg_{f(x)} \circ df_x$ envoie alors h sur :

$$g'(f(x)) \cdot (f'(x) \cdot h) = (g'(f(x)) \cdot f'(x)) \cdot h$$

Ce qui correspond bien à la multiplication par le scalaire $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$.

1.8 Gradient associé à une fonction numérique

Soit $f : U \subset E \rightarrow \mathbb{R}$ une fonction numérique, supposée différentiable en a . On peut définir le **gradient** de f au point a par ses coordonnées dans une b.o.n. $(e_1, \dots, e_p)$:

$$\nabla f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_p}(a) \end{pmatrix}$$

Lien / Remarque

Le théorème de Riesz fournit une définition intrinsèque : on appelle gradient de f en a , noté $\nabla f(a)$, le vecteur associé à la forme linéaire df_a . Pour tout $h \in E$:

$$df_a(h) = \langle \nabla f(a), h \rangle = \nabla f(a) \cdot h$$

Formule

Si $f : U \subset E \rightarrow \mathbb{R}$ est de classe C^1 , alors pour tout $x \in U$:

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + o(\|h\|)$$

1.9 Gradient et lignes de niveaux

Définition : Ligne de niveau

Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction numérique de classe C^1 . Pour tout $c \in \mathbb{R}$, on appelle **ligne de niveau** (ou **surface de niveau** si $n \geq 3$) de f associée à la valeur c l'ensemble :

$$\mathcal{C}_c = \{x \in U \mid f(x) = c\}$$

Propriété géométrique : Orthogonalité du gradient aux lignes de niveau

Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction numérique de classe C^1 et $a \in U$ tel que $\nabla f(a) \neq \vec{0}$.

Alors le gradient $\nabla f(a)$ est **orthogonal** à la ligne de niveau $\mathcal{C}_{f(a)}$ au point a .

Lien / Remarque : Démonstration

Soit $\gamma :]-\varepsilon, \varepsilon[\rightarrow U$ un arc de classe C^1 contenu dans la ligne de niveau $\mathcal{C}_{f(a)}$ et passant par a en $t = 0$, c'est-à-dire :

$$\gamma(0) = a \quad \text{et} \quad \forall t \in]-\varepsilon, \varepsilon[, f(\gamma(t)) = f(a)$$

La fonction $t \mapsto f(\gamma(t))$ est donc constante, égale à $f(a)$. En la dérivant en $t = 0$, on obtient :

$$\left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = 0$$

Or, d'après la règle de dérivation le long d'un arc (vue précédemment) :

$$\left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = df_a(\gamma'(0)) = \langle \nabla f(a), \gamma'(0) \rangle$$

On en déduit que :

$$\langle \nabla f(a), \gamma'(0) \rangle = 0$$

Le vecteur $\gamma'(0)$ est un vecteur tangent à la ligne de niveau en a . Comme ce raisonnement s'applique à tout arc contenu dans la ligne de niveau, le gradient $\nabla f(a)$ est orthogonal à tous les vecteurs tangents à $\mathcal{C}_{f(a)}$ en a , donc orthogonal à la ligne de niveau elle-même.

Idée clé : Interprétation géométrique

Le gradient $\nabla f(a)$ pointe dans la direction de la **plus forte pente** de f au point a . En effet, la dérivée directionnelle de f en a selon un vecteur unitaire u vaut :

$$D_u(f)(a) = \langle \nabla f(a), u \rangle = \|\nabla f(a)\| \cos(\theta)$$

où θ est l'angle entre $\nabla f(a)$ et u . Cette quantité est maximale lorsque $\theta = 0$, c'est-à-dire lorsque u est colinéaire à $\nabla f(a)$.

À l'inverse, le long d'une ligne de niveau, la fonction ne varie pas, donc la dérivée directionnelle est nulle, ce qui correspond exactement à $\theta = \pi/2$: le gradient est perpendiculaire à la ligne de niveau.

Calcul / Exemple : Exemple : fonction $f(x, y) = x^2 + y^2$

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

Le gradient vaut :

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

Les lignes de niveau sont les cercles concentriques $x^2 + y^2 = c$ (pour $c > 0$).

En un point (x_0, y_0) de la ligne de niveau, le gradient vaut $(2x_0, 2y_0)$, qui est bien le vecteur radial, orthogonal au cercle en ce point.

Attention

Si $\nabla f(a) = \vec{0}$, on ne peut rien dire sur l'orthogonalité : le point a est un point critique, et la ligne de niveau peut présenter une singularité (point double, rebroussement, etc.) en a .

1.10 Dérivées le long d'un arc et vecteurs tangents

Proposition (Intégration le long d'un arc) : Si $f : U \subset E \rightarrow F$ et $\gamma : [0, 1] \rightarrow E$ sont de classe C^1 avec $\gamma(0) = a$ et $\gamma(1) = b$, alors :

$$f(b) - f(a) = \int_0^1 df_{\gamma(t)}(\gamma'(t))dt$$

Si U est un ouvert connexe par arcs et $f : U \rightarrow F$, f est constante sur U si et seulement si pour tout $a \in U$, $df_a = 0_{\mathcal{L}(E, F)}$.

Définition : Vecteur tangent et Hyperplan tangent

- **Vecteur tangent :** Si $X \subset E$ et $x \in X$, un vecteur $v \in E$ est tangent à X en x s'il existe un arc γ défini sur $] -\varepsilon, \varepsilon[$ à valeurs dans X , dérivable en 0, tel que $\gamma(0) = x$ et $\gamma'(0) = v$. On note $T_x X$ l'ensemble des vecteurs tangents.
- **Hyperplan tangent :** Soit g une fonction numérique de classe C^1 sur U , $X = \{x \in U \mid g(x) = 0\}$ et $x \in X$. Si $dg_x \neq 0$, alors $T_x X = \ker(dg_x) = (\nabla g(x))^\perp$.

Lien / Remarque : Lien avec les lignes de niveaux

La notion d'hyperplan tangent est directement liée à celle de ligne de niveau. En effet, si $X = \mathcal{C}_c = \{x \in U \mid f(x) = c\}$ est une ligne de niveau de f , alors X est aussi l'ensemble des zéros de la fonction $g = f - c$.

L'hyperplan tangent à $\mathcal{C}_c$ en un point a (lorsque $\nabla f(a) \neq \vec{0}$) est donc :

$$T_a \mathcal{C}_c = (\nabla f(a))^\perp = \{h \in \mathbb{R}^n \mid \langle \nabla f(a), h \rangle = 0\}$$

Ce qui confirme que le gradient est orthogonal à la ligne de niveau.

2 Applications de classe C^k

2.1 Dérivées partielles d'ordres supérieurs

On définit par récurrence les dérivées partielles d'ordres supérieurs :

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \cdots \partial x_{i_1}} \right)$$

Définition : Application de classe C^k

Une application est dite de **classe** C^k sur un ouvert U si toutes ses dérivées partielles d'ordre k existent et sont continues.

Théorème de Schwarz : Soit $f : U \subset \mathbb{R}^2 \rightarrow F$ une application de classe C^2 sur un ouvert U . Alors :

$$\forall a \in U, \quad \frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a)$$

2.2 Hessienne

La **hessienne** au point $a \in U$ de la fonction numérique $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 est la matrice symétrique :

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leq i, j \leq n}$$

Formule

Formule de Taylor-Young à l'ordre 2 : Pour tout $x \in U$:

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle H_f(x)h, h \rangle + o(\|h\|^2)$$

3 Optimisation

Toutes les fonctions considérées sont des fonctions numériques.

3.1 Condition d'ordre 1

Définition : Point critique

Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^1 sur l'ouvert U . On dit que $a \in U$ est un **point critique** de f si :

$$df_a = 0 \quad \text{c'est-à-dire} \quad \nabla f(a) = \vec{0}$$

Théorème (CN d'existence d'un extremum) : Si $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^1 sur l'ouvert U admet un extremum au point $a \in U$, alors a est un point critique ($\nabla f(a) = \vec{0}$).

Attention

La propriété est fausse ailleurs que sur un ouvert. De plus, tout point critique ne correspond pas nécessairement à un extremum (cas des points selles).

3.2 Condition d'ordre 2

Théorème (CS d'existence d'un extremum) : Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 sur l'ouvert U . Si a est un point critique de f et $H_f(a) \in S_n^{++}(\mathbb{R})$ (matrice définie positive), alors f atteint un **minimum local strict** en a .

Lien / Remarque

Pour $n = 2$, le déterminant et la trace de la hessienne permettent de trouver facilement le signe des valeurs propres.

3.3 Optimisation sous contrainte

Théorème (Multiplicateur de Lagrange) : Soient f et g deux fonctions numériques de classe C^1 sur l'ouvert U de E et $X = \{x \in U \mid g(x) = 0\}$. Si la restriction de f à X admet un extremum local en $a \in X$ et $dg_a \neq 0$, alors df_a est colinéaire à dg_a .

Formule

L'existence de $\lambda \in \mathbb{R}^*$ tel que $\nabla f(a) = \lambda \nabla g(a)$ en un point critique a est une simple **condition nécessaire**. On mène ensuite une étude locale ou on donne un argument de compacité pour justifier la présence d'un extremum.

Lien / Remarque : Interprétation géométrique des multiplicateurs de La

Le théorème de Lagrange peut se reformuler ainsi : si f atteint un extremum sur la contrainte $X = \{g = 0\}$ en a , alors les gradients $\nabla f(a)$ et $\nabla g(a)$ sont colinéaires. Géométriquement, cela signifie que la ligne de niveau de f passant par a est **tangente** à la contrainte X en ce point. En effet, $\nabla f(a)$ est orthogonal à la ligne de niveau de f , et $\nabla g(a)$ est orthogonal à X . Si ces deux gradients sont colinéaires, alors les deux orthogonaux sont confondus, donc les deux ensembles sont tangents.

4 Intégrales multiples et changement de variables

Idée clé

La notion de différentielle et de matrice jacobienne, développée dans les sections précédentes, trouve l'une de ses applications les plus puissantes dans le calcul d'intégrales multiples. En effet, lorsqu'on souhaite intégrer une fonction sur un domaine de $\mathbb{R}^n$ dont la géométrie se prête mal aux coordonnées cartésiennes (disques, sphères, cônes...), on effectue un **changement de variables**. Le calcul différentiel intervient alors de manière cruciale : c'est le **déterminant de la matrice jacobienne** du changement de coordonnées qui donne le facteur correctif à appliquer à l'élément de volume.

4.1 Le théorème de changement de variables

Théorème : Changement de variables dans une intégrale multiple

Soit $\Phi : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ un $\mathcal{C}^1$ -difféomorphisme entre deux ouverts de $\mathbb{R}^n$ (c'est-à-dire une bijection de classe $\mathcal{C}^1$ dont la réciproque est aussi de classe $\mathcal{C}^1$).

Pour toute fonction $f : V \rightarrow \mathbb{R}$ continue et intégrable, on a :

$$\int_V f(x_1, \dots, x_n) dx_1 \cdots dx_n = \int_U f(\Phi(u_1, \dots, u_n)) |\det J_\Phi(u_1, \dots, u_n)| du_1 \cdots du_n$$

Lien / Remarque

Le facteur $|\det J_\Phi|$ mesure la **dilatation locale des volumes** induite par le changement de coordonnées. Géométriquement, la différentielle $d\Phi_u$ approxime localement Φ par une application linéaire, et le déterminant d'une application linéaire est exactement le facteur par lequel elle multiplie les volumes.

On retient la notation symbolique (à comprendre comme une conséquence du théorème) :

$$dx_1 \cdots dx_n = |\det J_\Phi(u_1, \dots, u_n)| du_1 \cdots du_n$$

4.2 Coordonnées cylindriques

On définit le changement de variables $\Phi : (r, \theta, z) \mapsto (x, y, z)$ par :

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \\ z = z \end{cases} \quad \text{avec } r > 0, \theta \in]0, 2\pi[, z \in \mathbb{R}.$$

Calcul / Exemple : Calcul de la jacobienne en coordonnées cylindriques

La matrice jacobienne de Φ est :

$$J_{\Phi}(r, \theta, z) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

En développant par rapport à la troisième ligne (ou troisième colonne) :

$$\det J_{\Phi}(r, \theta, z) = 1 \cdot (r \cos^2 \theta + r \sin^2 \theta) = r$$

Formule : Élément de volume en coordonnées cylindriques

Puisque $r > 0$, on a $|\det J_{\Phi}| = r$, d'où :

$$dx dy dz = r dr d\theta dz$$

4.3 Coordonnées sphériques

On définit le changement de variables $\Phi : (\rho, \theta, \varphi) \mapsto (x, y, z)$ par :

$$\begin{cases} x = \rho \sin \varphi \cos \theta \\ y = \rho \sin \varphi \sin \theta \\ z = \rho \cos \varphi \end{cases} \quad \text{avec } \rho > 0, \theta \in]0, 2\pi[, \varphi \in]0, \pi[.$$

Calcul / Exemple : Calcul de la jacobienne en coordonnées sphériques

La matrice jacobienne de Φ est :

$$J_{\Phi}(\rho, \theta, \varphi) = \begin{pmatrix} \sin \varphi \cos \theta & -\rho \sin \varphi \sin \theta & \rho \cos \varphi \cos \theta \\ \sin \varphi \sin \theta & \rho \sin \varphi \cos \theta & \rho \cos \varphi \sin \theta \\ \cos \varphi & 0 & -\rho \sin \varphi \end{pmatrix}$$

En développant le déterminant (par exemple par rapport à la troisième colonne puis en simplifiant à l'aide de $\cos^2 + \sin^2 = 1$), on obtient :

$$\det J_{\Phi}(\rho, \theta, \varphi) = -\rho^2 \sin \varphi$$

Formule : Élément de volume en coordonnées sphériques

Puisque $\varphi \in]0, \pi[$, on a $\sin \varphi > 0$ et donc $|\det J_{\Phi}| = \rho^2 \sin \varphi$, d'où :

$$dx dy dz = \rho^2 \sin \varphi d\rho d\theta d\varphi$$

Attention

Il ne faut **jamais** retenir ces formules par cœur sans comprendre d'où elles viennent. Elles sont la conséquence directe du théorème de changement de variables et du calcul de la matrice jacobienne, notions centrales du calcul différentiel.

De plus, attention aux conventions : certains auteurs (notamment en physique) échangent les rôles de θ et φ . La convention utilisée ici (angle φ par rapport à l'axe Oz , angle θ dans le plan xOy) est la plus courante en mathématiques.

Exercice

Calculer le volume de la boule unité $B = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\}$ en coordonnées sphériques.

Correction

En coordonnées sphériques, la boule unité est décrite par $\rho \in [0, 1]$, $\theta \in [0, 2\pi]$, $\varphi \in [0, \pi]$.

Le volume vaut donc :

$$\begin{aligned}\mathcal{V}(B) &= \int_0^{2\pi} \int_0^\pi \int_0^1 \rho^2 \sin \varphi \, d\rho \, d\varphi \, d\theta \\ &= \left(\int_0^{2\pi} d\theta \right) \times \left(\int_0^\pi \sin \varphi \, d\varphi \right) \times \left(\int_0^1 \rho^2 \, d\rho \right) \\ &= 2\pi \times [-\cos \varphi]_0^\pi \times \left[\frac{\rho^3}{3} \right]_0^1 \\ &= 2\pi \times 2 \times \frac{1}{3} = \frac{4\pi}{3}\end{aligned}$$

On retrouve bien la formule classique du volume de la boule unité.