

Calcul différentiel

I. La différentielle de Fréchet dans les espaces normés

Dans tout le chapitre, E et F désignent deux espaces vectoriels normés réels et U un ouvert de E . On note $\mathcal{L}_c(E, F)$ l'espace des applications linéaires continues de E dans F , muni de la norme subordonnée. En dimension finie, toute application linéaire est continue ; en revanche, en dimension infinie, la continuité fait partie de la définition.

L'idée est simple. Une fonction dérivable d'une variable réelle s'approche, près d'un point, par une fonction affine. On demande ici la même chose : l'accroissement $f(a+h) - f(a)$ doit ressembler à une application linéaire continue de E à F , à un reste négligeable devant $|h|$ près.

Définition :

Soit $f : U \rightarrow F$ et $a \in U$. On dit que f est **différentiable en** a s'il existe $L \in \mathcal{L}_c(E, F)$ telle que

$$f(a+h) = f(a) + L(h) + o(|h|) \text{ lorsque } h \rightarrow 0.$$

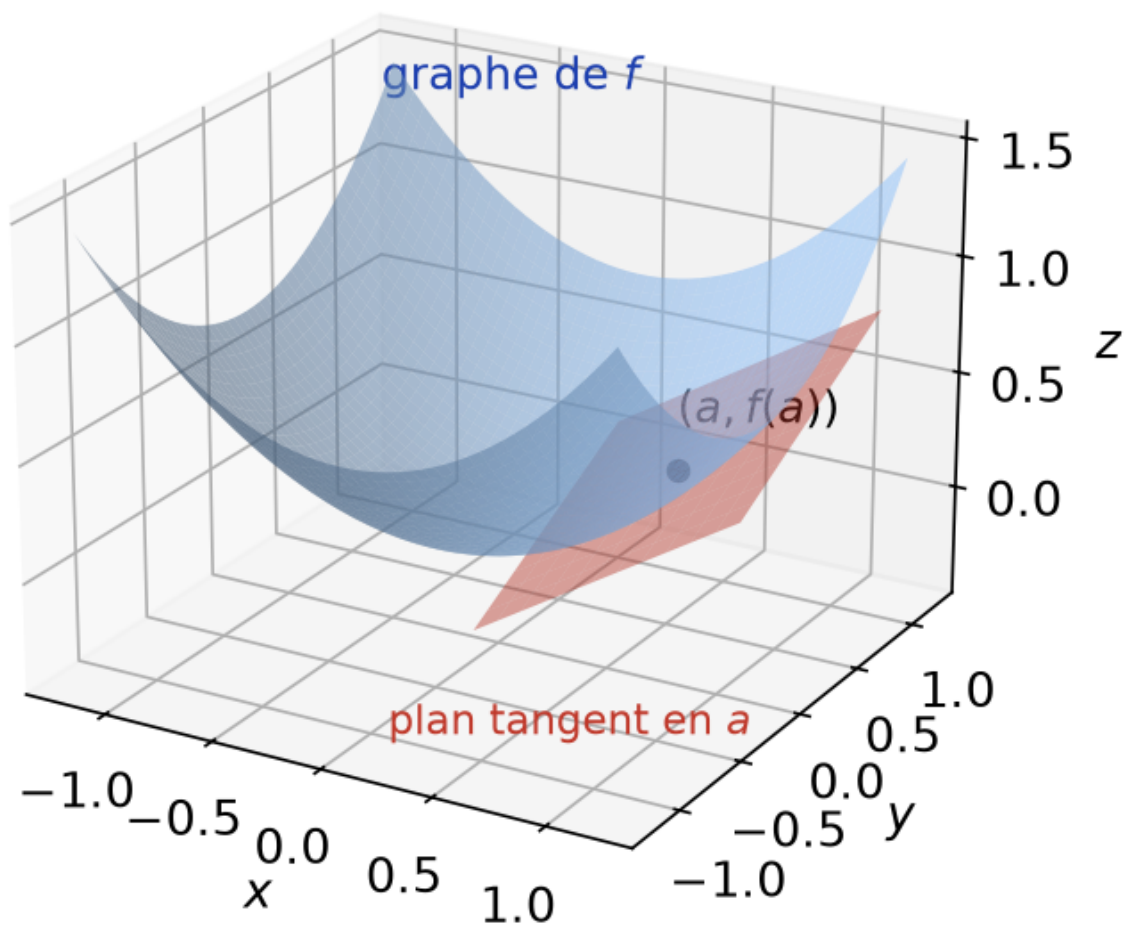
Autrement dit, $|f(a+h) - f(a) - L(h)| / |h| \rightarrow 0$. L'application L est alors unique : c'est la **différentielle** de f en a , notée $df(a)$.

L'unicité se démontre facilement. Si L_1 et L_2 conviennent, alors $(L_1 - L_2)(h) = o(|h|)$. Pour v fixé et $t \rightarrow 0^+$, on obtient $t(L_1 - L_2)(v) = o(t)$, donc $(L_1 - L_2)(v) = 0$. De plus, une application différentiable en a y est continue, puisque $df(a)$ est continue.

Propriété :

La différentiabilité et la différentielle ne changent pas si l'on remplace les normes de E et de F par des normes équivalentes. En particulier, en dimension finie, elles ne dépendent pas des normes choisies.

Géométriquement, pour $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, la différentiabilité en a signifie que le graphe de f admet un plan tangent en $(a, f(a))$. Ce plan a pour équation $z = f(a) + df(a)(x-a)$, comme le montre la figure ci-dessous pour $f(x,y) = (x^2+y^2)/2$.



Exemple :

Voici trois différentielles à connaître.

- Si $u \in \mathcal{L}_c(E, F)$, alors $u(a+h) = u(a) + u(h)$: ainsi $du(a) = u$ en tout point.
- Si $B : E_1 \times E_2 \rightarrow F$ est bilinéaire continue, alors
 $B(x+h, y+k) = B(x, y) + B(x, k) + B(h, y) + B(h, k)$. Or
 $|B(h, k)| \leq |B| |h| |k| = o(|(h, k)|)$. Par conséquent
 $dB(x, y)(h, k) = B(x, k) + B(h, y)$.
- Sur $\mathcal{M}_n(\mathbb{R})$ muni d'une norme d'algèbre, $f(M) = M^2$ vérifie
 $(M+H)^2 = M^2 + MH + HM + H^2$. Donc $df(M)(H) = MH + HM$,
 qui diffère de $2MH$ dès que M et H ne commutent pas.

La définition s'applique aussi en dimension infinie. Par exemple, sur $E = C([0, 1], \mathbb{R})$ muni de $\|\cdot\|_\infty$, la fonctionnelle $\varphi(u) = \int_0^1 u(t)^2 dt$ vérifie $\varphi(u+h) = \varphi(u) + 2 \int_0^1 u h + \int_0^1 h^2$. Le dernier terme est majoré par $\|h\|_\infty^2$. Ainsi $d\varphi(u)(h) = 2 \int_0^1 u(t)h(t) dt$, forme linéaire continue de norme au plus $2\|u\|_\infty$.

II. Dérivées directionnelles, dérivées partielles et classe C^1

1. Dérivée selon un vecteur

Définition :

Soit $v \in E$. La **dérivée de f en a selon v** est, si elle existe, la limite

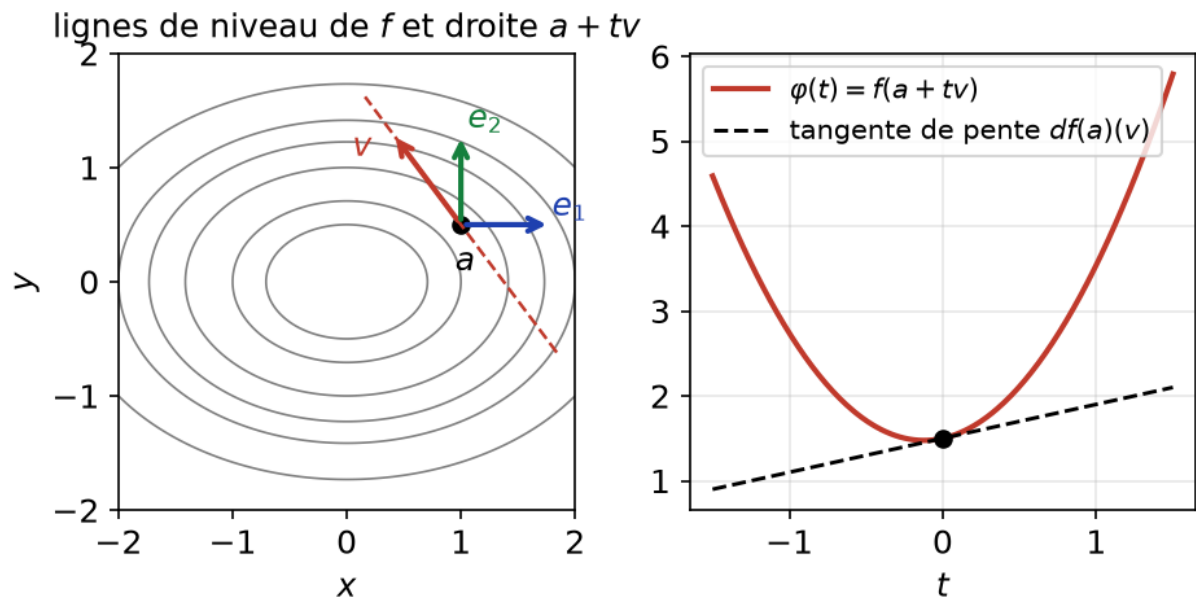
$$D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}.$$

C'est la dérivée en 0 de la fonction d'une variable $\varphi_v(t) = f(a+tv)$.

Proposition :

Si f est différentiable en a , elle admet une dérivée selon tout vecteur v , et $D_v f(a) = df(a)(v)$. En particulier, $v \rightarrow D_v f(a)$ est linéaire.

En effet, $f(a+tv) = f(a) + t df(a)(v) + o(t)$. La figure suivante illustre ce lien : la pente de la tangente à φ_v en 0 vaut $df(a)(v)$.



Attention :

La réciproque est fausse. Pour $f(x,y) = \frac{x^2y}{x^4+y^2}$ et $f(0,0) = 0$, toutes les dérivées

directionnelles existent en 0 . Cependant $f(x, x^2) = \frac{1}{2}$ pour $x \neq 0$: la fonction n'est même pas continue en 0 . L'existence de dérivées dans toutes les directions ne suffit donc jamais.

2. Dérivées partielles et matrice jacobienne

Supposons maintenant $E = \mathbb{R}^n$, muni de sa base canonique $(e_1, \dots, e_n)$. La dérivée selon e_j s'appelle la j -ième dérivée partielle, notée $\partial_j f(a)$ ou $\frac{\partial f}{\partial x_j}(a)$. Si f est différentiable en a , la linéarité donne

$$df(a)(h) = \sum_{j=1}^n h_j \partial_j f(a).$$

Définition :

Soit $f = (f_1, \dots, f_p) : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ différentiable en a . La **matrice jacobienne** de f en a est la matrice de $df(a)$ dans les bases canoniques :

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{\substack{1 \leq i \leq p, \\ 1 \leq j \leq n}} \in \mathcal{M}_{p,n}(\mathbb{R}).$$

Lorsque $p = 1$, le vecteur $\nabla f(a)$ défini par $df(a)(h) = \langle \nabla f(a), h \rangle$ est le **gradient**.

3. Applications de classe C^1

Définition :

L'application $f : U \rightarrow F$ est **de classe C^1** si elle est différentiable en tout point de U et si $df : U \rightarrow \mathcal{L}_c(E, F)$ est continue, l'espace d'arrivée étant muni de la norme subordonnée.

Théorème :

Soit $f : U \subset \mathbb{R}^n \rightarrow F$. Si les dérivées partielles $\partial_1 f, \dots, \partial_n f$ existent en tout point de U et sont continues sur U , alors f est de classe C^1 sur U . La réciproque est vraie.

Démonstration :

Traitons $n = 2$, le cas général étant identique. Soit $a \in U$ et $h = (h_1, h_2)$ petit. On écrit

$$f(a+h) - f(a) = [f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2)] + [f(a_1, a_2+h_2) - f(a)].$$

Le second crochet vaut $h_2 \partial_2 f(a) + o(h_2)$ par définition de $\partial_2 f(a)$. Pour le premier, on applique l'inégalité des accroissements finis de la partie IV à $t \rightarrow f(a_1+t, a_2+h_2) - t \partial_1 f(a)$ sur le segment d'extrémités 0 et h_1 . Sa dérivée est $\partial_1 f(a_1+t, a_2+h_2) - \partial_1 f(a)$. Par continuité de $\partial_1 f$, elle est de norme $\varepsilon(h) \rightarrow 0$. Ainsi le premier crochet vaut $h_1 \partial_1 f(a) + o(|h|)$. Finalement, f est différentiable en a . Enfin, df est continue car ses « coordonnées » $\partial_j f$ le sont.

Méthode :

Pour montrer qu'une application est de classe C^1 :

- en dimension finie, calculez les dérivées partielles et prouvez leur continuité, en isolant les points délicats (souvent l'origine) ;
- en dimension infinie, calculez d'abord $df(a)$ par un développement de $f(a+h)$, puis majorez $|df(a) - df(b)|$ en norme subordonnée ;
- sinon, utilisez les théorèmes d'opérations : sommes, produits, composées de fonctions C^1 sont C^1 .

Remarque :

Une fonction peut être différentiable sans être de classe C^1 . Par exemple,

$g(x,y) = (x^2+y^2)\sin(1/\sqrt{x^2+y^2})$, prolongée par $g(0,0) = 0$, est différentiable en 0 de différentielle nulle. Pourtant, ses dérivées partielles n'ont pas de limite en 0 .

III. Différentielle d'une composée et opérations

Théorème :

Soient E, F, G trois espaces normés, $f: U \subset E \rightarrow F$ différentiable en a et $g: V \subset F \rightarrow G$ différentiable en $b = f(a)$, avec $f(U) \subset V$. Alors $g \circ f$ est différentiable en a et

$$d(g \circ f)(a) = dg(f(a)) \circ df(a).$$

En dimension finie, cela s'écrit $J_{g \circ f}(a) = J_g(f(a)) J_f(a)$.

Démonstration :

Posons $k = f(a+h) - f(a) = df(a)(h) + |h|\varepsilon_1(h)$, avec $\varepsilon_1(h) \rightarrow 0$. En particulier $|k| \leq (|df(a)| + |\varepsilon_1(h)|)|h|$, donc $k \rightarrow 0$. De même, $g(b+k) = g(b) + dg(b)(k) + |k|\varepsilon_2(k)$. En remplaçant,

$$g(f(a+h)) - g(f(a)) = dg(b)(df(a)(h)) + |h| dg(b)(\varepsilon_1(h)) + |k| \varepsilon_2(k).$$

Le deuxième terme est un $o(|h|)$ par continuité de $dg(b)$. Le troisième aussi, puisque $|k| = O(|h|)$ et $\varepsilon_2(k) \rightarrow 0$. D'où le résultat.

On en déduit aussitôt les règles d'opérations. Une combinaison linéaire de fonctions différentiables est différentiable. De plus, si B est bilinéaire continue, $x \rightarrow B(f(x), g(x))$ a pour différentielle $h \rightarrow B(df(x)h, g(x)) + B(f(x), dg(x)h)$. C'est la composée de $x \rightarrow (f(x), g(x))$ et de B .

Exemple :

Pour $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ différentiable, posons $F(r, \theta) = f(r \cos \theta, r \sin \theta)$. La règle de la chaîne donne

$$\frac{\partial F}{\partial r} = \cos \theta \partial_1 f + \sin \theta \partial_2 f, \quad \frac{\partial F}{\partial \theta} = -r \sin \theta \partial_1 f + r \cos \theta \partial_2 f,$$

les dérivées de f étant prises en $(r \cos \theta, r \sin \theta)$.

1. Différentiel sur un espace de matrices

L'espace $\mathcal{M}_n(\mathbb{R})$ est de dimension finie : toutes les normes y sont équivalentes. On choisit donc une norme d'algèbre, qui vérifie $|AB| \leq |A| |B|$. Ensuite, on développe $f(M+H)$ et on isole la partie linéaire en H .

Proposition :

L'ensemble $GL_n(\mathbb{R})$ est un ouvert de $\mathcal{M}_n(\mathbb{R})$, et l'inversion $\iota : M \rightarrow M^{-1}$ y est différentiable, avec

$$d\iota(M)(H) = -M^{-1} H M^{-1}.$$

De plus, $\det$ est différentiable et $d\det(M)(H) = \text{tr}({}^t \text{com}(M) H)$; en particulier, $d\det(I_n)(H) = \text{tr}(H)$.

Démonstration :

D'abord, $GL_n(\mathbb{R}) = \det^{-1}(\mathbb{R}^*)$ est ouvert, car $\det$ est polynomiale donc continue. Ensuite, si $|H| < 1$, la série $\sum_k (-H)^k$ converge et $I+H$ est inversible d'inverse $I - H + \sum_{k \geq 2} (-H)^k$. Le reste est majoré par $|H|^2/(1-|H|)$: ainsi $(I+H)^{-1} = I - H + o(|H|)$. Enfin, $(M+H)^{-1} = (I + M^{-1}H)^{-1} M^{-1}$, ce qui donne $M^{-1} - M^{-1}HM^{-1} + o(|H|)$. Pour le déterminant, voir l'exercice consacré à cette formule.

Méthode :

Pour calculer une différentielle :

- écrivez $f(a+h)$ et développez ;
- repérez le terme linéaire en h et vérifiez qu'il est continu ;
- majorez le reste par $C|h|^2$ ou montrez qu'il est un $o(|h|)$;
- si f est une composée, préférez la règle de la chaîne.

IV. Inégalité des accroissements finis

Pour une fonction à valeurs vectorielles, l'égalité des accroissements finis tombe en défaut. Par exemple, $t \rightarrow e^{it}$ vaut 1 en 0 et en 2π , alors que sa dérivée ne s'annule jamais. Il reste cependant une inégalité, qui suffit presque toujours.

Lemme :

Soit $g : [0,1] \rightarrow F$ continue, dérivable sur $[0,1[$, avec $|g'(t)| \leq k$ pour tout t . Alors $|g(1)-g(0)| \leq k$.

Démonstration :

Fixons $\varepsilon > 0$. Soit A l'ensemble des $t \in [0,1]$ tels que $|g(s)-g(0)| \leq (k+\varepsilon)s$ pour tout $s \in [0,t]$. C'est un intervalle contenant 0, fermé par continuité de g . Posons $c = \max A$ et supposons $c < 1$. Par dérivabilité en c , pour $h > 0$ assez petit, $|g(c+h)-g(c)-hg'(c)| \leq \varepsilon h$. Par conséquent,

$$|g(c+h)-g(0)| \leq (k+\varepsilon)c + kh + \varepsilon h = (k+\varepsilon)(c+h).$$

Donc $c+h \in A$, ce qui contredit la maximalité de c . Ainsi $c = 1$, puis on fait tendre ε vers 0.

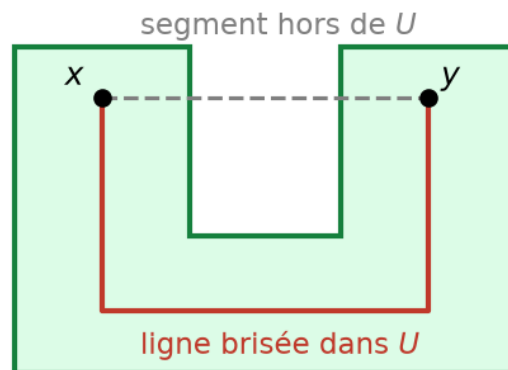
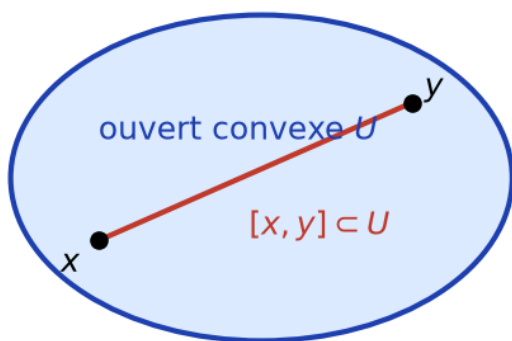
Théorème :

Inégalité des accroissements finis. Soit $f : U \rightarrow F$ différentiable et $x, y \in U$ tels que le segment $[x,y]$ soit inclus dans U . Si $|df(z)| \leq k$ pour tout $z \in [x,y]$, alors

$$|f(y)-f(x)| \leq k |y-x|.$$

En particulier, si U est convexe et $\sup_U |df| \leq k$, alors f est k -lipschitzienne sur U .

La preuve consiste à appliquer le lemme à $g(t) = f(x+t(y-x))$. En effet, par la règle de la chaîne, $g'(t) = df(x+t(y-x))(y-x)$, de norme au plus $k|y-x|$. L'hypothèse de segment est essentielle. Sur un ouvert non convexe, on relie deux points par une ligne brisée, comme sur la figure ci-dessous ; la constante obtenue dépend alors de la longueur du chemin.



Corollaire :

Si U est un ouvert connexe et si $df(x) = 0$ pour tout $x \in U$, alors f est constante sur U .

Démonstration :

Fixons $a \in U$ et posons $A = \{x \in U : f(x) = f(a)\}$. D'abord, A est fermé dans U par continuité de f . Ensuite, A est ouvert : si $x \in A$ et $B(x, r) \subset U$, la boule est convexe, donc f y est constante par l'inégalité avec $k = 0$. Enfin A est non vide. Par connexité, $A = U$.

Exemple :

L'application $F(x, y) = \left(\frac{1}{3}\sin(x+y), \frac{1}{3}\cos(x-y)\right)$ a une jacobienne de norme euclidienne subordonnée au plus $\frac{\sqrt{2}}{3}$. Elle est donc $\frac{\sqrt{2}}{3}$ -lipschitzienne sur $\mathbb{R}^2$, c'est-à-dire contractante.

V. Différentielles d'ordre supérieur et théorème de Schwarz

Si f est différentiable sur U , sa différentielle df est une application de U dans $\mathcal{L}_c(E, F)$. On peut donc chercher à la différentier à son tour.

Définition :

On dit que f est **deux fois différentiable en** a si df est différentiable en a . On note $d^2f(a) = d(df)(a) \in \mathcal{L}_c(E, \mathcal{L}_c(E, F))$, que l'on identifie à l'application bilinéaire continue $(h, k) \rightarrow d^2f(a)(h, k) = (d(df)(a)(k))(h)$. L'application f est de classe C^2 si d^2f existe et est continue sur U . On définit de même la classe C^k .

Concrètement, $d^2f(a)(h, k)$ est la dérivée selon k de la fonction $x \rightarrow df(x)(h)$. En dimension finie, si f est à valeurs réelles, $d^2f(a)(h, k) = \sum_{i,j} \partial_i \partial_j f(a) h_i k_j$. La matrice $H_f(a) = (\partial_i \partial_j f(a))$ est la **matrice hessienne**.

Exemple :

Pour $f(M) = M^2$, on a $df(M)(H) = MH + HM$. Cette expression est linéaire en M pour H fixé. Sa dérivée selon K vaut donc $KH + HK$. Ainsi $d^2f(M)(H, K) = HK + KH$, qui est bien symétrique en (H, K) .

Théorème :

Théorème de Schwarz. Si f est deux fois différentiable en a , alors $d^2f(a)$ est une application bilinéaire symétrique. En particulier, si $f : U \subset \mathbb{R}^n \rightarrow F$ est de classe C^2 , alors $\partial_i \partial_j f = \partial_j \partial_i f$ pour tous i, j , et la hessienne est symétrique.

Démonstration :

Donnons l'idée pour $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^2 . Posons

$\Delta(t) = f(a+te_1+te_2) - f(a+te_1) - f(a+te_2) + f(a)$. D'une part, avec $\psi(s) = f(a+se_1+te_2) - f(a+se_1)$, on a $\Delta(t) = \psi(t) - \psi(0)$. Le théorème des accroissements finis appliqué deux fois donne $\Delta(t) = t^2 \partial_2 \partial_1 f(c_t)$, avec $c_t \rightarrow a$. D'autre part, en échangeant les rôles des variables, $\Delta(t) = t^2 \partial_1 \partial_2 f(c'_t)$. Enfin, on divise par t^2 et on utilise la continuité des dérivées secondes.

Attention :

Sans hypothèse de régularité, les dérivées croisées peuvent différer. Pour $f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2}$ et $f(0,0) = 0$, on trouve $\partial_2 \partial_1 f(0,0) = -1$ alors que $\partial_1 \partial_2 f(0,0) = 1$. Cette fonction est C^1 mais pas C^2 .

VI. Formules de Taylor

Les formules de Taylor décrivent $f(a+h)$ jusqu'à l'ordre 2. Elles servent à étudier la position d'une fonction par rapport à son plan tangent et à approcher f près d'un point.

Théorème :

Formule de Taylor-Young. Si $f : U \rightarrow F$ est deux fois différentiable en a , alors

$$f(a+h) = f(a) + df(a)(h) + \frac{1}{2} d^2 f(a)(h,h) + o(|h|^2).$$

Pour $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, le terme d'ordre deux s'écrit $\frac{1}{2} {}^t h H_f(a) h$.

Théorème :

Taylor avec reste intégral. Si F est complet, si f est de classe C^2 et si $[a, a+h] \subset U$, alors

$$f(a+h) = f(a) + df(a)(h) + \int_0^1 (1-t) d^2 f(a+th)(h,h) dt.$$

Inégalité de Taylor-Lagrange. Si de plus $|d^2 f(z)| \leq M$ sur $[a, a+h]$, alors
 $|f(a+h) - f(a) - df(a)(h)| \leq \frac{M}{2} |h|^2$.

Démonstration :

On pose $g(t) = f(a+th)$, de classe C^2 sur $[0,1]$. La règle de la chaîne donne
 $g'(t) = df(a+th)(h)$ et $g''(t) = d^2 f(a+th)(h,h)$. Il suffit alors d'appliquer la formule de Taylor avec reste intégral en dimension un : $g(1) = g(0) + g'(0) + \int_0^1 (1-t) g''(t) dt$. Enfin,

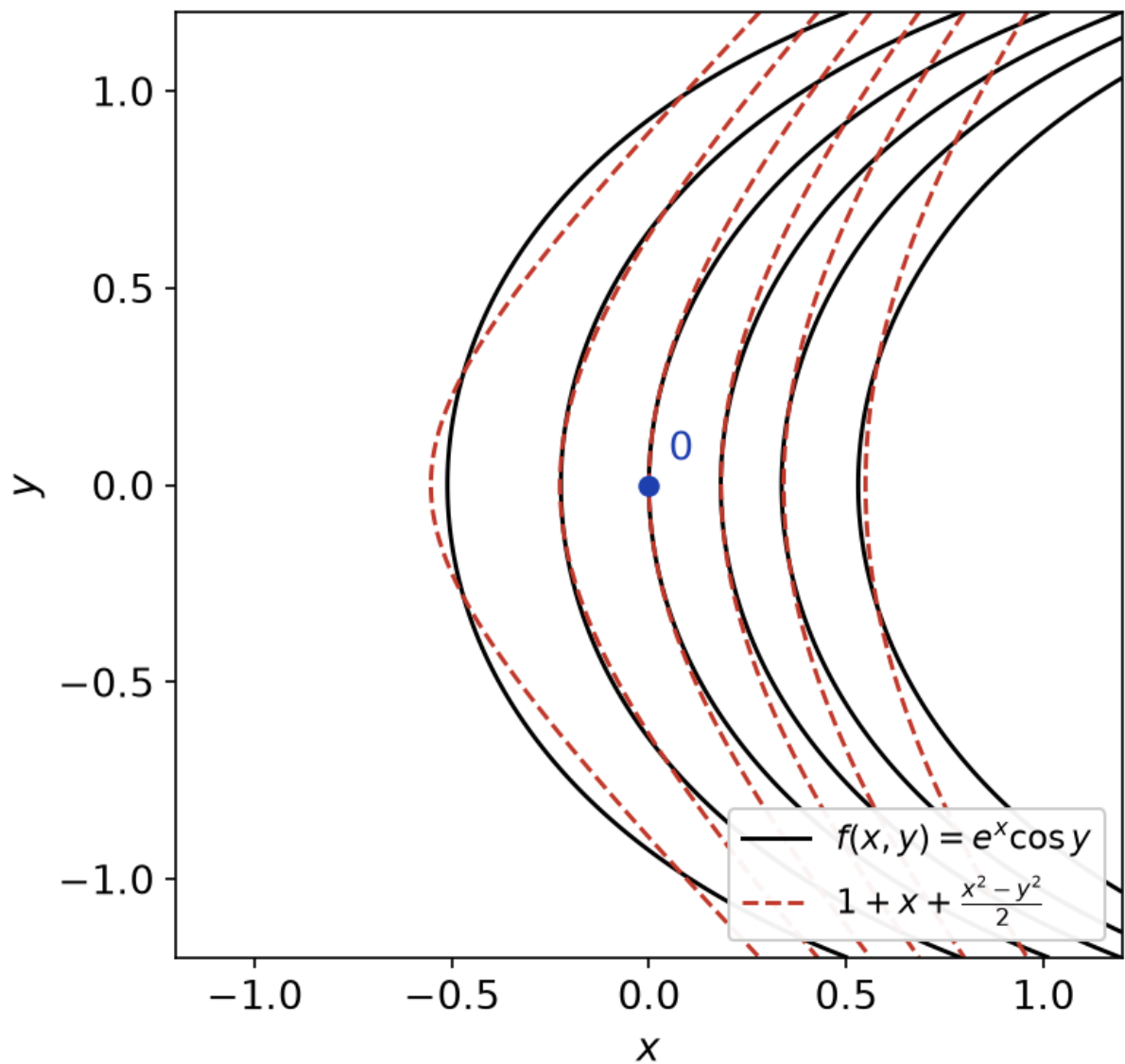
l'inégalité découle de $\int_0^1 (1-t) dt = \frac{1}{2}$.

Exemple :

Pour $f(x,y) = e^x \cos y$ en 0 , on a $\nabla f(0) = (1,0)$ et $H_f(0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
Donc

$$e^x \cos y = 1 + x + \frac{x^2 - y^2}{2} + o(x^2 + y^2).$$

La figure ci-dessous superpose les lignes de niveau de f et celles de ce polynôme : elles se confondent près de l'origine.



Méthode :

Pour écrire une formule de Taylor à l'ordre 2 en a :

- calculez le gradient et la hessienne en a , ou bien $df(a)$ et $d^2f(a)(h, h)$ directement ;
- vérifiez la symétrie de la hessienne, qui contrôle vos calculs grâce au théorème de Schwarz ;
- à défaut, développez $f(a+h)$ à l'aide des développements limités usuels et gardez les termes de degré au plus 2.

VII. Convexité et différentiabilité

Soit U un ouvert convexe de E et $f:U \rightarrow \mathbb{R}$. Rappelons que f est convexe si $f(tx+(1-t)y) \leq tf(x) + (1-t)f(y)$ pour tous $x, y \in U$ et $t \in [0,1]$. Lorsque f est différentiable, cette propriété se lit sur la différentielle.

Théorème :

Soit $f:U \rightarrow \mathbb{R}$ différentiable sur l'ouvert convexe U . Les assertions suivantes sont équivalentes :

1. f est convexe ;
2. pour tous $x, y \in U$, $f(y) \geq f(x) + df(x)(y-x)$: le graphe est au-dessus de ses plans tangents ;
3. pour tous $x, y \in U$, $(df(y) - df(x))(y-x) \geq 0$.

Si f est de classe C^2 , elles équivalent aussi à : $d^2f(x)(h,h) \geq 0$ pour tous $x \in U$ et $h \in E$.

Démonstration :

Montrons (1) $\Rightarrow$ (2). Pour $t \in]0,1]$, la convexité donne

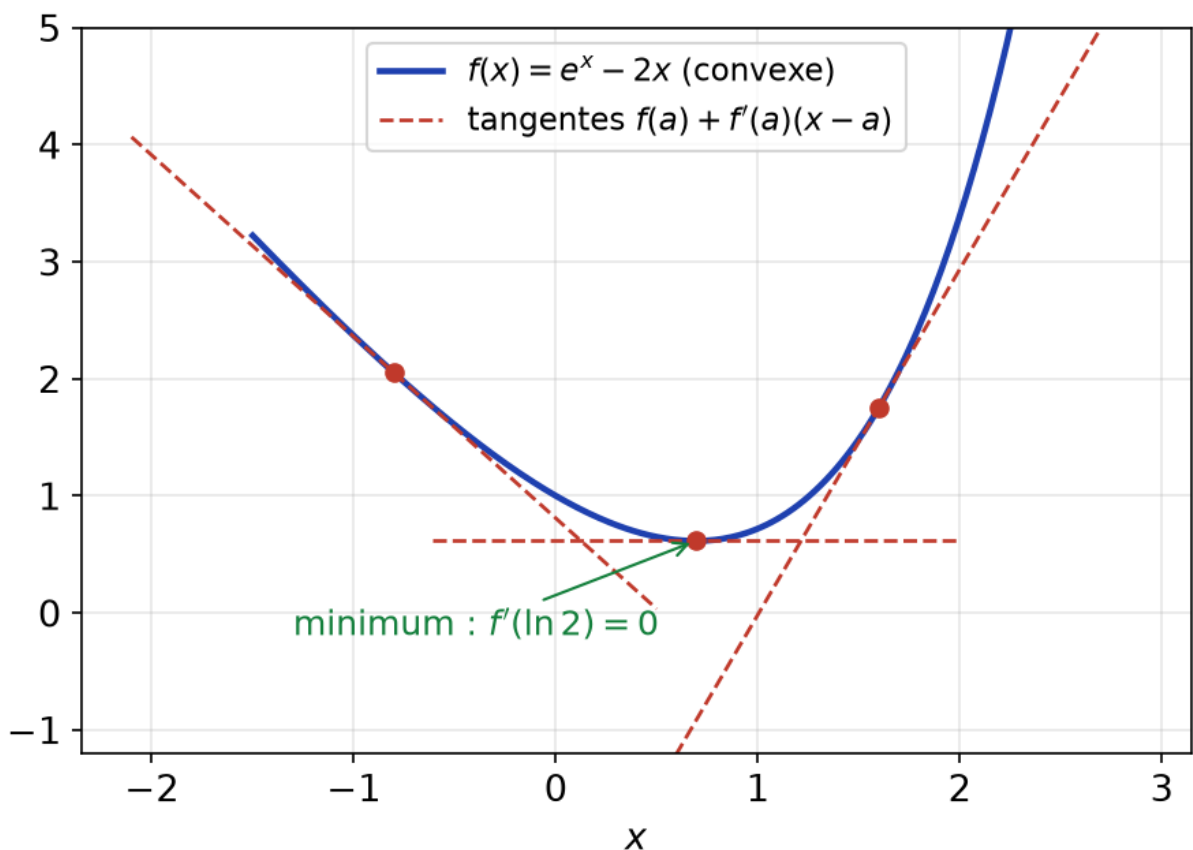
$f(x+t(y-x)) \leq f(x) + t(f(y)-f(x))$. Donc $\frac{f(x+t(y-x))-f(x)}{t} \leq f(y)-f(x)$. En faisant tendre t vers 0^+ , on obtient $df(x)(y-x) \leq f(y)-f(x)$.

Réciproquement, supposons (2) et posons $z = tx+(1-t)y$. On a

$f(x) \geq f(z) + df(z)(x-z)$ et $f(y) \geq f(z) + df(z)(y-z)$. Multiplions par t et $1-t$, puis ajoutons. Comme $t(x-z)+(1-t)(y-z) = 0$, il reste $tf(x)+(1-t)f(y) \geq f(z)$.

Ensuite, $(2) \Rightarrow (3)$ s'obtient en ajoutant les inégalités écrites en x et en y . Pour $(3) \Rightarrow (1)$, on constate que $g(t) = f(x+t(y-x))$ a une dérivée croissante, donc g est convexe. Enfin, dans le cas C^2 , $g''(t) = d^2 f(x+th)(h,h)$, et une fonction d'une variable est convexe si et seulement si $g'' \geq 0$.

La figure suivante illustre la caractérisation (2) en dimension un : chaque tangente reste sous la courbe.



Corollaire :

Si f est convexe et différentiable sur l'ouvert convexe U , alors a est un point de minimum global de f si et seulement si $df(a) = 0$.

En effet, si $df(a) = 0$, la caractérisation (2) donne $f(y) \geq f(a)$ pour tout y . Par exemple, si A

est symétrique définie positive, $f(x) = \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle$ a pour hessienne A . Elle est donc convexe, et son unique point critique $A^{-1}b$ réalise son minimum.

Remarque :

L'étude des extrema locaux d'une fonction quelconque et des extrema sous contrainte utilise les mêmes outils. Elle est développée dans le chapitre sur l'inversion locale et les fonctions implicites.

Ce qu'il faut retenir

- La différentielle $df(a)$ est l'unique application linéaire continue telle que $f(a+h) = f(a) + df(a)(h) + o(|h|)$.
- En dimension finie, elle ne dépend pas des normes ; en dimension infinie, la continuité de $df(a)$ doit être vérifiée.
- Différentiable implique dérivable selon tout vecteur avec $D_v f(a) = df(a)(v)$; la réciproque est fausse.
- Dérivées partielles continues sur un ouvert de $\mathbb{R}^n$: l'application est de classe C^1 .
- Règle de la chaîne : $d(g \circ f)(a) = dg(f(a)) \circ df(a)$, soit un produit de matrices jacobiniennes.
- Sur les matrices : $d(M^2)(H) = MH + HM$, $d(M^{-1})(H) = -M^{-1}HM^{-1}$, $d\det(I_n)(H) = \text{tr } H$.
- Inégalité des accroissements finis : $|f(y) - f(x)| \leq \sup_{[x,y]} |df| |y - x|$; une différentielle nulle sur un connexe donne une fonction constante.
- Théorème de Schwarz : pour une fonction C^2 , la différentielle seconde est symétrique.
- Taylor-Young : $f(a+h) = f(a) + df(a)(h) + \frac{1}{2}d^2f(a)(h,h) + o(|h|^2)$.
- Une fonction différentiable est convexe si et seulement si son graphe est au-dessus de ses plans tangents ; ses points critiques sont alors des minima globaux.

Questions fréquentes sur calcul différentiel

Quelle différence entre dérivées partielles et différentielle ?

Les dérivées partielles décrivent les variations de f le long des axes seulement. La différentielle est une application linéaire continue qui approche $f(a+h) - f(a)$ dans toutes les directions à la fois, avec un reste en $o(\|h\|)$. Une fonction peut avoir toutes ses dérivées partielles sans être différentiable, ni même continue. En revanche, des dérivées partielles continues sur un ouvert de $\mathbb{R}^n$ entraînent la classe C^1 .

Pourquoi la norme compte-t-elle en dimension infinie ?

En dimension finie, toutes les normes sont équivalentes : la différentiabilité ne dépend donc pas de la norme. En dimension infinie, ce n'est plus vrai. Par exemple, $u \mapsto \int u^2$ est différentiable sur $C([0,1])$ pour la norme uniforme, mais nulle part pour la norme $\int |u|$.

Comment différentier une fonction de matrices ?

On développe $f(M+H)$, on isole le terme linéaire en H et on majore le reste par une norme d'algèbre. On obtient ainsi $d(M^2)(H) = MH + HM$, $d(M^{-1})(H) = -M^{-1}HM^{-1}$ et $d(\det)(I)(H) = \text{tr } H$. Attention à ne pas faire commuter M et H .

Quand utiliser l'inégalité des accroissements finis ?

Elle sert à majorer $\|f(y) - f(x)\|$ par $\sup \|df\| \cdot \|y - x\|$, pour montrer qu'une application est lipschitzienne ou contractante. Elle exige que le segment $[x, y]$ soit inclus dans le domaine. Sur un ouvert non convexe, on la combine avec des lignes brisées.

Pour aller plus loin en L3

- Les énoncés : [exercices de maths en L3 sur calcul différentiel](#)
- À maîtriser avant : [Espaces normés et applications linéaires continues](#)

- Chapitre précédent : [Réduction de Jordan et décomposition de Dunford](#)
- Chapitre suivant : [Inversion locale, fonctions implicites et extrema liés](#)
- Le même thème en maths spé : [calcul différentiel, cours de maths spé](#)
- Tester vos connaissances : [QCM de maths en L3 par chapitre](#)
- Le sommaire : [tous les chapitres de maths de L3](#) et la [licence de maths de L1 à L3](#)