Chap 0 - Rappels de calcul différentiel

Où l'on rappelle que les applications linéaires, c'est mieux

1 De la dérivée à la différentielle

Dérivée classique: les fonctions réelles

Définition 1

Soit $f : I \subset \mathbb{R} \to \mathbb{R}$ une fonction définie sur un intervalle, et soit $a \in I$.

On dit que f est dérivable *en a si le* taux d'accroissement *de f en a, $\dfrac{f(a+h)-f(a)}{h}$ a une limite quand $h \to 0$.*

On note cette limite $f'(a)$; on a donc

$$\left| \frac{f(a+h)-f(a)}{h} - f'(a) \right| \xrightarrow[h\to 0]{} 0$$

et si f est dérivable en tout point $a \in I$, ceci définit une fonction $f' : a \in I \mapsto f'(a) \in \mathbb{R}$, qu'on appelle la dérivée *de f.*

$\rightsquigarrow$ Comment peut-on généraliser à des fonctions définies sur des e.v.n. ?

Un cas "facile": fonctions vectorielles à une variable

Soit $(F, \|.\|_F)$ un e.v.n., $f : I \subset \mathbb{R} \to F$ une fonction à valeurs dans F, et $t_0 \in I$.

Comme on sait calculer des limites dans un espace vectoriel normé, commme F, on peut généraliser directement la définition précédente:

Définition 2

On dit que f est dérivable *en t_0 s'il existe un* **vecteur** *$f'(t_0) \in F$ tel que*

$$\left\| \frac{f(t_0+h)-f(t_0)}{h} - f'(t_0) \right\|_F \xrightarrow[h\to 0]{} 0$$

On dit alors que $f'(t_0)$ est la dérivée *de f en t_0.*

$\rightsquigarrow$ Que faire si $f : U \subset E \to F$ est définie sur un e.v.n. $(E, \|.\|_E)$?

Dérivée directionnelle

Soit donc $f : U \subset E \to F$ une fonction définie entre deux e.v.n. Cette fois, la définition précédente ne peut pas être généralisée: si $a \in U \subset E$ et si $h \in E$ est tel que $a + h \in U$, on ne peut pas définir

$$\cancel{\frac{f(a+h) - f(a)}{h}}$$

car on ne peut pas diviser par un vecteur comme h. C'est un problème.

Solution 1: On se ramène de force à une seule dimension en étudiant les variations de f dans *une seule direction* donnée par un vecteur $v \in E$: autrement dit, au lieu d'étudier f sur U, on regarde la fonction

$$f_v : t \in I_a \mapsto f(a + tv)$$

où I_a est un intervalle tel que pour tout $t \in I_a$, $a + tv \in U$.

Question bonus : Pourquoi est-ce que I_a est un intervalle ouvert qui contient 0 ?

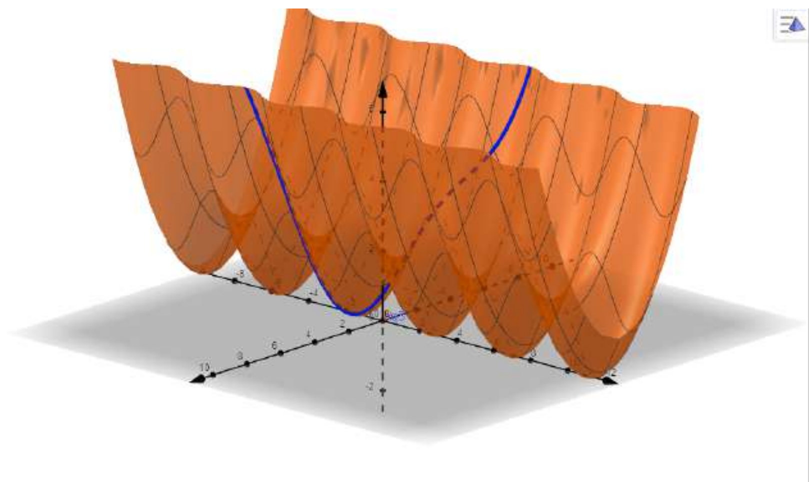


Figure 1: Source: `https://www.geogebra.org/m/t2q3pvtz`

Ici, on voit en orange le graphe de la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto 1 + \frac{x^2}{4} + \sin\left(\frac{3}{2}y\right)$, et en bleu la courbe qui représente la fonction f_v avec $v = (2, -3)$ et $a = (1, -1)$.

Définition 3

Pour $a \in U$ et $v \in E$, on dit que f est dérivable en a dans la direction de v si f_v est dérivable en 0, autrement dit, s'il existe un vecteur $u \in F$ tel que

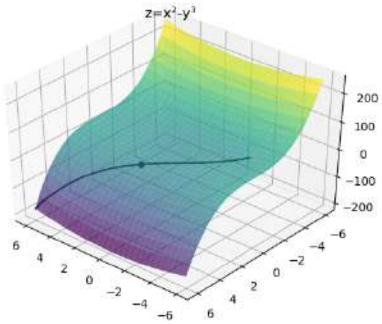
$$\left\| \frac{f(a + tv) - f(a)}{t} - u \right\|_F \xrightarrow[t \to 0]{} 0.$$

On note alors $u = \frac{\partial f}{\partial v}(a) \in F$ et on l'appelle la dérivée directionnelle de f en a dans la direction de v.

Une illustration: `https://mathinsight.org/applet/directional_derivative_mountain`

Exemple

Considérons $a = (1,2), v = (1,1)$ et : $(x,y) \in \mathbb{R}^2 \mapsto x^2 - y^3 \in \mathbb{R}$: on a alors



$$\begin{aligned}
\frac{f(a+tv)-f(a)}{t} &= \frac{1}{t}((1+t)^2 - (2+t)^3 - (-7)) \\
&= \frac{1}{t}(-10t - 5t^2 - t^3) \\
&= -10 - 5t - t^2 \\
&\xrightarrow[t\to 0]{} -10 = \frac{\partial f}{\partial v}(a)
\end{aligned}$$

Dans le cas particulier ou $E = \mathbb{R}^n$: Dérivées partielles

Si $E = \mathbb{R}^n$, il y a des directions particulièrement intéressantes: celles des vecteurs de la *base canonique* $(e_1, \ldots, e_n)$.

Définition 4

Si f admet une dérivée en a dans la direction de e_i, on l'appelle la i-ième dérivée partielle de f, et on la note $\frac{\partial f}{\partial x_i}(a)$. On a donc

$$\left\| \frac{f(a+te_i)-f(a)}{t} - \frac{\partial f}{\partial x_i}(a) \right\|_F \xrightarrow[t\to 0]{} 0 \text{ i.e. } \left\| \frac{f(a_1 \ldots, a_i + t, \ldots, a_n) - f(a)}{t} - \frac{\partial f}{\partial x_i}(a) \right\|_F \xrightarrow[t\to 0]{} 0.$$

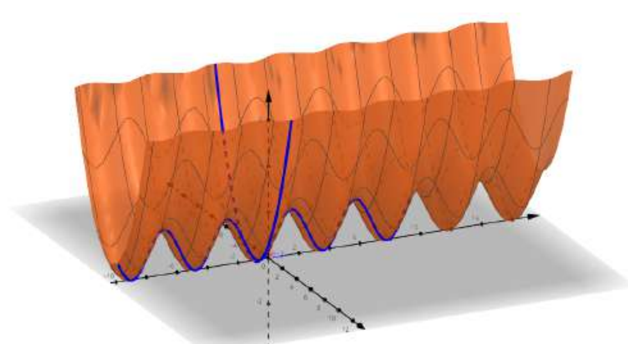


Figure 2: Source: `https://www.geogebra.org/m/t2q3pvtz`

Ici, on voit à nouveau en orange le graphe de la fonction $f : (x,y) \in \mathbb{R}^2 \mapsto 1 + \frac{x^2}{4} + \sin\left(\frac{3}{2}y\right)$, et en bleu les graphes des fonctions partielles en $a = (1,-1)$.

Remarque 5

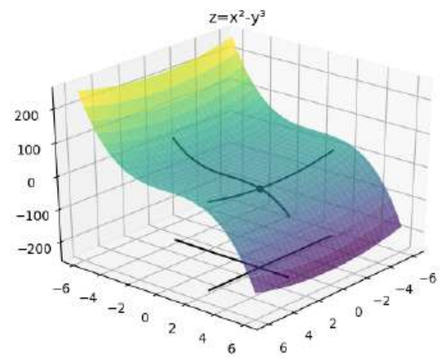
Calculer la i-ème dérivée partielle revient à considérer la fonction

$$\varphi_i : t \in I \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots a_n) \in F$$

obtenue en fixant toutes les variables à $x_j = a_j$ sauf la i-ème à la valeur qui nous intéresse : si elle existe, la i-ème dérivée dérivée partielle de f en a est la dérivée de φ_i en a_i.

Exemple: Toujours avec $f(x,y) = x^2 - y^3$, on a

$$\begin{aligned}
&\frac{1}{t}(f((x,y)+te_1) - f(x,y)) \\
&\quad = \frac{1}{t}((x+t)^2 - y^3 - (x^2 - y^3)) \\
&\quad = 2x + t \xrightarrow[t\to 0]{} 2x = \frac{\partial f}{\partial x}(x,y) \\
&\frac{1}{t}(f((x,y)+te_2) - f(x,y)) \\
&\quad = \frac{1}{t}(x^2 - (y+t)^3 - (x^2 - y^3)) \\
&\quad = -3y^2 - 3yt - t^2 \xrightarrow[t\to 0]{} -3y^2 = \frac{\partial f}{\partial y}(x,y)
\end{aligned}$$



$\rightsquigarrow$ Très bien, mais y a t-il moyen de "dériver" en tenant compte de toutes les directions en même temps ?

DL à l'ordre 1 et approximation par une fonction affine

Revenons à la dérivée classique pour les fonctions $\mathbb{R} \to \mathbb{R}$. Géométriquement, la dérivée en a détermine la tangente au graphe de f en a: c'est la droite qui "approche" le mieux f au voisinage de a. Plus formellement, si on note

$$R(h) = \underbrace{f(a+h)}_{f \text{ au voisinage de } a} - \underbrace{(f(a) + hf'(a))}_{\text{affine}} \tag{1}$$

alors f est dérivable en a si $R(h)$ tend vers 0 "assez vite" quand $h \to 0$, autrement dit, si

$$\frac{1}{h}R(h) \xrightarrow[h\to 0]{} 0, \text{ i.e. } R(h) = o(|h|).$$

C'est cette idée qu'on va généraliser:

Définition 6

Soit $f : U \subset E \to F$ une application entre deux e.v.n. On dit que f est différentiable en $a \in U$ s'il existe $L \in \mathcal{L}(E,F)$ telle que, pour tout $h \in E$ assez petit,

$$\boxed{f(a+h) = f(a) + L(h) + R(h) \text{ avec } \frac{1}{\|h\|_E}\|R(h)\|_F \xrightarrow[h\to 0]{} 0}$$

On appelle L la différentielle de f en a, et on la note $Df(a)$.

- Pour chaque $a \in U$ où f est différentiable, $Df(a)$ est une application linéaire continue sur E :

$$\boxed{Df(a) \in \mathcal{L}(E, F)}$$

- Pour $h \in E$, on note $Df(a)(h)$ son image par cette application: donc

$$\boxed{Df(a)(h) \in F}$$

  est un vecteur de F.

  $\rightsquigarrow$ Attention à ne pas les confondre !

Différentielles et dérivées directionnelles

Si f est différentiable en $a \in U$, elle admet des dérivées directionnelles en a dans toutes les directions: c'est en ce sens que la différentielle permet de dériver "dans toutes les directions en même temps".

Proposition 1

Si f est différentiable en a, alors pour tout $v \in E$, f admet une dérivée directionnelle dans la direction de v et

$$\boxed{\frac{\partial f}{\partial v}(a) = Df(a)(v)}$$

Preuve: Puisque f est différentiable en a, on a, pour tout $h \in E$,

$$f(a+h) = f(a) + Df(a)(h) + R(h) \text{ avec } \frac{1}{\|h\|_E}\|R(h)\|_F \xrightarrow[h\to 0]{} 0$$

Soit $v \in E$, on calcule

$$\frac{1}{t}(f(a+tv) - f(a)) = \frac{1}{t}(Df(a)(tv) + R(tv)) = Df(a)(v) + \underbrace{\frac{1}{t}R(tv)}_{\xrightarrow[t\to 0]{} 0}$$

$\square$

Remarque 7

En particulier, si $E = \mathbb{R}^n$ et si f est différentiable en a, alors f admet une dérivée directionnelle dans la direction de e_i pour $i = 1, \ldots, n$. Autrement dit, si f est différentiable en $a \in \mathbb{R}^n$, alors f admet des dérivées partielles en a.

De plus, par linéarité, pour tout $v = (v_1 \ldots, v_n) \in \mathbb{R}^n$, on a

$$Df(a)(v) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(a). \tag{2}$$

Quelques cas particuliers

- Lien dérivée-différentielle pour les fonctions d'une seule variable: Si $E = \mathbb{R}$, et si f est différentiable en a, alors, par linéarité de $Df(a)$, on a, pour tout $h \in \mathbb{R}$, $Df(a)(h) = hDf(a)(1) \in F$, donc

$$\frac{1}{|h|}\|f(a+h) - f(a) - Df(a)(h)\|_F = \left\| \frac{f(a+h) - f(a)}{h} - Df(a)(1) \right\|_F$$

  On obtient donc que $Df(a)(1) = f'(a)$.

  Réciproquement, si f est dérivable en a, alors pour tout $h \in \mathbb{R}$ tel que $a + h \in U$, on a, par la formule de Taylor à l'ordre 1,

$$f(a+h) = f(a) + hf'(a) + o(h) = f(a) + L(h) + R(h)$$

  avec $L(h) = f'(a)h$ et $R(h) = o(h)$ donc $\frac{1}{|h|}R(h) \to 0$.

  $\rightsquigarrow$ Donc si f est dérivable en a, f est différentiable, et $Df(a) : h \in \mathbb{R} \mapsto f'(a)h \in F$.

- Gradient des fonctions à valeurs dans $\mathbb{R}$: Si $F = \mathbb{R}$ et $E = \mathbb{R}^n$, alors $Df(a) : \mathbb{R}^n \to \mathbb{R}$ est une forme linéaire. Mais alors, par le théorème de représentation de Riesz, il existe un unique vecteur $v \in \mathbb{R}^n$ tel que

$$\forall\, h \in \mathbb{R}^n, Df(a)(h) = \langle v, h \rangle$$

  où $\langle ., . \rangle$ est le produit scalaire habituel sur $\mathbb{R}^n$.

  On appelle ce vecteur gradient de f en a, et on le note $\nabla f(a)$. On a donc

$$\nabla f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix} \in \mathbb{R}^n$$

  Remarque: On peut en fait faire ça dès que E est un *espace de Hilbert* (mais pas si E est n'importe quel préhilbertien: il faut qu'il soit complet. On en reparlera.)

- Jacobienne des fonctions $\mathbb{R}^n \to \mathbb{R}^p$ Si $E = \mathbb{R}^n$, $F = \mathbb{R}^p$, notons

$$f : x = (x_1, \dots, x_n) \in U \subset \mathbb{R}^n \mapsto (f_1(x), \dots, f_p(x)) \in \mathbb{R}^p$$

  Alors; si f est différentiable en a, $Df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ admet une représentation matricielle (dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$): on appelle cette matrice la matrice jacobienne, notée $\operatorname{Jac} f(a)$.

  Donc, la j-ième colonne de $\operatorname{Jac} f(a)$ est donnée par les coordonnées du $Df(a)(e_j)$ dans la base canonique de $\mathbb{R}^p$.

  Or, $Df(a)(e_j)$ est la dérivée directionnelle dans la direction de e_j, autrement dit c'est la j-ième dérivée partielle de f en a:

$$Df(a)(e_j) = \frac{\partial f}{\partial x_j}(a) = \left(\frac{\partial f_1}{\partial x_j}(a), \dots, \frac{\partial f_p}{\partial x_j}(a) \right)$$

  La matrice jacobienne est donc donnée par

$$\operatorname{Jac} f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

- La i-ième ligne de $\mathrm{Jac}\, f(a)$ est la matrice de $Df_i(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$.
- La j-ième colonne de $\mathrm{Jac}\, f(a)$ est la j-ième dérivée partielle $\frac{\partial f}{\partial x_j}(a) \in \mathbb{R}^p$.

Lien entre toutes les façons de dériver

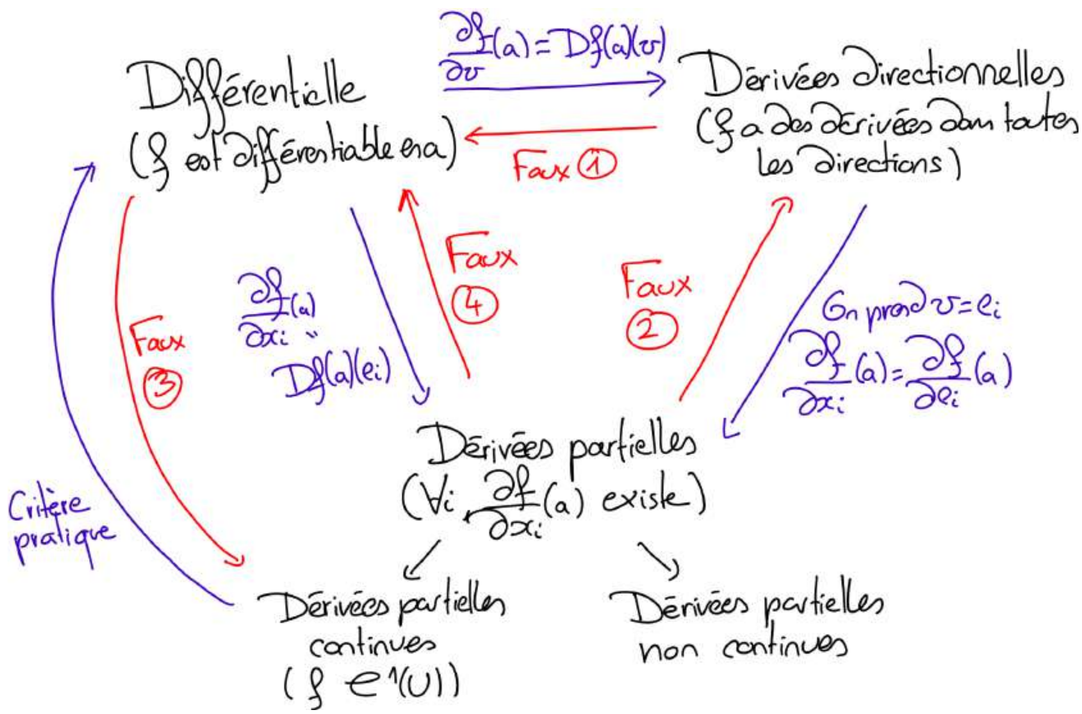


Figure 3: Les contre-exemples 1,2,3,4 sont sur la feuille de TD0

2 Applications de classe $\mathcal{C}^1$

Si f est différentiable en a pour tout $a \in U$, on peut donc définir une application

$$Df : a \in U \mapsto Df(a) \in \mathcal{L}(E, F)$$

où l'e.v. $\mathcal{L}(E, F)$ est muni de la norme des applications linéaires $\|L\|_{\mathcal{L}(E,F)} = \sup_{\|x\|_E=1} \|L(x)\|_F$.

⚠ Attention, DF n'est pas la différentielle de f: c'est une fonction définie sur l'ouvert U, *à valeurs dans l'e.v.n des applications linéaires continues $E \to F$*. Donc, pour tout $a \in U$, $Df(a)$ est linéaire, mais Df, par contre, n'est pas (nécessairement) linéaire.

Définition 8

On dit que f est continûment différentiable ($\mathcal{C}^1$) sur U si f est différentiable en tout $a \in U$ et si l'application

$$a \in U \subset E \mapsto Df(a) \in \mathcal{L}(E, F)$$

est continue. On note $\mathscr{C}^1(U)$ l'ensemble des applications continûment différentiables sur U.

Cas particulier: Dans le cas où $E = \mathbb{R}^n$, $F = \mathbb{R}^p$, $f \in \mathcal{C}^1(U)$ ssi l'application $a \in U \mapsto \operatorname{Jac} f(a) \in \mathcal{M}_{p,n}(\mathbb{R})$ est continue. C'est le cas si, et seulement si, les coefficients de la matrice $\operatorname{Jac} f(a)$ dépendent continûment de a sur U. Ce qui nous amène à:

Théorème 9

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. Alors

$$\begin{aligned} f \in \mathscr{C}^1(U) &\iff \forall\, i \in [\![1, n]\!],\ \left(x \in U \mapsto \frac{\partial f}{\partial x_i}(x) \in \mathbb{R}^p \right) \in \mathscr{C}^0(U, \mathbb{R}^p) \\ &\iff \forall\, i \in [\![1, n]\!],\ \forall\, j \in [\![1, p]\!],\ \frac{\partial f_j}{\partial x_i} \in \mathscr{C}^0(U, \mathbb{R}). \end{aligned}$$

Opérations sur les différentielles

Linéarité Soient $f, g : U \subset E \to F$, $\lambda, \mu \in \mathbb{R}$. Si f et g sont diff. en $a \in U$ alors $\lambda f + \mu g$ aussi et

$$\boxed{D(\lambda f + \mu g)(a) = \lambda Df(a) + \mu Dg(a)}$$

Composition 1 Soient $f : U \subset E \to F$ diff. en a et $g : V \subset F \to G$ t.q. $f(U) \subset V$ et g diff. en $f(a)$.
Alors $g \circ f : U \to G$ est diff. en a et

$$\boxed{D(g \circ f)(a) = Dg(f(a)) \circ Df(a)}$$

Composition 2 Si $E = \mathbb{R}^n$,$F = \mathbb{R}^p$, $G = \mathbb{R}^q$, alors cela donne:

$$\boxed{\operatorname{Jac} g \circ f(a) = \operatorname{Jac} g(f(a)) \cdot \operatorname{Jac} f(a)}$$

Composition 3 Si $E = \mathbb{R}^n$,$F = \mathbb{R}^p$, $G = \mathbb{R}$, alors $g \circ f : \mathbb{R}^n \to \mathbb{R}$ a des dérivées partielles en a données par:

$$\boxed{\frac{\partial (g \circ f)}{\partial x_i}(a) = \sum_{j=1}^{p} \frac{\partial g}{\partial y_j}(f(a)) \frac{\partial f_j}{\partial x_i}(a)}$$

Exemple: Considérons $f : x \in \mathbb{R}^n \setminus \{0\} \mapsto \|x\|_2 \in \mathbb{R}$.
$\rightsquigarrow$ On peut écrire $f = g \circ q$ avec

$$\begin{aligned} q : x \in \mathbb{R}^n \setminus \{0\} &\mapsto \langle x, x \rangle \in \mathbb{R}_+^* \\ g : u \in \mathbb{R}_+^* &\mapsto \sqrt{u} \in \mathbb{R} \end{aligned}$$

Alors, pour $x \in \mathbb{R}^n \setminus \{0\}$

- q est quadratique, donc diff. en $x \in \mathbb{R}^n \setminus \{0\}$ et $Dq(x)(h) = 2\langle x, h \rangle$;
- g est dérivable, donc différentiable en $u = q(x) = \langle x, x \rangle$ et $Dg(u)(t) = g'(u)(t) = \frac{t}{2\sqrt{u}}$;

donc f est différentiable en x et on a

$$Df(x)(h) = Dg(q(x))(Dq(x)(h)) = Dg(q(x))(2\langle x, h \rangle) = \frac{\langle x, h \rangle}{\sqrt{\langle x, x \rangle}} \text{ et}$$
$$\frac{\partial f}{\partial x_i}(x) = g'(f(x)) \frac{\partial q}{\partial x_i}(x) = \frac{1}{2\sqrt{\langle x, x \rangle}} 2\langle x, e_i \rangle = \frac{x_i}{\|x\|_2}$$

Dérivées partielles d'une application composée

Proposition 2

$$f : U \subset \mathbb{R}^n \to \mathbb{R}^p \text{ différentiable en } a \in U$$
$$g : V \in \mathbb{R}^p \to \mathbb{R} \text{ différentiable en } f(a) \in V$$

Alors $g \circ f : \mathbb{R}^n \to \mathbb{R}$ a des dérivées partielles en a données par:

$$\text{Pour } 1 \leq i \leq n, \frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{j=1}^{p} \frac{\partial g}{\partial y_j}(f(a)) \frac{\partial f_j}{\partial x_i}(a)$$

Exemple: Coordonnées polaires

$$\text{Considérons } f : (r, \theta) \in U = \mathbb{R}_+^* \times]0, 2\pi[\; \mapsto (r\cos\theta, r\sin\theta) \in \mathbb{R}^2$$
$$g : f(U) \to \mathbb{R} \text{ différentiable, et } h = g \circ f : (r, \theta) \mapsto g(r\cos\theta, r\sin\theta)$$

Alors pour $a = (r\cos\theta, r\sin\theta)$,

$$\frac{\partial h}{\partial r}(a) = \cos\theta \frac{\partial g}{\partial x}(a) + \sin\theta \frac{\partial g}{\partial y}(a) \quad \frac{\partial h}{\partial \theta}(a) = -r\sin\theta \frac{\partial g}{\partial x}(a) + r\cos\theta \frac{\partial g}{\partial y}(a)$$

Preuve: On a

$$\operatorname{Jac} g(f(a)) = \left(\frac{\partial g}{\partial y_1}(f(a)) \ldots, \frac{\partial g}{\partial y_p}(f(a)) \right),$$
$$\operatorname{Jac} f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \ldots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

En appliquant la formule pour le produit matriciel, on obtient

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = (\operatorname{Jac} g \circ f(a))_{1,i} = \sum_{k=1}^{p} (\operatorname{Jac} g(f(a))_{1k} \operatorname{Jac} f(a)_{ki})$$
$$= \sum_{k=1}^{p} \frac{\partial g}{\partial y_k}(f(a)) \cdot \frac{\partial f_k}{\partial x_i}(a).$$

□

3 Accroissements finis

Pour les fonctions réelles, le TAF permet de passer de l'information *locale* donnée par la dérivée à une information *globale*:

Théorème 10 (*Théorème des accroissements finis*)

Soit $f : [a, b] \to \mathbb{R}$ dérivable sur $]a, b[$ et continue sur $[a, b]$. Alors

$$\exists\, c \in]a, b[\;\; t.q. \;\; f(b) - f(a) = f'(c)(b - a).$$

$\rightsquigarrow$ Peut-on généraliser ce théorème ? En appliquant le TAF à la fonction

$$g : t \in [0,1] \mapsto f(a + t(b-a))$$

on démontre la généralisation suivante pour les fonctions définies sur E *à valeurs réelles*:

Proposition 3

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable sur un ouvert convexe U. Alors pour tous $a, b \in U$, il existe $c \in [a,b] = \{x \in U, \exists t \in [0,1], x = a + t(b-a)\}$ tel que

$$f(b) - f(a) = df(c)(b-a)$$

Toutefois, on ne pourra pas obtenir d'égalité pour les fonctions *à valeurs vectorielles.*

Contre-exemple: Considérons $\phi : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t)) \in \mathbb{R}^2$: ϕ est continue sur $[0, 2\pi]$ et dérivable sur $]0, 2\pi[$, mais

$$\phi(2\pi) - \phi(0) = (0,0) \neq 2\pi\phi'(c)$$

quel que soit $c \in]0, 2\pi[$, puisque $\|\phi'(c)\| = \|(-\sin(c), \cos(c))\| = 1$, donc $\phi'(c) \neq (0,0)$.

Inégalité des accroissements finis

On obtient en revanche l'inégalité suivante

Théorème 11

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une application $\mathscr{C}^1$ sur un ouvert convexe $U \subset \mathbb{R}^n$. Soient $a, b \in U$.

On suppose que $\sup_{c \in [a,b]} \|Df(c)\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} < \infty$. Alors

$$\|f(b) - f(a)\| \leq \max_{0 \leq t \leq 1} \|Df(a + t(b-a))\| \cdot \|b - a\|.$$

On l'utilise généralement sous la forme

Corollaire 1

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p \in \mathscr{C}^1(U)$. On suppose qu'il existe $C > 0$ tel que pour tout $x \in U$, $\|Df(x)\| \leq C$. Alors pour tous $a, b \in U$,

$$\|f(b) - f(a)\| \leq C\|b - a\|.$$

En particulier, si $Df(x) = 0$ pour tout $x \in U$, alors f est constante.

Preuve du théorème: Notons $M = \max_{0 \leq t \leq 1} \|Df(a + t(b-a))\|$. Il s'agit donc de montrer

$$\|f(b) - f(a)\| \leq M\|b - a\|.$$

Soit $\varepsilon > 0$, on va montrer que

$$\|f(b) - f(a)\| \leq (M + \varepsilon)\|b - a\|.$$

autrement dit que, pour $t = 1$,

$$\|f(a + t(b-a)) - f(a)\| \leq t(M + \varepsilon)\|b - a\|$$

autrement autrement dit, en posant

$$A_\varepsilon = \{t \in [0,1], \|f(a+t(b-a)) - f(a)\| \leq t(M+\varepsilon)\|b-a\|\},$$

on veut montrer que $1 \in A_\varepsilon$.

Remarquons que $A_\varepsilon = g_\varepsilon^{-1}([0,+\infty[)$, avec

$$g_\varepsilon(t) = t(M+\varepsilon)\|b-a\| - \|f(a+t(b-a)) - f(a)\| \in \mathcal{C}^0([0,1],\mathbb{R}).$$

Donc A_ε est un fermé de $[0,1]$ compact, donc c'est aussi un compact: A_ε admet donc un maximum t_0.
$\rightsquigarrow$ On va montrer que $t_0 = 1$. Supposons, par l'absurde, que $t_0 < 1$. Alors on a

$$\begin{aligned} \text{pour tout } t \in]t_0, 1], &\|f(a+t(b-a)) - f(a)\| > t(M+\varepsilon)\|b-a\| \\ \text{tandis que } &\|f(a+t_0(b-a)) - f(a)\| \leq t_0(M+\varepsilon)\|b-a\| \end{aligned}$$

donc, en soustrayant ces inégalités, par inégalité triangulaire inversée, on a:

$$\|f(a+t(b-a)) - f(a) - (f(a+t_0(b-a)) - f(a))\| \geq (t-t_0)(M+\varepsilon)\|b-a\|$$

ceci donne, en divisant par $t - t_0$,

$$\|Df(a+t_0(b-a))(b-a)\| \geq (M+\varepsilon)\|b-a\| \Rightarrow \|Df(a+t_0(b-a))\| \geq M+\varepsilon$$

ce qui est absurde. Donc $t_0 = 1$ et $1 \in A_\varepsilon$, ce qui donne

$$\|f(a+(b-a)) - f(a)\| = \|f(b) - f(a)\| \leq (M+\varepsilon)\|b-a\|$$

Ceci étant vrai pour tout ε, on obtient l'inégalité requise en faisant $\varepsilon \to 0$. □

4 $\mathscr{C}^1$-difféomorphismes

Définition 12

Soient U, V ouverts dans $\mathbb{R}^n$. Une application $f : U \to V$ est un $\mathscr{C}^1$-difféomorphisme si

- *f est $\mathscr{C}^1$ sur U*
- *f est bijective et l'application réciproque f^{-1} est de classe $\mathscr{C}^1$ sur V.*

⚠ Il ne suffit pas que f soit $\mathscr{C}^1$ et bijective. Par exemple, $f : x \in \mathbb{R} \mapsto x^3 \in \mathbb{R}$ est C^1 sur $\mathbb{R}$, bijective, mais la réciproque $f^{-1} : x \in \mathbb{R} \mapsto x^{\frac{1}{3}} \in \mathbb{R}$ n'est pas différentiable en 0.

Soit $f : U \to V$ un $\mathscr{C}^1$-difféomorphisme, alors $f^{-1} \circ f = \mathrm{Id}_U$. Donc, pour tout $x \in U$,

$$D(f^{-1})(f(x)) \circ Df(x) = D(f^{-1} \circ f)(x) = D\,\mathrm{Id}_U(x) = \mathrm{Id}_U$$

Autrement dit, $D(f^{-1})(f(x))$ est l'application linéaire inverse de $Df(x)$:

$$\boxed{D(f^{-1})(y) = Df(x)^{-1} \text{ pour } y = f(x) \in V.}$$

En particulier, $\mathrm{Jac}\, f(x)$ est une matrice inversible: on ne peut donc pas avoir de difféomorphisme entre des ouverts de $\mathbb{R}^n$ et $\mathbb{R}^p$ pour $n \neq p$.