Leçon 222 : Fonctions de plusieurs variables, dérivées partielles, différentiabilité, fonctions de classe $\mathcal{C}^1$. Notes.

Table des matières

1 Prérequis **2**

2 Fonctions d'une variable réelle **2**
2.1 Droite tangente 2

3 Dérivées directionnelles **3**
3.1 Dérivée directionnelle 3
3.2 Dérivée partielle 3
3.3 Quelques contre-exemples 3

4 Différentielle **4**
4.1 Définition 4
4.2 Matrice jacobienne 5
4.3 Gradient 5

5 Fonctions de classe $\mathcal{C}^1$ **6**
5.1 Définition 6
5.2 Théorème principal 6

6 Propriétés de la différentielle **8**

7 Somme, composition, produit **8**
7.1 Exemple : Fonctions à valeurs dans une sphère 10
7.2 Exemple : La norme 10
7.3 Exemple : Le déterminant 11

8 Autour du théorème des accroissements finis **12**

9 Références **13**

1 Prérequis

Connaissances préalables :
- ▷ Équivalence des normes en dimension finie.
- ▷ Dérivée d'une fonction réelle d'une variable réelle.
- ▷ Intégrales et primitives.

Dans ce qui suit, on se donne :
- ▷ p, $q \geq 1$ deux entiers.
- ▷ $U \subset \mathbb{R}^p$ et $V \subset \mathbb{R}^q$ deux ouverts non vides.
- ▷ $(\vec{e}_1, \ldots, \vec{e}_p)$ et $(\vec{f}_1, \ldots, \vec{f}_q)$ les bases canoniques de $\mathbb{R}^p$ et $\mathbb{R}^q$ respectivement.
- ▷ $f : U \to V$.

Exceptionnellement, comme en géométrie affine, nous utiliserons la notation $\vec{v}$ afin de mieux distinguer les *points* et les *vecteurs tangents*. Cette notion n'est pas à connaître, et en particulier les flèches peuvent être ignorées.

2 Fonctions d'une variable réelle

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. La dérivée en un point a de f est la limite, si elle existe,

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

Cette définition s'étend sans peine à des fonctions à valeurs vectorielle, disons dans $\mathbb{R}^q$. Ainsi, la dérivée en un point d'une fonction $f : \mathbb{R} \to \mathbb{R}^q$ est un vecteur (dont nous noterons les coordonnées en colonne), et la fonction dérivée une nouvelle fonction $\vec{f'} : \mathbb{R} \to \mathbb{R}^q$. La question se pose alors de définir la dérivée d'une fonction de plusieurs variables réelles.

2.1 Droite tangente

Remarquons au passage que la notion de dérivée permet de calculer les équations paramétriques de droites tangentes. Soit $f \in \mathcal{C}(\mathbb{R}, \mathbb{R}^q)$, dérivable en un point a, et telle que $\vec{f'(a)} \neq 0$. Alors la courbe représentative Γ de f admet une droite tangente T_a en a, d'équation paramétrique

$$T_a := \{f(a) + t\vec{f'(a)},\ t \in \mathbb{R}\}.$$

On peut voir cette équation comme une version linéarisée de la représentation paramétrique de Γ. En effet,

$$\Gamma = \{f(t),\ t \in \mathbb{R}\},$$

et linéariser f en a, c'est remplacer f par son approximation au premier ordre $f(a) + (t-a)\vec{f'(a)}$ en a. Cette idée se retrouve à de nombreuses reprises en calcul différentiel, par exemple dans le calcul d'équations de plans tangents.

Exemple 2.1.
Soit $f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$. Le graphe de f est la courbe $\Gamma = \{(t, f(t)),\ t \in \mathbb{R}\}$, c'est-à-dire la courbe représentative de la fonction $g : t \mapsto (t, f(t))$. Or $g'(t) = (1, f'(t))$, donc la tangente au graphe de f en $(a, f(a))$ a pour équation paramétrique

$$T_a = \{(a, f(a)) + t(1, f'(a)),\ t \in \mathbb{R}\} = \{(t, f(a) + (t-a)f'(a)),\ t \in \mathbb{R}\}.$$

Si $f'(a) = 0$, la situation est nettement plus compliquée ; la courbe représentative peut par exemple exhiber des points de rebroussement. On pourra dessiner en exercice la courbe représentative de la fonction $t \mapsto (t^2, t^3)$, et en approfondissement se reporter au cours sur les séries de Taylor.

3 Dérivées directionnelles

Une première tentative consiste à faire apparaître des fonctions d'une seule variable réelle en prenant des "tranches" de l'espace de départ. Cela mène aux notions de dérivée directionnelle et de dérivée partielle.

3.1 Dérivée directionnelle

Définition 3.1 (Dérivée directionnelle)**.**
*Soit $a \in U$ et $\vec{u} \in \mathbb{R}^p$. La **dérivée directionnelle** de f en a dans la direction $\vec{u}$ est la dérivée en 0 de la fonction $t \mapsto f(a + t\vec{u})$, c'est-à-dire la limite, si elle existe,*

$$\lim_{t \to 0} \frac{f(a + t\vec{u}) - f(a)}{t}.$$

On note alors cette quantité $D_a f(\vec{v})$.

Attention : D'autres notations pour la dérivée directionnelle, en particulier les notations $D_{\vec{v}} f(a)$ et $d_a f(\vec{v})$, existent. La notation précédente me semble l'une des plus cohérentes avec ce qui suit.

3.2 Dérivée partielle

Définition 3.2 (Dérivée partielle)**.**
*Soit $1 \leq i \leq p$. La i-ème **dérivée partielle** de f est la dérivée directionnelle de f dans la direction $\vec{e}_i$. On la note*

$$\frac{\partial f}{\partial x_i} = \lim_{t \to 0} \frac{f(a + t\vec{e}_i) - f(a)}{t}.$$

Autrement dit, pour calculer $\frac{\partial f}{\partial x_i}$ en un point, on fixe les coordonnées $(x_j)_{j \neq i}$, et on dérive la fonction d'une variable réelle $x_i \mapsto f(x_1, \ldots, x_p)$.
Attention : Là encore, d'autres notations existent, comme par exemple $\partial_{x_i} f$. On prendra garde aussi à ce que, dans une expression du type

$$\frac{\partial f}{\partial x_i}(x_1, \ldots, x_n),$$

le symbole x_i joue deux rôles distincts : direction dans laquelle on dérive au dénominateur, et une des coordonnées en laquelle on évalue la fonction $\frac{\partial f}{\partial x_i}$ au numérateur.

3.3 Quelques contre-exemples

Exemple 3.3.
La fonction f définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$ par $f(x,y) = \frac{xy}{x^2+y^2}$ et telle que $f(0,0) = 0$ admet des dérivées partielles en tout point :

- ▷ *Si $(x,y) \neq (0,0)$, alors $\frac{\partial f}{\partial x}(x,y) = \frac{y(y^2-x^2)}{(x^2+y^2)^2}$, et $\frac{\partial f}{\partial y}(x,y) = \frac{x(x^2-y^2)}{(x^2+y^2)^2}$.*
- ▷ *$f(x,0) = 0$ pour tout x, donc $\frac{\partial f}{\partial x}(0,0) = 0$. De même, $\frac{\partial f}{\partial y}(0,0) = 0$.*

Cependant, f n'est pas continue. En effet, en coordonnées polaires (r, θ), on calcule $f(x,y) = \frac{\sin(2\theta)}{2}$, donc l'ensemble des valeurs d'adhérence de f en $(0,0)$ est $[-1/2, 1/2]$.

Exemple 3.4.
La fonction f définie sur $\mathbb{R}^2 \setminus \{0\} \times \mathbb{R}$ par $f(x,y) = \frac{y^2}{x}$ et telle que $f(0,y) = 0$ pour tout y admet des dérivées directionnelles en toute direction en $(0,0)$, mais n'est pas continue en $(0,0)$. En effet, pour tout vecteur non nul $\vec{u} = (\alpha, \beta)$ et tout $t \in \mathbb{R}^*$, on a $f(t\vec{u}) = t\beta^2/\alpha$ si $\alpha \neq 0$, et $f(t\vec{u}) = 0$ si $\alpha = 0$. Dans tous les cas, $D_{\vec{u}} f(0,0) = 0$.

Exemple 3.5.
La fonction f définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$ par $f(x,y) = \frac{x^2y}{x^4+y^2}$ et telle que $f(0,0) = 0$ admet des dérivées directionnelles en tout point et toute direction, mais n'est pas continue en $(0,0)$.

La simple existence de dérivées directionnelle est donc trop faible pour contrôler une fonction.

4 Différentielle

Une seconde approche consiste à linéariser la définition de la dérivée.

4.1 Définition

Soit $f \in \mathcal{C}(\mathbb{R}, \mathbb{R}^q)$. Remarquons que

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

si et seulement si

$$f(a+h) = f(a) + f'(a)h + o(|h|),$$

autrement dit si f admet un développement limité à l'ordre 1 en a. Cette seconde version a l'avantage de ne pas diviser par un vecteur, et donc de pouvoir se généraliser à des fonctions de plusieurs variables réelles. Un développement limité à l'ordre 1 admet alors une partie constante et une partie linéaire (polynômiale homogène de degré 1).

Définition 4.1 (Différentiabilité)**.**
Soit $a \in U$. On dit que f est ***différentiable*** *en a s'il existe une application linéaire $D_af : \mathbb{R}^p \to \mathbb{R}^q$ telle que*

$$f\left(a + \vec{h}\right) = f\left(a\right) + D_af\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right).$$

Attention : N'essayez pas d'affaiblir cette définition, cela ne marchera pas ! Même si f admet des dérivées directionnelles dans toute direction et si ces dérivées directionnelles dépendent linéairement de la direction, il se peut encore que f ne soit pas différentiable !

Lemme 4.2.
Si f est différentiable en a, alors l'application D_af est unique. On l'appelle ***différentielle*** *de f en a.*

Démonstration.
Le point crucial est que f est définie sur un ouvert [1]. En prenant $\vec{h} = \vec{e}_i$, on montre que $D_af_j(\vec{e}_i)) = \frac{\partial f_j}{\partial x_i}(a)$. Par conséquent, la matrice de D_af dans les bases canoniques de $\mathbb{R}^p$ et $\mathbb{R}^q$ sont déterminées par f. Voir aussi ce qui suit sur la notion de matrice jacobienne. □

Lemme 4.3.
Si f est différentiable en a, alors f est continue en a.

Démonstration.
En effet, D_af est de norme d'opérateur finie, donc $f\left(a + \vec{h}\right) = f\left(a\right) O\left(\left\|\vec{h}\right\|\right)$. □

1. Ou, à tout le moins, un ensemble "assez gros en tout point" (techniqueement, qui en tout point a un cône tangent d'intérieur non vide). Par exemple, tout se passe bien pour une fonction $f : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$, mais on peut avoir des soucis en $(0,0)$ avec une fonction $f : \{(x,y) : \ |y| \leq x^2\} \to \mathbb{R}$.

Exemple 4.4.
La fonction $f : x \mapsto \|x\|^2$ *de* $\mathbb{R}^p$ *dans* $\mathbb{R}$ *est différentiable en tout point. En effet, pour tous* x *et* $\vec{h}$*,*

$$\begin{aligned} f(x+\vec{h}) &= \langle x+\vec{h}, x+\vec{h}\rangle \\ &= \langle x, x\rangle + 2\langle x, \vec{h}\rangle + \langle \vec{h}, \vec{h}\rangle \\ &= f(x) + L_x(\vec{h}) + O\left(\left\|\vec{h}\right\|^2\right), \end{aligned}$$

où $L_x(\vec{h}) = 2\langle x, \vec{h}\rangle$*. Par conséquent,* f *est différentiable en* x*, de différentielle* $\vec{h} \mapsto 2\langle x, \vec{h}\rangle$*.*

Exemple 4.5.
La fonction $g : M \mapsto (M^T)M$ *de* $\mathcal{M}_n(\mathbb{R})$ *dans* $\mathcal{M}_n(\mathbb{R})$ *est différentiable en tout point. En effet, par un calcul similaire au précédent, pour tous* M *et* H*,*

$$\begin{aligned} g(M+H) &= (M^T)M + (H^T)M + (M^T)H + (H^T)H \\ &= g(M) + (H^T)M + (M^T)H + O\left(\|H\|^2\right), \end{aligned}$$

Par conséquent, f *est différentiable en* M*, de différentielle* $H \mapsto (H^T)M + (M^T)H$*.*

4.2 Matrice jacobienne

Définition 4.6 (Matrice jacobienne)**.**
Supposons que f *admette des dérivées partielles en un point* a*. La* ***matrice jacobienne*** *de* f *en* a *est la matrice de ses dérivées partielles en* a *:*

$$Jac_a f = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1\le i\le q \\ 1\le j\le p}}.$$

Lemme 4.7.
Si f *est différentiable en* a*, alors* $Jac_a f$ *est la matrice de* $D_a f$ *dans les bases canoniques de* $\mathbb{R}^p$ *et* $\mathbb{R}^q$*. Autrement dit,*

$$f(a+\vec{h}) = f(a) + \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a) \cdot \vec{h}_i.$$

Remarque 4.8.
Il faut éviter de confondre les lignes et les colonnes ! Pour s'en souvenir :
- ▷ *Si* $f : \mathbb{R} \to \mathbb{R}^q$*, alors la différentielle est un vecteur colonne, comme vu dans le cadre de fonctions d'une variable réelle. Il n'y a qu'une seule variable, donc* $Jac_a f$ *a bien une colonne par variable.*
- ▷ *Si* $f : \mathbb{R}^p \to \mathbb{R}$*, alors la différentielle est un vecteur ligne. Donc* $Jac_a f$ *a bien une ligne par dimension de l'espace d'arrivée.*

4.3 Gradient

Définition 4.9 (Gradient)**.**
Soit $f : U \to \mathbb{R}$ *différentiable en* a*. Dans ce cas,* $D_a f$ *est une forme linéaire sur* $\mathbb{R}^p$*, dont la matrice dans la base canonique est un vecteur ligne. Le* ***gradient*** *de* f *en* a *est le vecteur* $\vec{\nabla} f \in \mathbb{R}^p$ *tel que, pour tout* $\vec{h} \in \mathbb{R}^p$*,*

$$\langle \vec{\nabla} f, \vec{h}\rangle = D_a f(\vec{h})$$

Les coordonnées du gradient dans une bases orthonormée sont les transposées des coordonnées de $D_a f$:

$$\vec{\nabla} f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_p} \end{pmatrix}.$$

Le gradient est moins naturel, mais a plusieurs avantages :

- ▷ À un niveau élémentaire, travailler avec des vecteurs colonnes, là où les étudiants ne sont pas toujours à l'aise avec des vecteurs lignes.
- ▷ Donner une interprétation géométrique des lignes de niveau (elles sont orthogonales à la ligne de gradient, c'est-à-dire à la ligne de plus forte pente), similaire à l'interprétation apportée aux équations de droites ou de plans à l'aide des vecteurs normaux.
- ▷ À plus haut niveau, permettre de dériver des champs de vecteurs en géométrie riemannienne.

Exemple 4.10.
Reprenons le cas de la fonction $f : x \mapsto \|x\|^2$. Sa différentielle en x est la forme linéaire $x \mapsto 2\langle x, \vec{h}\rangle$. Par identification, $\vec{\nabla} f(x) = 2x$. Ce résultat se retrouve aussi en calculant les dérivées partielles, même s'il faut alors admettre – pour l'instant – que f est effectivemment dérivable. En effet, $f(x) = \sum_{k=1}^{p} x_k^2$, donc, pour tout k, on a $\frac{\partial f}{\partial x_k} = 2x_k$. On retrouve alors l'identité $\vec{\nabla} f(x) = 2x$.

5 Fonctions de classe $\mathcal{C}^1$

5.1 Définition

Définition 5.1 (Fonctions de classe $\mathcal{C}^1$)**.**
*On dit que f est **de classe** $\mathcal{C}^1$ si f est différentiable en tout point, et si la fonction*

$$\begin{cases} U & \rightarrow & \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ a & \mapsto & D_a f \end{cases}$$

est continue. On définit récursivement les fonctions de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N} \cup \{\infty\}$.

5.2 Théorème principal

Le résultat le plus important de cette leçon est le suivant :

Théorème 1.
Une fonction f est de classe $\mathcal{C}^1$ si et seulement si ses dérivées partielles existent et sont continues.

Grâce à ce théorème, on peut montrer facilement que des fonctions sont de classe $\mathcal{C}^1$.

Remarque 5.2.
Récursivement, une fonction est de classe $\mathcal{C}^k$ si et seulement si toutes ses dérivées partielles d'ordre $\leq k$ existent et sont continues, où une dérivée partielle d'ordre ℓ est de la forme

$$\frac{\partial^\ell f}{\partial x_{i(1)} \ldots \partial x_{i(\ell)}} = \frac{\partial}{\partial x_{i(1)}} \cdots \frac{\partial f}{\partial x_{i(\ell)}}.$$

Comme nous le verrons dans une leçon ultérieure, l'ordre des opérateurs de différentiation partielle n'a alors pas d'importance (lemme de Schwartz). En particulier, pour calculer la différentielle k-ième d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, il suffit de calculer non pas n^k dérivées partielles, mais seulement C_k^{n+k-1} dérivées partielles[2].

2. Petit exercice de combinatoire.

La démonstration la plus fréquente passe par le théorème des acroissements finis. Pour des raisons esthétiques personnelles, la démonstration qui suit utilise seulement le théorème fondamental de l'analyse.

Démonstration.
Le sens direct est facile : si f est de classe $\mathcal{C}^1$, alors $D_a f$ existe et dépend continûment de a, donc les matrices de $D_a f$ dans la base canonique $\mathrm{Jac}_a f$ existent et dépendent continûment de a, donc les dérivées partielles existent et dépendent continûment de a.

Démontrons le sens indirect. Il faut réussir à contrôler $f(a+\vec{h})$ pour tous les vecteurs $\vec{h}$ suffisamment petits. Cependant, les dérivées partielles ne permettent de contrôler f que le long de droites parallèles à un des axes de $\mathbb{R}^p$. Nous allons donc utiliser la décomposition $\vec{h} = \sum_{k=1}^p h_k \vec{e}_k$. Pour tous $1 \leq i \leq q$, tout $a \in U$ et tout $\vec{h}$ suffisamment petit,

$$\begin{aligned} f_i(a+\vec{h}) - f_i(a) &= \sum_{k=1}^p \left[f_i \left(a + \sum_{j=1}^k h_j \vec{e}_j \right) - f_i \left(a + \sum_{j=1}^{k-1} h_j \vec{e}_j \right) \right] \\ &= \sum_{k=1}^p \int_0^{h_k} \frac{\partial f_i}{\partial x_k} \left(a + \sum_{j=1}^{k-1} h_j \vec{e}_j + t\vec{e}_k \right) \, \mathrm{d}t. \end{aligned}$$

Soit $L_a f$ l'opérateur linéaire dont la matrice dans la base canonique est $\mathrm{Jac}_a f$. Alors

$$\begin{aligned} \left| f_i(a+\vec{h}) - f_i(a) - L_a f(\vec{h})_i \right| &= \left| \sum_{k=1}^p \left[\int_0^{h_k} \frac{\partial f_i}{\partial x_k} \left(a + \sum_{j=1}^{k-1} h_j \vec{e}_j + t\vec{e}_k \right) \, \mathrm{d}t - h_k \frac{\partial f_i}{\partial x_k}(a) \right] \right| \\ &\leq \sum_{k=1}^p \int_0^{|h_k|} \left| \frac{\partial f_i}{\partial x_k} \left(a + \sum_{j=1}^{k-1} h_j \vec{e}_j + t\vec{e}_k \right) - \frac{\partial f_i}{\partial x_k}(a) \right| \, \mathrm{d}t. \end{aligned}$$

Soit $\varepsilon > 0$. Par continuité des dérivées partielles, il existe $\delta > 0$ tel que, si $\|b-a\|_\infty \leq \delta$, alors $\left| \frac{\partial f_i}{\partial x_k}(b) - \frac{\partial f_i}{\partial x_k}(a) \right| \leq \varepsilon$ pour toux i et k. Alors, si $\left\| \vec{h} \right\|_\infty \leq \delta$,

$$\left| f_i(a+\vec{h}) - f_i(a) - L_a f(\vec{h})_i \right| \leq \sum_{k=1}^p \int_0^{|h_k|} \varepsilon \, \mathrm{d}t = \varepsilon \sum_{k=1}^p |h_k| \leq p\varepsilon \left\| \vec{h} \right\|_\infty.$$

Autrement dit, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, si $\left\| \vec{h} \right\|_\infty < \delta$, alors

$$\left| f_i(a+\vec{h}) - f_i(a) - L_a f(\vec{h})_i \right| \leq \varepsilon \left\| \vec{h} \right\|_\infty \ ;$$

ou, en résumé,

$$\left\| f(a+\vec{h}) - f(a) - L_a f(\vec{h}) \right\|_\infty = o\left(\left\| \vec{h} \right\|_\infty \right).$$

Ainsi, f est différentiable en a, de différentielle $L_a(f)$. □

Exemple 5.3.
Grâce à ce théorème, on peut par exemple montrer que la fonction $x \mapsto \|x\|^2$ est de classe $\mathcal{C}^1$ seulement à partir de ses dérivées partielles.

Exemple 5.4.
La fonction

$$f : \begin{cases} \mathbb{R}_+^* \times (-\pi, \pi) & \to \quad \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\}) \\ (r, \theta) & \mapsto \quad \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix} \end{cases}$$

a pour matrice jacobienne :

$$Jac_{(r,\theta)} f = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

Cette dernière est une fonction continue de r et θ, donc f est de classe $\mathcal{C}^1$. De plus, $\det(Jac_{(r,\theta)} f) = r$ *ne s'annule pas.*

6 Propriétés de la différentielle

7 Somme, composition, produit

Propriété 7.1 (Somme)**.**
Soient f, $g : U \to \mathbb{R}^q$. Soit $a \in U$. Si f et g sont différentiables en a, alors $f + g$ aussi, et

$$D_a(f+g) = D_a f + D_a g.$$

En particulier, une somme de fonctions de classe $\mathcal{C}^1$ est de classe $\mathcal{C}^1$.

Démonstration.
Pour tout vecteur $\vec{h}$ suffisamment petit,

$$\begin{aligned} (f+g)\left(a+\vec{h}\right) &= f(a) + D_a f\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right) + g(a) + D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right) \\ &= (f+g)(a) + (D_a f + D_a g)(\vec{h}) + o\left(\left\|\vec{h}\right\|\right), \end{aligned}$$

donc $f + g$ est différentiable en a et $D_a(f+g) = D_a f + D_a g$. □

Propriété 7.2 (Composition)**.**
Soient p, q, r trois entiers strictement positifs. Soient U, V, W trois ouverts non vides de $\mathbb{R}^p$, $\mathbb{R}^q$, $\mathbb{R}^r$ respectivement. Soient $f : U \to V$ et $g : V \to W$. Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a, et

$$D_a(g \circ f) = (D_{f(a)} g)(D_a f).$$

En particulier, la composée de fonctions de classe $\mathcal{C}^1$ est de classe $\mathcal{C}^1$.

Démonstration.
Pour tout vecteur $\vec{h}$ suffisamment petit,

$$\begin{aligned} f \circ g\left(a+\vec{h}\right) &= f\left(g(a) + D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right) \\ &= f(g(a)) + D_{g(a)} f\left(D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right) + o\left(D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right) \\ &= (g(a)) + D_{g(a)} f \cdot D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right), \end{aligned}$$

donc $f \circ g$ est différentiable en a et $D_a(f \circ g) = D_{g(a)} f \cdot D_a g$. □

Remarque 7.3.
Il est ici important de ne pas inverser les lignes et les colonnes de la matrice jacobienne !

Exemple 7.4.
Soit $f : U \to \mathbb{R}$ et $\gamma : [a,b] \to U$ un chemin. Supposons f et γ de classe $\mathcal{C}^1$. Alors

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b (f \circ \gamma)'(t)\ \mathrm{d}t = \int_a^b (D_{\gamma(t)} f)(\vec{\gamma}'(t))\ \mathrm{d}t = \int_a^b \langle \vec{\nabla}_{\gamma(t)} f, \vec{\gamma}'(t) \rangle\ \mathrm{d}t.$$

En coordonnées,

$$\begin{aligned}(f \circ \gamma)'(t) &= \langle \vec{\nabla}_{\gamma(t)} f, \vec{\gamma}'(t) \rangle \\ &= \sum_{k=1}^{p} \frac{\partial f}{\partial x_k}(\gamma(t)) \cdot \gamma_k'(t).\end{aligned}$$

Souvent, le chemin γ est noté $\gamma = (x_k(t))_{1 \leq k \leq p}$, de telle sorte que $f \circ \gamma(t) = f(x_1(t), \ldots, x_p(t))$. Alors, en notation plus "physicienne", on retrouve la formule de dérivation en chaîne :

$$(f \circ \gamma)' = \sum_{k=1}^{p} \frac{\partial f}{\partial x_k} \cdot \frac{\mathrm{d}x_k}{\mathrm{d}t}.$$

Propriété 7.5 (Produit)**.**
Soient f, $g : U \to \mathbb{R}$. Soit $a \in U$. Si f et g sont différentiables en a, alors fg aussi, et

$$D_a(fg) = f(a) \cdot D_a g + g(a) \cdot D_a f.$$

En particulier, un produit de fonctions de classe $\mathcal{C}^1$ est de classe $\mathcal{C}^1$.

Démonstration.
Nous fournissons deux démonstrations, une directe, et une utilisant les outils précédents. Pour la première,

$$\begin{aligned}(fg)\left(a + \vec{h}\right) &= (fg)(a) + \left[f\left(a + \vec{h}\right) - f(a)\right] g(a) + f(a)\left[g\left(a + \vec{h}\right) - g(a)\right] \\ &\quad + \left[f\left(a + \vec{h}\right) - f(a)\right]\left[g\left(a + \vec{h}\right) - g(a)\right] \\ &= (fg)(a) + g(a)\left[D_a f\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right] + f(a)\left[D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right] \\ &\quad + \left[D_a f\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right]\left[D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right)\right] \\ &= (fg)(a) + g(a) D_a f\left(\vec{h}\right) + D_a g\left(\vec{h}\right) + o\left(\left\|\vec{h}\right\|\right),\end{aligned}$$

donc fg est dérivable en a et $D_a(fg) = f(a) \cdot D_a g + g(a) \cdot D_a f$.
La seconde démonstration utilise le théorème de dérivation en chaîne. Soit $P : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $P(x,y) = xy$. Alors P est $\mathcal{C}^1$ et sa dérivée en (x,y) est l'application linéaire $D_{(x,y)}P$: $(h,k) \mapsto yh + xk$ (la technique de calcul est la même que pour la dérivée de $x \mapsto \|x\|^2$). Soit $Q : \mathbb{R}^p \to \mathbb{R}^2$ la fonction définie par $x \mapsto (f(x), g(x))$. Alors Q est dérivable en a et sa dérivée est $(D_a f, D_a g)$. Par la formule de dérivation des fonctions composées, $fg = P \circ Q$ est différentiable en a, et sa différentielle vaut

$$\begin{aligned}D_a(fg)\left(\vec{h}\right) &= D_{Q(a)}P \cdot D_a Q\left(\vec{h}\right) \\ &= D_{(f(a),g(a))}P \cdot \left(D_a f\left(\vec{h}\right), D_a g\left(\vec{h}\right)\right) \\ &= g(a) D_a f\left(\vec{h}\right) + f(a) D_a g\left(\vec{h}\right).\end{aligned}$$

□

7.1 Exemple : Fonctions à valeurs dans une sphère

Soient $M \in \mathbb{R}^n$ et $\gamma \in \mathcal{C}^1(I, \mathbb{R}^n)$. Soit $t_0 \in I$; posons $R := d(\gamma(t_0), M) = \left\| \overrightarrow{M\gamma(t_0)} \right\|$.

Proposition 7.6.
La courbe paramétrée γ est à valeurs dans la sphère $S(M, R)$ si et seulement si $\overrightarrow{\gamma'(t)} \perp \overrightarrow{M\gamma(t)}$ pour tout $t \in I$.

Démonstration.
Posons $f(t) := \left\| \overrightarrow{M\gamma(t)} \right\|^2$ pour tout $t \in I$. On sait que $f(t_0) = R^2$. Donc γ est à valeurs dans $S(M, R)$ si et seulement si $f' \equiv 0$. Or, par composition de fonctions,

$$f'(t) = 2\langle \overrightarrow{M\gamma(t)}, \overrightarrow{\gamma'(t)} \rangle,$$

ce qui finit la démonstration. $\square$

Corollaire 7.7.
Soit $a \in M_n(\mathbb{R})$. Il y a équivalence entre :
- ▷ *A est antisymétrique ;*
- ▷ *e^{tA} est une isométrie pour tout $t \in \mathbb{R}$.*

Démonstration.
Soit $x \in \mathbb{R}^n$. On applique la proposition précédente à $\gamma(t) := e^{tA}x$, en remarquant que $\overrightarrow{\gamma'(t)} = Ae^{tA}x = A\gamma(t)$. Alors $\left\| e^{tA}x \right\|$ est constant si et seulement si $\langle A\gamma(t), \gamma(t) \rangle = 0$ pour tout $t \in \mathbb{R}$.

Supposons que A soit antisymétrique. Alors $\langle A\gamma(t), \gamma(t) \rangle = -\langle \gamma(t), A\gamma(t) \rangle = -\langle A\gamma(t), \gamma(t) \rangle$, donc $\langle A\gamma(t), \gamma(t) \rangle = 0$. Ceci étant vrai pour tout $t \in \mathbb{R}$, par le paragraphe précédent, $\left\| e^{tA}x \right\| = \|x\|$ pour tout $t \in \mathbb{R}$. Enfin, ceci étant valable pour tout $x \in \mathbb{R}^n$, on a montré que e^{tA} était une isométrie pour tout t.

Supposons maintenant que e^{tA} soit une isométrie pour tout $t \in \mathbb{R}$. Alors, en particulier, $\langle Ax, x \rangle = \langle A\gamma(0), \gamma(0) \rangle = 0$, et ce pour tout $x \in \mathbb{R}^n$. La forme quadratique $x \mapsto \langle Ax, x \rangle$ est donc nulle, donc A est antisymétrique. $\square$

Remarque 7.8.
Si l'un des conditions du théorème précédent est satisfaite, alors $e^{tA} \in SO_n(\mathbb{R})$ pour tout $t \in \mathbb{R}$. En effet, comme e^{tA} est une isométrie, on en déduit que $\det(e^{tA}) \in \{\pm 1\}$ pour tout t. Or la fonction $t \mapsto \det(e^{tA})$ est continue, et $\det(I) = 1$. Par connexité, $\det(e^{tA}) = 1$ pour tout t.

7.2 Exemple : La norme

Sur $\mathbb{R}^n \setminus \{(0,0)\}$, la fonction $x \mapsto \|x\|$ est la composée de $g : x \mapsto \|x\|^2$ (de $\mathbb{R}^n \setminus \{(0,0)\}$ dans $\mathbb{R}_+^*$) et de $f : y \mapsto \sqrt{y}$ (de $\mathbb{R}_+^*$ dans $\mathbb{R}_+^*$), qui sont toutes les deux de classe $\mathcal{C}^\infty$. Sa différentielle vaut donc

$$\begin{aligned} D_x f \circ g \left(\vec{h}\right) &= D_{g(x)} \cdot D_x g \left(\vec{h}\right) \\ &= \frac{1}{2\sqrt{g(x)}} \cdot 2\langle x, \vec{h} \rangle \\ &= \frac{1}{\|x\|} \cdot \langle x, \vec{h} \rangle. \end{aligned}$$

En particulier, le gradient de la fonction $\|\cdot\|$ vaut $\frac{x}{\|x\|}$, c'est-à-dire que le gradient de la norme en x est le vecteur unitaire de direction x.

La fonction norme n'est pas dérivable en 0 ; cela peut se voir en prenant une fonction partielle, où l'on retrouve la fonction valeur absolue, qui n'est pas dérivable en 0.

Exercice 7.9.
Retrouvez ce qui précède en calculant directement les dérivées partielles de la fonction $f(x) = \sqrt{\sum_{k=1}^{p} x_k^2}$.

7.3 Exemple : Le déterminant

La fonction det : $\mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est polynômiale en les coefficients de la matrice, donc ses dérivées partielles le sont aussi, donc cette fonction est de classe $\mathcal{C}^1$ par ce qui précède. Calculons pour commencer sa dérivée en I.

Une base de $\mathcal{M}_n(\mathbb{R})$ est $(A_{ij})_{1\le i,j\le n}$, où A_{ij} est la matrice dont le coefficient (i,j) vaut 1 et les autres coefficients 0. Alors $\det(I + tA_{ij}) = 1$ si $i \neq j$, car la matrice $I + tA_{ij}$ est alors triangulaire supérieure ou inférieure avec seulement des 1 sur la diagonale. De plus, $\det(I + tA_{ii}) = 1 + t$ pour tout i. Par conséquent, en dérivant ces fonctions de t,

$$\frac{\partial \det}{\partial x_{ij}}(I) = \delta_{ij},$$

et donc

$$D_I \det(H) = \sum_{1\le i,j\le n} \frac{\partial \det}{\partial x_{ij}}(I) H_{ij} = \sum_{1\le i,j\le n} \delta_{i,j} H_{ij} = \sum_{i=1}^{n} H_{ii} = \mathrm{Tr}(H).$$

On peut utiliser cette formule pour dériver le déterminant ailleurs. Soit $M \in \mathrm{GL}_n(\mathbb{R})$. Alors $\det(M + H) = \det(M)\det(I + M^{-1}H) = \det(M) + \det(M)\mathrm{Tr}(M^{-1}H) + o(\|H\|)$, donc

$$D_M \det(H) = \det(M)\mathrm{Tr}(M^{-1}H).$$

Cette dernière formule n'a cependant pas de sens si M n'est pas inversible. Cependant, on peut se souvenir de la formule de Cramer :

$$M^{-1} = \det(M)^{-1}\mathrm{Com}(M)^T \quad \forall M \in \mathrm{GL}_n(\mathbb{R}),$$

pour réécrire

$$D_M \det(H) = \mathrm{Tr}(\mathrm{Com}(M)^T H) \quad \forall M \in \mathrm{GL}_n(\mathbb{R}).$$

Finalement, le déterminant étant polynômial (donc de classe $\mathcal{C}^1$, la fonction $M \mapsto D_M \det$ est continue. De plus, la fonction $M \mapsto \mathrm{Tr}(\mathrm{Com}(M)^T \cdot)$ est aussi continue [3], et ces deux fonctions coïncident sur l'ensemble $\mathrm{GL}_n(\mathbb{R})$ qui est dense dans $M_n(\mathbb{R})$. Par densité, ces deux fonctions coïncident sur $M_n(\mathbb{R})$:

$$D_M \det(H) = \mathrm{Tr}(\mathrm{Com}(M)^T H) \quad \forall M \in M_n(\mathbb{R}).$$

Remarque 7.10.
On observe que la base $(A_{ij})_{1\le i,j\le n}$ *est orthonormée pour le produit scalaire canonique sur* $M_n(\mathbb{R})$ *défini par* $\langle A, B\rangle = \mathit{Tr}(A^T B)$. *Mais alors* $D_M \det(H) = \langle \mathit{Com}(M), H\rangle$, *c'est-à-dire que*

$$\vec{\nabla}_M \det = \mathit{Com}(M) \quad :$$

la comatrice est le gradient du déterminant.

3. Ces fonctions vont de $M_n(\mathbb{R})$ dans $M_n(\mathbb{R})^*$.

Application 7.11.
Soit $A \in M_n(\mathbb{R})$. Considérons la fonction

$$f : \begin{cases} \mathbb{R} & \to & M_n(\mathbb{R}) \\ t & \mapsto & \det(e^{tA}) \end{cases}$$

La fonction $\gamma : t \mapsto e^{tA}$ a pour dérivée $\gamma' : t \mapsto Ae^{tA}$. Par la formule de dérivation en chaîne, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} f'(t) &= D_{e^tA} \det \cdot \gamma'(t) \\ &= \det(e^{tA})\, Tr(e^{-tA} \cdot Ae^{tA}) \\ &= Tr(e^{tA}e^{-tA}A) \det(e^{tA}) \\ &= Tr(A) f(t). \end{aligned}$$

Par conséquent, la fonction f est solution de l'équation différentielle $u' = \det(A)u$ avec condition initiale $u(0) = \det(I) = 1$, et donc, pour tout $t \in \mathbb{R}$,

$$f(t) = \det(e^{tA}) = e^{Tr(A)t}.$$

En choisissant $t = 1$, on retrouve l'identité $\det(e^A) = e^{Tr(A)}$.

Corollaire 7.12.
Soit $A \in M_n(\mathbb{R})$. Il y a équivalence entre :
- ▷ *$Tr(A) = 0$;*
- ▷ *e^{tA} préserve le volume pour tout $t \in \mathbb{R}$.*

8 Autour du théorème des accroissements finis

Le théorème des accroissements finis est faux si la dimension de l'espace d'arrivée q est supérieure ou égale à 2 ; considérer par exemple une fonction de $\mathbb{R}$ dans $\mathbb{R}^3$ qui paramètre une spirale, ou une fonction de $\mathbb{R}$ dans $\mathbb{R}^2$ qui paramètre la lettre γ. Il existe des variantes plus techniques, que nous n'aborderons pas ; nous ne mentionnerons que l'*inégalité* des accroissements finis. De plus, la géométrie du domaine est importante.

Théorème 2 (Inégalité des accroissements finis)**.**
Supposons U convexe. *Supposons de plus que f est de classe $\mathcal{C}^1$. Alors, pour tous $a, b \in U$,*

$$\|f(b) - f(a)\| \leq \max_{c \in [a,b]} \|D_c f\| \cdot \left\| \vec{ab} \right\|.$$

Ici, les normes sont toutes les normes euclidiennes standards [a].

[a] La norme appliquée à la matrice $D_c f$ peut être remplacée par la norme subordonnée aux normes euclidiennes standards de $\mathbb{R}^p$ et $\mathbb{R}^q$.

Démonstration.
Choisissons $\gamma(t) := a + t\vec{ab}$. Alors $\gamma(0) = a$ et $\gamma(1) = b$, donc

$$f(b) - f(a) = \int_0^1 (f \circ \gamma)'(t)\, \mathrm{d}t = \int_0^1 D_(\gamma(t)) f \cdot \vec{ab}\, \mathrm{d}t.$$

Par conséquent,

$$\|f(b) - f(a)\| \leq \int_0^1 \left\| D_(\gamma(t)) f \cdot \vec{ab} \right\| \,\mathrm{d}t \leq \int_0^1 \left\| D_(\gamma(t)) f \right\| \cdot \left\| \vec{ab} \right\| \,\mathrm{d}t \leq \max_{c \in [a,b]} \|D_c f\| \cdot \left\| \vec{ab} \right\|. \quad \square$$

Corollaire 8.1.
Si U est convexe, f est de classe $\mathcal{C}^1$ et $a \mapsto D_a f$ est bornée, alors f est lipschitzienne, de semi-norme lipschitzienne au plus $\sup_{a \in U} \|D_a f\|$.

Corollaire 8.2. *Si f est de classe $\mathcal{C}^1$, alors f est localement lipschitzienne.*

9 Références

Dans les références, ce contenu est parfois traité trop vite sur le chemin de théorèmes plus difficiles (inversion locale, fonctions implicites...). Je conseille par exemple le livre de **Skandalis**, ou **[AC] :** ***Topologie et analyse.*** **Auliac, Caby**, même si d'autres références peuvent convenir.

Un des obstacles à cette leçon est le manque d'uniformité des notations dans les références. Choisissez une référence, et utilisez les notations de cette référence.