Module A1
Différentiabilité sur un espace de Hilbert
Différentiabilité de Gateaux

Dans ce cours, sauf mention contraire, $\mathcal{X}$ et $\mathcal{Y}$ désignent des espaces euclidiens, c'est-à-dire des espaces vectoriels de dimension finie. Rappelons que l'on peut munir tout espace euclidien E de dimension $n \in \mathbb{N}^*$ du produit scalaire (bilinéaire et continu) suivant

$$\langle \cdot, \cdot \rangle : \left\{ \begin{array}{ccc} E \times E & \to & \mathbb{R} \\ (x, z) & \mapsto & \displaystyle\sum_{i=1}^{n} x_i \, z_i \end{array} \right.$$

pour tous éléments (x, z) de E tels que $x = x_1 \, e_1 + \cdots + x_n \, e_n$ et $z = z_1 \, e_1 + \cdots + z_n \, e_n$, avec $(e_i)_{1 \leq i \leq n}$ une base de E. On peut lui associer la norme

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2}$$

Lorsque $E = \mathbb{R}^n$ et que $(e_i)_{1 \leq i \leq n}$ est la base canonique de E, le produit scalaire est qualifié d'*usuel* ou de *canonique* et la norme associée, que l'on notera $\| \cdot \|_2$, est la norme euclidienne.

L'objectif de ce module est de rappeler la définition et les principales propriétés de la différentielle usuelle (dite de Fréchet), puis de considérer le cas particulier des fonctions définies sur un ouvert d'un espace euclidien. En ouverture, la troisième partie s'intéressera à d'autres notions de différentiabilité, en particulier celle dite de Gateaux. Tout au long de ce module, on gardera comme fil rouge la généralisation de la dérivée d'une fonction réelle dans un espace de dimension quelconque.

1 Différentielle de Fréchet sur un espace de Banach

1.1 Différentiabilité d'une application

Soit $U \subset \mathbb{R}$ un ouvert et $f : U \to \mathbb{R}$ une **fonction réelle**. On rappelle que f est *dérivable* en $a \in U$ si le taux de variation suivant

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

admet une limite finie, notée $f'(a)$, lorsque h tend vers 0. Autrement dit,

$$\forall \, \varepsilon > 0, \exists \, \delta > 0, \qquad 0 < |h| < \delta \quad \Longrightarrow \quad \left| f'(a) - \frac{f(a+h) - f(a)}{h} \right| < \varepsilon$$

Qu'en est-il lorsque l'on souhaite généraliser cette notion au cas d'*applications* plus générales, définies sur des espaces de dimension quelconque ou dont l'image appartient à un espace de dimension différente de 1 ? Une première généralisation simple est possible pour les applications à variable **réelle** mais dont l'image appartient à un espace de Banach $\mathcal{Y}$ quelconque : il suffit de remplacer dans la formule précédente la valeur absolue par une norme sur $\mathcal{Y}$. On peut ainsi définir de la même manière la dérivée en a

d'une application $f : U \to \mathcal{Y}$, avec U un ouvert réel et $\mathcal{Y}$ un espace vectoriel normé : il s'agit d'un **vecteur** $f'(a) \in \mathcal{Y}$ tel que

$$\forall\, \varepsilon > 0, \exists\, \delta > 0, \qquad 0 < |h| < \delta \quad \Longrightarrow \quad \left\| f'(a) - \frac{f(a+h) - f(a)}{h} \right\| < \varepsilon$$

En effet, puisque h est un **nombre réel**, la division reste bien définie tant que h ne s'annule pas.

Si l'on souhaite généraliser cette définition au cas d'une application **définie** sur un ouvert d'un espace vectoriel normé quelconque (donc différent de $\mathbb{R}$), c'est-à-dire introduire la notion de *différentielle*, on se heurte à une première difficulté : la division par un vecteur n'est pas bien définie. Aussi, l'idée est de réécrire de manière **équivalente** la définition de la dérivée de sorte de ne pas faire apparaître de division par un vecteur. Il y a de nombreuses manières de procéder, et qui ne définissent pas nécessairement les mêmes objets. Dans ce cours, on va étudier quelques unes de ces possibilités, qui vont définir des différentielles différentes.

On va commencer par la définition la plus courante, car c'est celle qui préserve le mieux les propriétés de la dérivation : la différentielle de Fréchet. On revient pour cela à la définition de la dérivée d'une fonction $f : U \to \mathbb{R}$ dérivable en $a \in U$. Soit $V \subset \mathbb{R}$ un voisinage de 0 tel que $a + V \subset U$. On considère alors $\varepsilon : V \to \mathbb{R}$ la fonction définie par

$$\forall\, h \in V, \qquad \varepsilon(h) = \begin{cases} \dfrac{f(a+h) - f(a)}{h} - f'(a) & \text{si } h > 0 \\ -\dfrac{f(a+h) - f(a)}{h} + f'(a) & \text{si } h < 0 \\ 0 & \text{sinon} \end{cases}$$

D'après ce qui précède, si f est dérivable en a, alors ε est continue en 0 et on a

$$\lim_{h \to 0} \varepsilon(h) = 0$$

Par ailleurs, par construction, on a

$$\forall\, h \in V, \qquad f(a+h) = f(a) + f'(a)\, h + \underbrace{|h|\, \varepsilon(h)}_{=o(|h|)}$$

avec $h \mapsto f'(a)\, h$ une fonction linéaire (donc continue). Cette écriture est équivalente à la définition de la dérivée de f en a, mais elle n'utilise plus de division : on peut donc la généraliser telle quelle à n'importe quelle application.

Définition 1 (Différentielle de Fréchet)

Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces de Banach munis chacun d'une norme notée $\|\cdot\|$. Soient $\mathcal{O} \subset \mathcal{X}$ un ouvert et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. On dit que f est *différentiable (au sens de Fréchet)* ou encore *Fréchet-différentiable* en a s'il existe une application linéaire continue $L : \mathcal{X} \to \mathcal{Y}$, un voisinage $\mathcal{V} \subset \mathcal{X}$ de 0 et une application $\varepsilon : \mathcal{V} \to \mathcal{Y}$ tels que $\varepsilon(h)$ tend vers 0 lorsque $\|h\|$ tend vers 0 et

$$\forall\, h \in \mathcal{V}, \qquad f(a+h) = f(a) + L(h) + \|h\|\, \varepsilon(h)$$

Cette application L (dont on montrera qu'elle est unique) est alors appelée *différentielle (de Fréchet)* de f en a.

Si f est différentiable en tout point de Ω, alors on dit que f est *différentiable sur* Ω. Si $f : \mathcal{O} \to \mathcal{Y}$ est différentiable en tout point de son domaine de définition $\mathcal{O}$, alors on dit que f est *différentiable*.

> En général, il est inutile de préciser “au sens de FRÉCHET” et, dans la littérature, à moins que le contexte ne suggère le contraire, la différentiabilité s'entend au sens de FRÉCHET.

> La définition de la différentielle au sens de FRÉCHET dépend des normes dont on munit les espace $\mathcal{X}$ et $\mathcal{Y}$. Ainsi, la différentiabilité d'une application dépend de ces choix. En revanche, en dimension finie, toutes les normes sont équivalentes : la différentiabilité ne dépend donc pas du choix de ces normes.

REMARQUE : Dans la définition 1, le terme

$$\eta(h) = \|h\| \, \varepsilon(h)$$

où $\varepsilon(h) \to 0$ lorsque $\|h\| \to 0$ vérifie la propriété suivante :

$$\frac{\eta(h)}{\|h\|} = \varepsilon(h) \underset{\|h\|\to 0}{\longrightarrow} 0$$

Aussi, on utilisera la notation de LANDAU

$$\|h\| \, \varepsilon(\|h\|) = o(\|h\|)$$

lorsque cela sera nécessaire. Toutefois, la notation “explicite” avec ε, bien que plus lourde, est à privilégier en cas de manipulations. En cas de doute sur les opérations possibles ou non sur les $o(\|h\|)$, il est vivement conseillé d'y revenir.

Dans le cas d'une fonction réelle, la fonction linéaire $h \mapsto f'(a)\, h$ est continue (car les applications linéaires sont continues en dimension finie). Or, en dimension infinie, il existe des applications linéaires non continues, c'est-à-dire non continues en 0. Se pose donc la question de savoir si, dans la définition de la différentielle, on écarte ou non le cas des applications L non continues. La raison pour laquelle on a retenu pour la différentielle de FRÉCHET l'hypothèse selon laquelle L est continue est la suivante : si L est continue, alors puisque

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

on peut passer à la limite et obtenir

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

(on rappelle que $L(0) = 0$). Autrement dit, si f est différentiable au sens de FRÉCHET en a, alors f est continue en a (proposition 1). Ce ne serait pas le cas si L n'est pas continue en 0. Or, une propriété de la dérivabilité qu'il est utile de préserver dans de nombreuses situations est le fait que la dérivabilité implique la continuité. Il s'agit donc d'un choix délibéré, mais pas obligatoire : si on ne retient pas cette hypothèse, alors on définit une autre forme de différentiabilité, qui est satisfaite par un autre ensemble de fonctions, et qui possède d'autres propriétés. Dans la section 3 de ce module, on verra d'autres définitions de différentiabilité qui ne préserve pas la relation “différentiable implique continue”.

Proposition 1

Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces de BANACH. Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. Si f est différentiable en a, alors f est continue en a.

REMARQUE : Cette propriété peut être utilisée pour démontrer qu'une fonction n'est pas différentiable en un point, en établissant qu'elle n'y est pas continue.

1.2 Application différentielle et différentielle en un point

Proposition 2 (Unicité de la différentielle)

Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces de BANACH. Soient $\mathcal{O} \subset \mathcal{X}$ un ouvert et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. On suppose que f est différentiable en a. Alors sa différentielle en a est unique.

DÉMONSTRATION : Laissée en exercice. On raisonne par l'absurde en supposant qu'il existe deux différentielles L_1 et L_2.

L'unicité de la différentielle est une propriété essentielle, qui permet entre autres sa détermination ; il suffit en effet de décomposer la fonction $h \mapsto f(a+h)$ en

$$\underbrace{f(a)}_{\text{terme constant}} \quad + \quad \underbrace{L(h)}_{\text{linéaire continue}} \quad + \underbrace{\|h\|\,\varepsilon(\|h\|)}_{o(\|h\|)}$$

pour obtenir la différentielle.

REMARQUE : En dimension **finie**, il suffit de vérifier la linéarité du terme central, car sa continuité est automatiquement acquise.

EXEMPLE

Différentielle d'une application linéaire continue. Soit $A : \mathcal{X} \to \mathcal{Y}$ une application linéaire continue. Soit $a \in \mathcal{X}$. Pour tout $h \in \mathcal{X}$, on a par linéarité de A

$$A(a+h) = A(a) + A(h) + 0$$

avec $L = A$ une application linéaire continue et $0 = o(\|h\|)$. Ainsi, A est différentiable en a, de différentielle A.

EXERCICE

Soit $\mathcal{X}$ un espace de HILBERT, muni de la norme euclidienne (issue du produit scalaire) $\|\cdot\|$. Montrer que la différentielle de $f = \|\cdot\|^2/2$ est l'identité.

L'unicité de la différentielle permet également d'introduire une notation pour l'application L : dans ce cours, on la notera $d_a f$ ou $df(a)$ (et on réservera la notation $f'(a)$ pour le cas d'une application définie sur un ouvert de $\mathbb{R}$). Par définition, on a donc

$$\forall\, h \in \mathcal{V}, \qquad f(a+h) = f(a) + d_a f(h) + \|h\|\,\varepsilon(h)$$

ou $$\forall\, h \in \mathcal{V}, \qquad f(a+h) = f(a) + df(a)\cdot h + \|h\|\,\varepsilon(h)$$

avec $\varepsilon(h)$ qui tend vers 0 lorsque $\|h\|$ tend vers 0 et $d_a f$ ou $df(a)$ une application linéaire continue. On utilise la notation $df(a)\cdot h$ pour remplacer la notation $df(a)(h)$, d'une part pour alléger les notations (en évitant la succession des parenthèses), mais également pour mettre l'accent sur la linéarité de $df(a)$.

> Les deux formules précédentes sont également connues sous le nom de *formule de Taylor–Young d'ordre 1* ou encore *développement limité* (DL) de f à l'ordre 1.

La notation $d_a f$ est plus adaptée lorsque l'on s'intéresse à l'application linéaire continue L qui apparaît dans la définition 1 pour un point a donné ; on a

$$d_a f : \left\{ \begin{array}{ccc} \mathcal{X} & \to & \mathcal{Y} \\ h & \mapsto & L(h) \end{array} \right.$$

c'est-à-dire que $d_a f$ est un élément de $\mathcal{L}_c(\mathcal{X};\mathcal{Y})$. La notation $df(a)$ est quant à elle plus adaptée lorsque l'on s'intéresse à l'application différentielle

$$df : \left\{ \begin{array}{ccc} \mathcal{O}' & \to & \mathcal{L}_c(\mathcal{X},\mathcal{Y}) \\ a & \mapsto & d_a f \end{array} \right.$$

qui à tout point $a \in \mathcal{O}'$ où f est différentiable, associe l'application linéaire continue $d_a f$. Cette distinction entre les deux objets "différentielle en un point" et "application différentielle" prendra son importance dans les modules **A2 : Différentiabilité sur les espaces euclidiens** et **A4 : Théorème du rang constant**, et peut être rapprochée de la différence qu'il existe entre la (fonction) dérivée f' et le nombre dérivée $f'(a)$ d'une fonction dérivable.

1.3 Règles de calcul

Dans ce paragraphe, on rappelle brièvement les règles du calcul différentiel. Pour plus d'informations sur la différentielle de Fréchet (et en particulier les preuves des résultats qui suivent), le lecteur est invité à consulter le polycopié du cours **3MA260 – Topologie et calcul différentiel II**.

Proposition 3 (Combinaison linéaire et composition d'applications différentiables)

Soient $\mathcal{X}$, $\mathcal{Y}$ et $\mathcal{Z}$ trois espaces de Banach. Soient $\mathcal{O} \subset \mathcal{X}$ et $\mathcal{O}' \subset \mathcal{Y}$ deux ouverts et $a \in \mathcal{O}$. Soient $f, g : \mathcal{O} \to \mathcal{Y}$ deux applications différentiables en a telles que $f(a) \in \mathcal{O}'$ et $j : \mathcal{O}' \to \mathcal{Z}$ une application différentiable en $f(a)$. Alors

- (multiplication par un scalaire) l'application $(\lambda\, f) : \mathcal{O} \to \mathcal{Y}$ est différentiable en a pour tout $\lambda \in \mathbb{R}$ et
$$\forall\, h \in \mathcal{X}, \qquad d(\lambda\, f)(a)\cdot h = \lambda\, df(a)\cdot h$$
- (somme) l'application $(f+g) : \mathcal{O} \to \mathcal{Y}$ est différentiable en a et
$$\forall\, h \in \mathcal{X}, \qquad d(f+g)(a)\cdot h = df(a)\cdot h + dg(a)\cdot h$$
- (composition) l'application $(j \circ f) : \mathcal{O} \to \mathcal{Z}$ est différentiable en a et
$$\forall\, h \in \mathcal{X}, \qquad d(j\circ f)(a)\cdot h = dj(f(a))\cdot (df(a)\cdot h)$$

En général, les applications que l'on étudie sont construites à l'aide de fonctions usuelles dérivables, mais de formes plus complexes qu'une simple somme ou composition.

Proposition 4 (Composantes d'une application)

Soient $\mathcal{X}$ et $(\mathcal{Y}_i)_{1\leq i\leq m}$ des espaces de Banach. Soient $\mathcal{O} \subset \mathcal{X}$ un ouvert et $a \in \mathcal{O}$. Soient $f_i : \mathcal{O} \to \mathcal{Y}_i$ des applications pour tout $i \in [\![1\,;m]\!]$. Alors l'application

$$f : \left\{ \begin{array}{ccc} \mathcal{O} & \to & \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_m \\ x & \mapsto & \big(f_1(x), \ldots, f_m(x)\big) \end{array} \right.$$

est différentiable en a si et seulement si les f_i sont différentiables en a, et on a dans ce cas :

$$\forall\, h \in \mathcal{X}, \qquad df(a) \cdot h = (df_1(a) \cdot h, \ldots, df_m(a) \cdot h)$$

Exercice

On considère l'application définie par

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (\exp(t), t^2) \end{array} \right.$$

Montrer que f est différentiable sur $\mathbb{R}^2$ et calculer sa différentielle.

Proposition 5 (Somme et produit séparables de fonctions différentiables)

Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces de Banach. Soient $\mathcal{O} \subset \mathcal{X}$ et $\mathcal{O}' \subset \mathcal{Y}$ deux ouverts et $(a,b) \in \mathcal{O} \times \mathcal{O}'$. Soient $f : \mathcal{O} \to \mathbb{R}$ une fonction différentiable en a et $g : \mathcal{O}' \to \mathbb{R}$ une fonction différentiable en b. Alors

- (somme séparable) la fonction

$$s : \left\{ \begin{array}{ccc} \mathcal{O} \times \mathcal{O}' & \to & \mathbb{R} \\ (x,z) & \mapsto & f(x) + g(z) \end{array} \right.$$

  est différentiable en (a,b) et

$$\forall\,(h,k) \in \mathcal{X} \times \mathcal{Y}, \qquad ds(a,b) \cdot (h,k) = df(a) \cdot h + dg(b) \cdot k$$

- (produit séparable) la fonction

$$s : \left\{ \begin{array}{ccc} \mathcal{O} \times \mathcal{O} & \to & \mathbb{R} \\ (x,z) & \mapsto & f(x)\, g(z) \end{array} \right.$$

  est différentiable en (a,b) et

$$\forall\,(h,k) \in \mathcal{X} \times \mathcal{Y}, \qquad dp(a,b) \cdot (h,k) = g(b)\, df(a) \cdot h + f(a)\, dg(b) \cdot k$$

Démonstration : Par hypothèse sur f et g, on a

$$f(a+h) = f(a) + df(a) \cdot h + \|h\|\, \varepsilon_1(h) \text{ et } g(b+k) = g(b) + dg(b) \cdot k + \|k\|\, \varepsilon_2(k)$$

avec $\varepsilon_i(t) \to 0$ lorsque $t \to 0$ pour $i \in \{1, 2\}$.

- (somme séparable) Par définition de s,

$$\begin{aligned} s(a+h,b+k) &= f(a+h)+g(b+k) \\ &= \underbrace{f(a)+g(b)}_{=s(a,b)} + \underbrace{df(a)\cdot h + dg(b)\cdot k}_{\substack{\text{linéaire et continue} \\ \text{en } (h,k)}} + \|h\|\,\varepsilon_1(h) + \|k\|\,\varepsilon_2(k) \end{aligned}$$

En particulier, puisque

$$\|h\| \leq \sqrt{\|h\|^2+\|k\|^2} \qquad \text{et} \qquad \|k\| \leq \sqrt{\|h\|^2+\|k\|^2}$$

l'inégalité triangulaire permet d'écrire

$$\big|\|h\|\,\varepsilon_1(h)+\|k\|\,\varepsilon_2(k)\big| \leq \|(h,k)\| \ \underbrace{\big(|\varepsilon_1(h)|+|\varepsilon_2(k)|\big)}_{\to 0 \text{ si } \|(h,k)\|\to 0}$$

Il s'ensuit que s est différentiable en (a,b) et on obtient l'expression de sa différentielle comme annoncé.

- (produit séparable) Par définition de p,

$$\begin{aligned} p(a+h,b+k) &= f(a+h)\,g(b+k) \\ &= \underbrace{f(a)\,g(b)}_{=p(a,b)} + \underbrace{g(b)\,df(a)\cdot h + f(a)\,dg(b)\cdot k}_{\substack{\text{linéaire et continue} \\ \text{en } (h,k)}} + \underbrace{o(\|h\|)+o(\|k\|)}_{=o(\|(h,k)\|)} \end{aligned}$$

et on conclut avec les mêmes arguments que pour la somme séparable. ■

Exemple

Fonctions polynomiales. Les fonctions polynomiales sur $\mathbb{R}^m$, c'est-à-dire de la forme

$$P : (x_1,\ldots,x_m) \mapsto \sum_{j=1}^{J} a_j \prod_{i=1}^{m} x_i^{d_{i,j}} \qquad\qquad \text{avec } d_{i,j}\in\mathbb{N}$$

sont différentiables. En effet, chaque monome définit une fonction $t\mapsto t^{d_{i,j}}$ dérivable sur $\mathbb{R}$, donc la fonction

$$(x_1,\ldots,x_m) \mapsto \prod_{i=1}^{m} x_i^{d_{i,j}}$$

est différentiable sur $\mathbb{R}^m$ pour tout $j\in[\![1\,;J]\!]$ (comme produit séparable de fonctions dérivables). La fonction P étant combinaison linéaire de telles fonctions, la proposition 3 assure qu'elle est différentiable sur $\mathbb{R}^m$.

> L'exemple précédent montre en particulier que le produit scalaire dans $\mathbb{R}^n$, le déterminant d'une matrice et la trace d'une matrice sont des fonctions différentiables.

2 Fonction différentiable sur un espace de Hilbert

Dans cette section, on s'intéresse au cas particulier de la différentiabilité des **fonctions** définies sur $\mathcal{X}$ (c'est-à-dire à valeurs dans $\mathbb{R}$).

> Dans ce qui suit et dans les modules suivants, la notation $\mathcal{X}$ (ainsi que $\mathcal{Y}$, $\mathcal{Z}$...) est utilisée lorsque les résultats associés restent valables **sans hypothèses supplémentaires** lorsque $\mathcal{X}$ (et $\mathcal{Y}$ etc.) sont des **espaces de Hilbert**. Ces espaces, qui généralisent les espaces euclidiens en dimension infinie, possèdent des propriétés communes avec ces derniers, notamment l'existence d'un produit scalaire bilinéaire et continu (noté $\langle \cdot, \cdot \rangle$) et d'une norme associée (notée $\| \cdot \|$). **Dans tous les énoncés de ce cours**, on pourra donc remplacer l'espace euclidien $\mathcal{X}$ par un espace de Hilbert $\mathcal{X}$. En revanche, la notation E (et F...) sera réservée aux espaces euclidiens uniquement.

2.1 Gradient d'une fonction différentiable

Lorsque $f : \mathcal{O} \to \mathbb{R}$ est différentiable en $a \in \mathcal{O}$, sa différentielle en a est l'application $d_a f : \mathcal{X} \to \mathbb{R}$, qui par définition est linéaire et continue. Si $\mathcal{X} = \mathbb{R}^n$, alors il s'agit d'une forme linéaire (et continue) de $\mathbb{R}^n$ dans $\mathbb{R}$. Il existe donc une unique matrice $M \in \mathcal{M}_{1,n}(\mathbb{R})$ telle que

$$\forall\, x \in \mathbb{R}^n, \qquad d_a f(x) = M\, x$$

Or, l'écriture explicite du produit matriciel $M\,x$ assure que $M\,x = \langle M^\top, x\rangle$, où $M^\top$ est la transposée de la matrice (vecteur) M. Il s'ensuit donc qu'il existe un unique vecteur $v = M^\top \in \mathbb{R}^n$ tel que

$$\forall\, x \in \mathbb{R}^n, \qquad d_a f(x) = \langle M^\top, x\rangle$$

> Ce résultat est généralisable à n'importe quel espace de Hilbert $\mathcal{X}$: il s'agit du théorème de représentation de Riesz, qui assure que, à toute forme linéaire continue $L : \mathcal{X} \to \mathbb{R}$, on peut associer un unique vecteur $v \in \mathcal{X}$ tel que
> $$\forall\, x \in \mathcal{X}, \qquad L(x) = \langle v, x\rangle$$
> Puisque la différentielle d'une fonction en a est une forme linéaire continue, on peut donc lui appliquer ce théorème.

En notant $\nabla f(a)$ l'unique vecteur $v \in \mathcal{X}$ déterminé plus haut, on définit :

Définition 2 (Gradient d'une fonction différentiable)

Soient $\mathcal{O} \subset \mathcal{X}$ un ouvert et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathbb{R}$ une fonction différentiable en a. On appelle *gradient de f en a* l'unique vecteur $\nabla f(a)$ vérifiant

$$\forall\, h \in \mathcal{X}, \qquad d_a f(h) = \langle \nabla f(a), h\rangle$$

La notion de gradient généralise de manière plus rigoureuse celle de dérivée :

Proposition 6 (Dérivée et gradient)

Soit $U \subset \mathbb{R}$ un ouvert et $a \in U$. Soit $f : U \to \mathbb{R}$ une fonction dérivable en a. Alors $\nabla f(a) = f'(a)$.

On voit donc que le gradient joue un rôle analogue à celui de la dérivée : tous les deux définissent de manière **implicite** la différentielle (le nombre dérivé est le coefficient directeur de la différentielle, tandis que le gradient définit la différentielle à l'aide du **produit scalaire**). On verra dans le paragraphe §2.3 que considérer le gradient à la place de la différentielle permet d'alléger certaines notations. Par ailleurs, l'étude des propriétés de l'application différentielle $df : \mathcal{O} \to \mathcal{L}_c(\mathcal{X}; \mathbb{R})$ peut se faire directement sur l'application gradient $\nabla f : \mathcal{O} \to \mathcal{X}$, dont l'image est généralement plus simple que celle de l'application différentielle, comme on va le voir dans le paragraphe suivant.

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

Commençons par insister sur le fait que, si $d_a f$ est par définition linéaire et continue, df n'est aucune raison de l'être en général. Elle est linéaire pour certaines applications quadratiques, alors qu'elle n'est continue que pour les applications de classe $\mathcal{C}^1$.

Définition 3 (Application $\mathcal{C}^1$)

Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces de Banach. Soient $\mathcal{O} \subset \mathcal{X}$ un ouvert et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. On suppose que f est différentiable au voisinage de a. On dit que f est de classe $\mathcal{C}^1$ en a si l'application df est continue en a.

Pour étudier la continuité éventuelle de df, il faut munir $\mathcal{L}_c(\mathcal{X}, \mathcal{Y})$ d'une norme. On choisit ici la norme d'opérateur :

$$\forall\, L \in \mathcal{L}_c(\mathcal{X}, \mathcal{Y}), \qquad |||L||| = \sup_{\|z\|=1} \|L(z)\|$$

On remarque que cette norme dépend du choix des normes sur les espaces $\mathcal{X}$ et $\mathcal{Y}$. On dit alors que df est continue en $a \in \mathcal{O}$ si

$$\forall\, \varepsilon > 0, \exists\, \eta > 0, \forall\, x \in \mathcal{X}, \qquad \|x - a\| \leq \eta \quad \Longrightarrow \quad |||df(x) - df(a)||| \leq \varepsilon$$

On voit donc que la continuité de la différentielle est en général difficile à établir, car la norme d'opérateur est généralement difficile à manipuler. Dans le cas d'une fonction définie sur un espace de Hilbert $\mathcal{X}$, la continuité de l'application différentielle df peut être caractérisée de manière plus simple, à l'aide du gradient, de la même manière que la continuité de la différentielle $a \mapsto (h \mapsto f'(a)\, h))$ d'une fonction réelle f dérivable est équivalente à la continuité de la fonction dérivée $a \mapsto f'(a)$.

Proposition 7

Soient $\mathcal{O} \subset \mathcal{X}$ un ouvert et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathbb{R}$ une fonction différentiable au voisinage de a. Alors on a équivalence entre les deux énoncés suivants :

(i) f est de classe $\mathcal{C}^1$ en a ;

(ii) ∇f est continue en a.

Démonstration : Puisque $\mathcal{X}$ est un espace de Hilbert et que f est différentiable au voisinage de a, on a pour tout x voisin de a

$$\forall\, h \in \mathcal{X}, \qquad df(x) \cdot h = \langle \nabla f(x), h \rangle$$

Ainsi, on a $$|||df(x) - df(a)||| = \sup_{\|h\|=1} |\langle \nabla f(x), h\rangle - \langle \nabla f(a), h\rangle|$$

La bilinéarité du produit scalaire assure que

$$|\langle \nabla f(x), h\rangle - \langle \nabla f(a), h\rangle| = |\langle \nabla f(x) - \nabla f(a), h\rangle|$$

tandis que l'inégalité de Cauchy–Schwarz permet d'obtenir la majoration suivante pour tout $h \in \mathcal{X}$ de norme 1 :

$$|\langle \nabla f(x) - \nabla f(a), h\rangle| \leq \|\nabla f(x) - \nabla f(a)\| \cdot \|h\| = \|\nabla f(x) - \nabla f(a)\|$$

Ainsi, on en déduit que si ∇f est continue en a, alors df est continue en a. Réciproquement, si df est continue en a, alors puisque pour tout $h \in \mathcal{X}$ de norme 1

$$|\langle \nabla f(x) - \nabla f(a), h\rangle| \leq |||df(x) - df(a)|||$$

on conclut en appliquant cette inégalité à $h = \dfrac{\nabla f(x) - \nabla f(a)}{\|\nabla f(x) - \nabla f(a)\|}$. ■

Dans le module **A2 : Différentiabilité sur les espaces euclidiens**, on pourra même étudier la différentiabilité de l'application gradient.

2.3 Règles de calcul

Les résultats établis dans le paragraphe §1.3 pour les fonctions différentiables sur un ouvert d'un espace de Hilbert peuvent s'exprimer à l'aide des gradients :

Proposition 8 (Combinaison linéaire et composition de fonctions différentiables)

Soient $\mathcal{O} \subset \mathcal{X}$ et $U \subset \mathbb{R}$ deux ouverts, $a \in \mathcal{O}$. Soient $f, g : \mathcal{O} \to \mathbb{R}$ deux fonctions différentiables en a telles que $f(a) \in U$ et $j : U \to \mathbb{R}$ une fonction dérivable en $f(a)$. Alors on a

- (homogénéité) $\nabla(\lambda\, f)(a) = \lambda\, \nabla f(a)$
- (somme) $\nabla(f+g)(a) = \nabla f(a) + \nabla g(a)$
- (produit) $\nabla(f \times g)(a) = g(a)\, \nabla f(a) + f(a)\, \nabla g(a)$
- (composition) $\nabla(j \circ f)(a) = j'(f(a)) \nabla f(a)$

Démonstration : Soit $h \in \mathcal{X}$. L'idée est d'utiliser les identités établies dans la proposition 8 pour écrire l'image de la différentielle sous la forme d'un produit scalaire. L'unicité du gradient permet alors de conclure.

- (homogénéité) La bilinéarité du produit scalaire permet d'écrire
$$\lambda\, df(a) \cdot h = \lambda\, \langle \nabla f(a), h\rangle = \langle \lambda\, \nabla f(a), h\rangle$$
- (somme) La bilinéarité du produit scalaire donne cette fois
$$df(a) \cdot h + dg(a) \cdot h = \langle \nabla f(a), h\rangle + \langle \nabla g(a), h\rangle = \langle \nabla f(a) + \nabla g(a), h\rangle$$
- (composition) En remplaçant successivement les différentielles par le gradient et la dérivée associés, on obtient
$$dj(f(a)) \cdot (df(a) \cdot h) = dj(f(a)) \cdot (\langle \nabla f(a), h\rangle) = j'(f(a)) \cdot \langle \nabla f(a), h\rangle$$
et on conclut en utilisant à nouveau la bilinéarité du produit scalaire. ■

3 Différentiabilité au sens de Gateaux [1]

L'objectif de cette section est d'explorer d'autres notions de différentiabilité, c'est-à-dire d'autres généralisations de la dérivabilité à des applications à variable vectorielle. On s'intéressera en particulier à la différentiabilité de Gateaux, qui possède des propriétés intéressantes, notamment dans le cas convexe.

3.1 Dérivée directionnelle suivant un vecteur

Nous avons vu dans la section 1 que la différentielle (au sens de Fréchet) est **une** généralisation de la dérivée dans le cas des applications définies sur des espaces de Banach quelconque, obtenue en réécrivant de manière équivalente la définition de la dérivée sous la forme d'un développement limité d'ordre 1, afin de supprimer la division qui apparaît dans le taux de variation.

Si l'on souhaite malgré tout conserver une définition avec un taux de variation, on peut commencer par faire la remarque suivante : si $v \in \mathbb{R}$ est non nul et que $f : U \to \mathbb{R}$ (avec $U \subset \mathbb{R}$ un ouvert) est dérivable en $a \in U$, alors en effectuant le changement $h = t\,v$, on obtient

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{t \to 0} \frac{f(a+t\,v) - f(a)}{t\,v} = \frac{1}{v} \lim_{t \to 0} \frac{f(a+t\,v) - f(a)}{t}$$

de sorte que f est dérivable en a si et seulement s'il existe $f'(a) \in \mathbb{R}$ tel qu'on a la limite suivante pour tout $v \neq 0$:

$$\lim_{t \to 0} \frac{f(a+t\,v) - f(a)}{t} = f'(a)\,v$$

Notons cette relation reste valable si $v = 0$ (la limite étant alors nulle).

L'intérêt de cette réécriture est de garder une expression proche du taux de variation, tout en ouvrant la possibilité de la généraliser à des applications non définies sur un ensemble réel. En effet, si $f : \mathcal{O} \to \mathcal{Y}$ avec $\mathcal{O} \subset \mathcal{X}$ un ouvert, $a \in \mathcal{O}$ et $v \in \mathcal{X}$ un **vecteur**, on peut s'intéresser à la limite de la quantité

$$\frac{f(a+t\,v) - f(a)}{t}$$

lorsque t tend vers 0, qui cette fois est bien définie (puisqu'on divise par un **réel** non nul). Il s'agit en réalité du taux de variation appliqué à l'application $g : t \mapsto f(a+t\,v)$ en 0 (puisque $g(0) = f(a)$), qui est définie sur un ensemble **réel**. Ainsi, cette quantité tend vers une limite finie lorsque $t \to 0$ si et seulement si g est dérivable en 0.

On peut donc introduire la définition suivante :

Définition 4 (Dérivée directionnelle suivant un vecteur)

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$, $a \in \mathcal{O}$ et $v \in \mathcal{X}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. Soit $U \subset \mathbb{R}$ un ouvert contenant 0 tel que l'application

$$\left\{ \begin{array}{ccc} U & \to & \mathcal{Y} \\ t & \mapsto & f(a+t\,v) \end{array} \right.$$

soit bien définie. On dit que f admet une *dérivée directionnelle suivant v au point a* si cette application est dérivable en 0, de dérivée notée $f'(a;v)$.

1. Certains auteurs écrivent Gâteaux, mais René-Eugène Gateaux (à qui le nom de cette notion rend honneur) signait Gateaux sans l'accent circonflexe, et c'est également ainsi que son nom est orthographié sur son acte de naissance.

Ainsi, f admet une dérivée directionnelle suivant v au point a si la limite suivante existe

$$\lim_{t\to 0} \frac{f(a+t\,v)-f(a)}{t} = f'(a;v) \in \mathcal{Y}$$

Notons que toute application admet une dérivée directionnelle suivant le vecteur nul en tout point où elle est définie, donc la question n'est pertinente qu'en v non nul.

Si $f : U \to \mathcal{Y}$ est une application d'une variable réelle dérivable en $a \in U$, alors elle est dérivable directionnellement suivant tout vecteur, et

$$\forall\, v \in \mathbb{R}, \qquad f'(a;v) = v\, f'(a) \in \mathcal{Y}$$

Dans le cas dérivable (et en particulier réel), il existe donc un lien simple entre la dérivée et les dérivées directionnelles. On pourrait donc être tenté de remplacer la différentiabilité au sens de Fréchet par la dérivabilité directionnelle suivant tout vecteur. Malheureusement, l'exemple suivant montre que la dérivabilité directionnelle est une notion trop faible, et qu'elle ne préserve en particulier pas la propriété “dérivable implique continue”.

Contre-exemple

Fonction non continue et dérivable directionnellement suivant tout vecteur. On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$\forall\,(x,y) \in \mathbb{R}^2, \qquad f(x,y) = \begin{cases} \dfrac{y^2}{x} & \text{si } x \neq 0 \\ y & \text{sinon} \end{cases}$$

Cette fonction n'est pas continue en $(0,0)$ car

$$\lim_{t\to 0} f(t^4,t) = +\infty \qquad \text{et} \qquad \lim_{t\to 0} f(t,t) = 0$$

En revanche, pour tout $(u,v) \in \mathbb{R}^2 \setminus \{(0,0)\}$, on a

$$\forall\, t \neq 0, \qquad f(0+t\,u, 0+t\,v) = \frac{t^2\,v^2}{t\,u} = t\,\frac{v^2}{u}$$

Ainsi, la fonction $t \mapsto f(t\,u, t\,v)$ est linéaire, donc dérivable en 0, de dérivée $\dfrac{v^2}{u}$.

Bien que la dérivabilité directionnelle ne soit pas une généralisation raisonnable de la dérivabilité en dimension supérieure, elle présente un intérêt non négligeable, qui est celui d'avoir une forme proche de celle de la dérivabilité en dimension 1 (puisqu'elle est définie à l'aide d'un taux de variation), et permet donc de généraliser plus facilement les résultats de dérivabilité (et leurs preuves) qui s'écrivent à l'aide d'un taux de variation. Elle repose aussi sur la définition de la dérivée d'une fonction, ce qui signifie que les résultats sur la dérivée (en particulier, les règles de calculs) peuvent s'y appliquer.

> Dans le module **A2 : Différentiabilité sur les espaces euclidiens**, on verra toutefois que, dans le cas des fonctions différentiables au sens de Fréchet, les notions de différentielle et de dérivée directionnelle sont étroitement liées.

3.2 Différentielle de Gateaux

Les dérivées directionnelles présentent l'avantage d'être simples à manipuler car elles ramènent le problème de la différentiabilité à des problèmes de dérivabilité. Cependant, leur simplicité en font une notion relativement faible, qui échoue à satisfaire certaines propriétés. On va donc rajouter des hypothèses pour renforcer la notion qu'elle définit.

Définition 5 (Différentielle de GATEAUX)

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. On dit que f est *différentiable au sens de* GATEAUX ou encore GATEAUX*-différentiable* au point a si elle admet une dérivée directionnelle suivant tout $v \in \mathcal{X}$ en a et que l'application

$$\left\{ \begin{array}{ccc} \mathcal{X} & \to & \mathcal{Y} \\ v & \mapsto & f'(a;v) \end{array} \right.$$

est une application **linéaire continue**. Dans ce cas, cette application (dont on montre qu'elle est unique) est notée $d_{\mathrm{G}}f(a)$ et est appelée *différentielle de* GATEAUX de f en a.

On dit alors que f est différentiable au sens de GATEAUX sur $\mathcal{U}$ si f est différentiable au sens de GATEAUX en tout point de $\mathcal{U}$ et que f est différentiable au sens de GATEAUX si elle est est différentiable au sens de GATEAUX sur son domaine de définition $\mathcal{O}$.

> D'autres définitions de la différentielle de GATEAUX peuvent être rencontrées dans la littérature mathématique. Parmi les variantes rencontrées, on peut citer l'hypothèse de linéarité ou de la continuité de $v \mapsto f'(a;v)$ qui ne sont pas toujours imposées, ou encore le remplacement de la dérivée directionnelle par une autre notion de dérivée. Le choix réalisé ici est relativement commun, et va être abondamment utilisé dans les preuves qui suivent, en particulier dans les résultats démontrant l'identité entre la différentielle de GATEAUX et celle de FRÉCHET (qui entraîne la continuité).

EXEMPLE

Applications linéaires. Soit $A : \mathcal{X} \to \mathcal{Y}$ une application linéaire. Alors A est différentiable au sens de GATEAUX si et seulement si A est continue, et on a

$$\forall\, a \in \mathcal{X}, \qquad d_{\mathrm{G}}A(a) = A = dA(a)$$

Proposition 9

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. Alors f est différentiable au sens de GATEAUX en a si et seulement s'il existe $L : \mathcal{X} \to \mathcal{Y}$ une application **linéaire continue** telle que pour tout $v \in \mathcal{X}$,

$$f(a + h\,v) = f(a) + L(h\,v) + o(|h|)$$

DÉMONSTRATION : Laissée en exercice.

On note une analogie entre cette égalité et celle définissant la différentielle de FRÉCHET. La différence principale réside dans le troisième terme, qui **dépend de** v mais ne converge par forcément vers 0 lorsque v tend vers 0. Notons par ailleurs que, dans cette définition, $L = d_{\mathrm{G}}f(a)$.

L'ensemble des applications GATEAUX-différentiables est stable par combinaison linéaire :

Proposition 10

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soient $f : \mathcal{O} \to \mathcal{Y}$ et $g : \mathcal{O} \to \mathcal{Y}$ deux applications différentiables au sens de GATEAUX en a. Alors $\lambda\, f + \mu\, g$ est différentiable au sens de GATEAUX en a quel que soit $(\lambda, \mu) \in \mathbb{R}^2$ et on a

$$d_{\mathrm{G}}(\lambda\, f + \mu\, g)(a) = \lambda\, d_{\mathrm{G}} f(a) + \mu\, d_{\mathrm{G}} g(a)$$

DÉMONSTRATION : Laissée en exercice.

3.3 Lien entre les différentiabilités de FRÉCHET et de GATEAUX

Intéressons-nous à certains cas de coïncidence entre la différentiabilité de FRÉCHET et la différentiabilité de GATEAUX. Le premier cas est celui des applications dérivables :

Proposition 11

Soient U un ouvert de $\mathbb{R}$ et $a \in U$. Soit $f : U \to \mathcal{Y}$ une application. Alors on a l'équivalence entre les énoncés suivants :

(i) f est dérivable au sens de FRÉCHET en a

(ii) f est différentiable au sens de GATEAUX en a

De plus, on a $\qquad \forall\, t \in \mathbb{R}, \qquad d_{\mathrm{G}} f(a)(t) = f'(a)\, t$

En dimension supérieure, la différentiabilité de FRÉCHET entraîne la différentiabilité de GATEAUX (mais pas l'inverse) :

Proposition 12

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. Si f est différentiable au sens de FRÉCHET en a, alors f est différentiable au sens de GATEAUX en a. On a alors

$$d_{\mathrm{G}} f(a) = df(a)$$

REMARQUE : En particulier, la proposition 12 entraîne qu'une application FRÉCHET-différentiable en a admet des dérivées directionnelles suivant tout vecteur en a.

DÉMONSTRATION : Soit $v \in \mathcal{X}$. Soit $U \subset \mathbb{R}$ un ouvert contenant 0 tel que $a + t\, v \in \mathcal{O}$ pour tout $t \in U$. On pose

$$g : \left\{ \begin{array}{ccc} U & \to & \mathcal{Y} \\ h & \mapsto & f(a + h\, v) \end{array} \right.$$

La définition de la différentiabilité au sens de FRÉCHET assure que

$$g(h) = f(a + h\, v) = f(a) + df(a) \cdot (h\, v) + \|h\, v\|\, \varepsilon(h\, v)$$

où $\varepsilon(h) \to 0$ lorsque $h \to 0$. Il s'ensuit que

$$f'(a; v) = g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = df(a) \cdot v$$

On en déduit la linéarité de $v \mapsto f'(a; v)$. ■

REMARQUE : Si f est une fonction (c'est-à-dire si $\mathcal{Y} = \mathbb{R}$) définie sur un ouvert d'un espace de HILBERT, alors on peut vérifier que si f est différentiable en a, alors

$$\forall\, v \in \mathcal{X}, \qquad f'(a;v) = \langle \nabla f(a), v \rangle$$

Attention, la réciproque de la proposition précédente est fausse : il existe des applications différentiables au sens de GATEAUX en a qui ne sont pas différentiables au sens de FRÉCHET, en particulier parce que la différentiabilité au sens de GATEAUX n'implique pas la continuité :

CONTRE-EXEMPLE

On considère la fonction définie sur $\mathbb{R}^2$ par

$$\forall\,(x,y) \in \mathbb{R}^2, \qquad f(x,y) = \begin{cases} 0 & \text{si } (x,y) = (0,0) \\ \dfrac{x^3\, y}{x^6 + y^2} & \text{sinon} \end{cases}$$

Cette fonction n'est pas continue en $(0,0)$, puisque

$$\lim_{t \to 0} f(t,t^3) = \frac{1}{2}$$

Or, si on considère $(u,v) \in \mathbb{R}^2 \setminus \{(0,0)\}$, alors on a

$$\forall\, h \neq 0, \qquad f(h\,u, h\,v) = \frac{h^4\, u^3\, v}{h^6\, u^6 + h^2\, v^2} = h\, \underbrace{\frac{h\, u^3\, v}{h^4\, u^6 + v^3}}_{\underset{h\to 0}{\longrightarrow}\, 0}$$

de sorte que $\qquad f(0 + h\,u, 0 + h\,v) = f(0,0) + 0 + o(|h|)$

Ainsi f est différentiable au sens de GATEAUX en $(0,0)$, mais n'y est pas continue.

Les notions de différentiabilité au sens de FRÉCHET et de GATEAUX coïncident si les applications sont de classe $\mathcal{C}^1$. Avant de démontrer ce résultat, on établit le lemme :

Lemme 1 (Inégalité des accroissements finis)

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. On suppose qu'il existe $(a,b) \in \mathcal{O}^2$ tel que le segment

$$[\,a\,;b\,] = \Big\{ \lambda\, a + (1-\lambda)\, b \mid \lambda \in [\,0\,;1\,] \Big\}$$

soit contenu dans $\mathcal{O}$ et tel que f soit continue sur $[\,a\,;b\,]$ et différentiable au sens de GATEAUX en tout point de $]\,a\,;b\,[= [\,a\,;b\,] \setminus \{a,b\}$. Alors on a

$$\|f(a) - f(b)\| \leq \sup_{\substack{x \in [\,a\,;b\,] \\ \|v\|=1}} \|f'(x;v)\| \cdot \|a-b\|$$

REMARQUE : Cet énoncé ressemble à l'inégalité des accroissements finis dans le cas réel ; cependant, il est nécessaire d'y ajouter l'hypothèse de continuité.

DÉMONSTRATION : On considère l'application

$$g : \left\{ \begin{array}{rcl} [\,0\,;1\,] & \to & \mathcal{Y} \\ t & \mapsto & f(b + t\,(a-b)) \end{array} \right.$$

qui vérifie $g(0) = f(b)$ et $g(1) = f(a)$. Remarquons que

$$b + t\,(a-b) = t\,a + (1-t)\,b \in [\,a\,;b\,]$$

Par composition, l'application g est Gateaux-différentiable sur $]\,0\,;1\,[$, et donc dérivable, avec

$$\forall\, t\ \in\,]\,0\,;1\,[\,,\qquad g'(t) = f'(b + t\,(a-b); a-b)$$

L'inégalité des accroissements finis (pour les applications définies sur un ensemble réel) assure donc que

$$\|g(0) - g(1)\| \leq \sup_{t \in [\,0\,;1\,]} \|g'(t)\|$$

soit
$$\|f(a) - f(b)\| \leq \sup_{t \in [\,0\,;1\,]} \|f'(b + t\,(a-b); a-b)\|$$

Puisque, par linéarité,

$$\begin{aligned} \|f'(b + t\,(a-b); a-b)\| &= \|f'\left(b + t\,(a-b); \frac{a-b}{\|a-b\|}\right)\| \cdot \|a-b\| \\ &\leq \sup_{\|v\|=1} \|f'(b + t\,(a-b); v)\| \cdot \|a-b\| \end{aligned}$$

on termine en considérant la borne supérieure sur les t dans $[\,0\,;1\,]$. ■

Proposition 13

Soient $\mathcal{O}$ un ouvert de $\mathcal{X}$ et $a \in \mathcal{O}$. Soit $f : \mathcal{O} \to \mathcal{Y}$ une application. Alors on a équivalence entre les deux énoncés suivants :

(i) f est une application de classe $\mathcal{C}^1$ en a ;

(ii) f est différentiable au sens de Gateaux sur un voisinage $\mathcal{V}$ de a et $d_{\mathrm{G}} f$ est continue en a.

Démonstration : On démontre séparément les deux sens de l'équivalence :

- **Sens direct.** Si f est de classe $\mathcal{C}^1$, alors elle est en particulier Fréchet-différentiable et on a $d_{\mathrm{G}} f(x) = df(x)$ pour tout $x \in \mathcal{V}$. La continuité de df en a assure celle de $d_{\mathrm{G}} f$ en a.
- **Réciproque.** Soit $x \in \mathcal{V}$ et $h \in \mathcal{X}$ voisin de 0 tel que $x + [\,0\,;1\,]\,h \subset \mathcal{V}$. On pose

$$g : \left\{ \begin{array}{ccc} \mathcal{U} & \to & \mathcal{Y} \\ z & \mapsto & f(z) - f'(x; z) \end{array} \right.$$

Puisque $d_{\mathrm{G}} f(x)(z) = f'(x; z)$ pour tout $z \in \mathcal{X}$ et que f est Gateaux-différentiable, il s'ensuit que $z \mapsto f'(x; z)$ est une application linéaire et continue, donc Gateaux-différentiable. Ainsi, en tant que somme de deux applications différentiables au sens de Gateaux, g est différentiable au sens de Gateaux sur $\mathcal{U}$; on a alors

$$g'(z; v) = f'(z; v) - f'(x; v)$$

et l'inégalité des accroissements finis (lemme 1) assure que

$$\|g(x+h) - g(x)\| \leq \sup_{\substack{z \in [\,x\,;x+h\,] \\ \|v\|=1}} \|f'(z; v) - f'(x; v)\| \cdot \|h\|$$

soit
$$\|f(x+h) - f(x) - f'(x; h)\| \leq \sup_{\substack{z \in [\,x\,;x+h\,] \\ \|v\|=1}} \|f'(z; v) - f'(x; v)\| \cdot \|h\|$$

La continuité de la différentielle de Gateaux assure alors que la borne supérieure tend vers zéro lorsque h tend vers zéro. Autrement dit, f est différentiable au sens de Fréchet en x et $df(x) \cdot h = f'(x; h)$ pour tout $h \in \mathcal{X}$. La continuité de df en a découle de celle de la différentielle de Gateaux en a. ■

En vertu de ce résultat, il n'est pas nécessaire de préciser selon quel sens de différentiabilité on considère la continuité de la différentielle.