Analyse dans R^n, chapitre II

Alexey Muranov

17 avril 2025

Table des matières

II. Calcul différentiel

II.1. Dérivée totale

II.1.1. Dérivée d'une fonction d'un argument réel à valeurs dans R^n

Rappelons-nous que si $A \subset \mathbf{R}$ et $f\colon A \to \mathbf{R}$, alors la *fonction dérivée* de f est la fonction g à valeurs réelles définie par la formule :

$$f(x+h) = f(x) + g(x)h + o(|h|) \qquad \text{lorsque } h \to 0.$$

La fonction dérivée de f est couramment notée « f' » (notation de Lagrange[1]) ou « Df » (notation d'Arbogast[2]). Il y a d'autres notations courantes, surtout dans des contextes particuliers.

On peut facilement généraliser cette définition au cas où $f\colon A \to \mathbf{R}^n$, mais toujours avec $A \subset \mathbf{R}$. Dans ce cas, la *fonction dérivée* de f sera une fonction à valeurs dans $\mathbf{R}^n$.

Définition. Soient $A \subset \mathbf{R}$ et $f\colon A \to \mathbf{R}^n$. Soit x un point à l'intérieur de A. Alors la valeur de la *fonction dérivée* de f en x est l'unique vecteur $\boldsymbol{a} \in \mathbf{R}^n$ tel que

$$f(x+h) = f(x) + h\boldsymbol{a} + o(|h|) \qquad \text{lorsque } h \to 0$$

(si un tel $\boldsymbol{a}$ existe). Si x n'est pas un point à l'intérieur de A, la valeur de la *fonction dérivée* de f n'est pas définie en x.[3]

Remarque. Le produit d'un vecteur $\boldsymbol{a} \in \mathbf{R}^n$ et d'un scalaire $h \in \mathbf{R}$ peut être noté « $h\boldsymbol{a}$ » ou « $\boldsymbol{a}h$ ».

Définition. Soient $A \subset \mathbf{R}$ et $f\colon A \to \mathbf{R}^n$. La fonction f est dite *dérivable* ou *différentiable* en $x \in A$ si et seulement si x est dans le domaine de définition de sa fonction dérivée.

Notation. La fonction dérivée de f sera notée « f' » ou « Df ».

[1] Joseph-Louis Lagrange. *Théorie des fonctions analytiques. Contenant les principes du calcul différentiel.* Paris : Imprimerie de la République, 1797. xiii+277.

[2] Louis François Antoine Arbogast. *Du calcul des dérivations.* Strasbourg : Levrault frères, 1800. xxii+404.

[3] On peut avoir l'intérêt d'adopter une définition plus générale de la fonction dérivée de f, où son domaine de définition n'est pas obligé à être contenu dans l'intérieur du domaine de définition de f. Pourtant, une telle approche plus générale rend la théorie plus compliquée.

Ainsi, la fonction dérivée $f'\colon B \to \mathbf{R}^n$, avec $B \subset A \subset \mathbf{R}$, est définie par la formule :

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

ou par :

$$f(x+h) - f(x) = f'(x)h + o(|h|),$$

ou par :

$$f(x) - f(x-h) = f'(x)h + o(|h|),$$

où chaque fois « $o(|h|)$ » signifie $o(|h|)$ lorsque $h \to 0$. Le domaine de définition B de f' est l'ensemble des points $x \in A$ pour lesquels ces formules donnent une valeur définie de $f'(x)$. La fonction dérivée f' peut aussi être définie par la formule suivante, plus explicite :

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

ainsi que par

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

Remarque. D'habitude on restreint le domaine de définition de f' à l'*intérieur* topologique du domaine de définition de f : on ne tente de définir la valeur de f' en x que si f est définie au voisinage de x. Pourtant il peut-être utile d'adopter une définition plus générale et admettre que la valeur $f'(x) = \boldsymbol{a} \in \mathbf{R}^n$ est définie si elle est uniquement déterminée par la formule

$$f(x+h) = f(x) + \boldsymbol{a}h + o(|h|) \qquad \text{lorsque } h \to 0.$$

Avec cette définition généralisée, une fonction constante définie sur un intervalle fermée $[a,b]$ sera dérivable sur $[a,b]$ tout entier (y compris en a et en b), alors qu'avec la définition habituelle elle n'est dérivable que sur $]a,b[$.

Proposition. *Soient $A \subset \mathbf{R}$ et $f\colon A \to \mathbf{R}^n$. Si f est dérivable en $x \in A$, alors f est continue en x.*

Démonstration.

$$f(x+h) = f(x) + f'(x)h + o(|h|) \to f(x) \qquad \text{lorsque } h \to 0. \qquad \square$$

La proposition suivante établit un lien entre dérivation d'une fonction à valeurs dans $\mathbf{R}^n$ et dérivation des ses *fonctions composantes* à valeurs dans $\mathbf{R}$:

Proposition. *Soient $A \subset \mathbf{R}$, $f\colon A \to \mathbf{R}^n$, et $f_1, \ldots, f_n\colon A \to \mathbf{R}$ tels que*

$$f(x) = (f_1(x),\, \ldots,\, f_n(x)) \qquad \text{pour tout } x \in A.$$

Alors, pour tout $x \in A$,

(1) *f est dérivable en x si et seulement si $f_1, \ldots, f_m$ sont toutes dérivables en x, et*

(2) *si f est dérivable en x, alors*

$$f'(x) = (f_1'(x), \ldots, f_n'(x)).$$

Exercice. Prouver cette proposition.

II.1.2. Dérivée d'une fonction d'un argument dans R^m à valeurs dans R^n

D'après la dernière définition de la dérivée f' d'une fonction $f\colon A \to \mathbf{R}^n$, avec $A \subset \mathbf{R}$, la valeur $\boldsymbol{a} = f'(x) \in \mathbf{R}^n$ est déterminée par la formule :

$$f(x+h) = f(x) + \boldsymbol{a}h + o(|h|) \qquad \text{lorsque } h \to 0.$$

Tout vecteur $\boldsymbol{a} \in \mathbf{R}^n$ détermine une application linéaire $L\colon \mathbf{R} \to \mathbf{R}^n$ par la formule :

$$L(h) = \boldsymbol{a}h.$$

Ainsi, si $\boldsymbol{a} = f'(x)$ et $L(h) = \boldsymbol{a}h$, on a :

$$f(x+h) = f(x) + L(h) + o(|h|) \qquad \text{lorsque } h \to 0.$$

Soient maintenant $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Dans ce cas on définie la valeur de la *dérivée* de f en $X \in A$ comme une application linéaire (*opérateur linéaire*) $\mathbf{R}^m \to \mathbf{R}^n$.

Notation. Si E et F sont deux espaces vectoriels, l'ensemble des applications linéaires $\mathsf{E} \to \mathsf{F}$ sera noté $\mathscr{L}(\mathsf{E}, \mathsf{F})$.

Si $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$, alors la *fonction dérivée* de f, aussi dite la *dérivée totale* de f, sera une fonction $Df = f'\colon B \to \mathscr{L}(\mathbf{R}^m, \mathbf{R}^n)$, avec $B \subset A$.

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A. Alors la valeur de la *dérivée totale* de f en X (ou de la *dérivée* de f en X, tout court) est l'unique application linéaire $L\colon \mathbf{R}^m \to \mathbf{R}^n$ telle que

$$f(X+\boldsymbol{h}) = f(X) + L(\boldsymbol{h}) + o(\|\boldsymbol{h}\|) \qquad \text{lorsque } \boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^m}$$

(si une telle L existe). Si X n'est pas un point à l'intérieur de A, la valeur de la *dérivée totale* de f n'est pas définie en X.

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. La fonction f est dite *totalement dérivable* ou (totalement) *différentiable* en X si et seulement si la valeur de sa dérivée totale en X est définie.

La *norme* $\|\cdot\|$ dans $\mathbf{R}^m$ qui intervient dans cette définition dans « $o(\|\boldsymbol{v}\|)$ » peut être choisie arbitrairement sans que le choix affecte la définition de la dérivée. Souvent on utilise la norme euclidienne canonique $\|\cdot\|_2$ définie ainsi :

$$\|(v_1, \ldots, v_n)\|_2 = \sqrt{v_1^2 + \cdots + v_n^2}.$$

On pourrais aussi bien utiliser la norme $\|\cdot\|_1$ définie par :

$$\|(v_1, \ldots, v_n)\|_1 = |v_1| + \cdots + |v_n|\,,$$

ou la norme $\|\cdot\|_\infty$ définie par :

$$\|(v_1, \ldots, v_n)\|_\infty = \max\{|v_1|\,, \ldots, |v_n|\},$$

ou toute autre norme dans $\mathbf{R}^m$.

Notation. La dérivée (totale) de f sera notée « f' » ou « Df ».

Ainsi, f' est définie par la formule :

$$f(X+\boldsymbol{h}) = f(X) + f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|),$$

ou par :

$$f(X+\boldsymbol{h}) - f(X) = f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|),$$

ou par :

$$f(X) - f(X-\boldsymbol{h}) = f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|),$$

où chaque fois « $o(\|\boldsymbol{h}\|)$ » signifie $o(\|\boldsymbol{h}\|)$ lorsque $\boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^m}$.

Si l'application linéaire $f'(X)\colon \mathbf{R}^m \to \mathbf{R}^n$ est définie, elle peut être donnée par la formule

$$f'(X)(\boldsymbol{v}) = \lim_{h\to 0} \frac{f(X+h\boldsymbol{v}) - f(X)}{h},$$

ou par

$$f'(X)(\boldsymbol{v}) = \lim_{h\to 0} \frac{f(X) - f(X-h\boldsymbol{v})}{h}.$$

La proposition suivante en quelque sorte généralise ces formules :

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A tel que f est différentiable en X. Soient $\boldsymbol{v} \in \mathbf{R}^m$, $I \subset \mathbf{R}$ un intervalle, et $\gamma\colon I \to \mathbf{R}^m$ tels que*

$$\gamma(0) = X, \qquad \gamma'(0) = \boldsymbol{v}.$$

Alors

$$f'(X)(\boldsymbol{v}) = (f\circ\gamma)'(0) = \lim_{h\to 0} \frac{f(\gamma(h)) - f(X)}{h} = \lim_{h\to 0} \frac{f(X) - f(\gamma(-h))}{h}.$$

Remarque. Dans cet énoncé, la valeur de « $\gamma'(0)$ » est entendue au sens d'un élément de $\mathbf{R}^m$, et la valeur de « $(f \circ \gamma)'(0)$ » est entendu au sens d'un élément de $\mathbf{R}^n$, comme dans II.1.1.

Exercice. Prouver la dernière proposition.

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Si f est dérivable en $X \in A$, alors f est continue en X.*

Démonstration.

$$f(X + \boldsymbol{h}) = f(X) + f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|) \to f(X) \qquad \text{lorsque } \boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^m}.$$ □

La proposition suivante établie un lien entre dérivation d'une fonction à valeurs dans $\mathbf{R}^n$ et dérivation des ses fonctions composantes à valeurs dans $\mathbf{R}$:

Proposition. *Soient $A \subset \mathbf{R}^m$, $f\colon A \to \mathbf{R}^n$, et $f_1, \dots, f_n\colon A \to \mathbf{R}$ tels que*

$$f(X) = (f_1(X), \dots, f_n(X)) \qquad \textit{pour tout } X \in A.$$

Alors, pour tout $X \in A$,

(1) *f est différentiable en X si et seulement si $f_1, \dots, f_m$ sont toutes différentiables en X, et*

(2) *si f est différentiable en X, alors*

$$f'(X)(\boldsymbol{v}) = (f_1'(X)(\boldsymbol{v}), \dots, f_n'(X)(\boldsymbol{v})) \qquad \textit{pour tout } \boldsymbol{v} \in \mathbf{R}^m.$$

Exercice. Prouver cette proposition.

II.1.3. Deux sens de la « dérivée » d'une fonction $A \to \mathbf{R}^n$ avec $A \subset \mathbf{R}$

Pour les fonctions $f\colon A \to \mathbf{R}^n$ d'un argument réel (avec $A \subset \mathbf{R}$), on a ainsi deux définitions qui donnent deux sens différents au terme « *dérivée* de f », et en plus les notations « f' » et « Df » utilisées sont les mêmes dans les deux cas. Cette ambiguïté des définitions et des notations pour les « *dérivées* » de fonctions d'un argument réel ne doit pas poser de problèmes en pratique.

Considérons, par exemple, la fonction logarithme naturel $f = \ln\colon \,]0, \infty[\to \mathbf{R}$. Selon la première définition de la dérivée, on a :

$$f'(x) = \frac{1}{x} \in \mathbf{R} \qquad \text{pour tout } x \in\,]0, \infty[,$$
$$f'(x)v = \frac{v}{x} \in \mathbf{R} \qquad \text{pour tous } x \in\,]0, \infty[\text{ et } v \in \mathbf{R}.$$

Selon la deuxième définition de la dérivée, on a :

$$f'(x)\colon \mathbf{R} \to \mathbf{R} \qquad \text{pour tout } x \in\,]0, \infty[,$$
$$f'(x)(v) = \frac{v}{x} \in \mathbf{R} \qquad \text{pour tous } x \in\,]0, \infty[\text{ et } v \in \mathbf{R}.$$

Considérons encore la fonction $g\colon \mathbf{R} \to \mathbf{R}^2$ définie ainsi :

$$g(t) = (\cos t, \sin t) \in \mathbf{R}^2 \qquad \text{pour tout } t \in \mathbf{R}.$$

Selon la première définition de la dérivée, on a :

$$g'(t) = (-\sin t, \cos t) \in \mathbf{R}^2 \qquad \text{pour tout } t \in \mathbf{R},$$
$$g'(t)v = (-v \sin t, v \cos t) \in \mathbf{R}^2 \qquad \text{pour tous } t \in \mathbf{R} \text{ et } v \in \mathbf{R}.$$

Selon la deuxième définition de la dérivée, on a :

$$g'(t)\colon \mathbf{R} \to \mathbf{R}^2 \qquad \text{pour tout } t \in \mathbf{R},$$
$$g'(t)(v) = (-v \sin t, v \cos t) \in \mathbf{R}^2 \qquad \text{pour tous } t \in \mathbf{R} \text{ et } v \in \mathbf{R}.$$

La raison sous-jacent pour laquelle les deux définitions différentes des dérivées de fonctions d'un argument réel ne doivent pas provoquer une confusion est ce qu'on peut « identifier »[4] les éléments de $\mathbf{R}^n$ avec les applications linéaires $\mathbf{R} \to \mathbf{R}^n$.

II.1.4. Différentiabilité forte

La définition ordinaire (donnée ci-dessus) de la *différentiabilité* d'une fonction en un point peut être considérée trop faible ou peu naturelle.[5] On peut avoir l'intérêt de se servir d'une notion un peu plus forte et éventuellement plus naturelle, définie ainsi :

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A. La fonction f est dite *totalement fortement dérivable* ou (totalement) *fortement différentiable* en X si et seulement si elle est différentiable en X et que

$$f(X + \boldsymbol{h}_2) - f(X - \boldsymbol{h}_1) = f'(X)(\boldsymbol{h}_1 + \boldsymbol{h}_2) + o(\|\boldsymbol{h}_1 + \boldsymbol{h}_2\|)$$

lorsque $\boldsymbol{h}_1, \boldsymbol{h}_2 \to \mathbf{0}_{\mathbf{R}^m}$.

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Si f est fortement différentiable en $X \in A$, alors f est continue dans un voisinage de X.*

Exercice. Prouver cette proposition.

[4] Il y a un *isomorphisme* canonique entre les espaces vectoriels $\mathbf{R}^n$ et $\mathscr{L}(\mathbf{R}, \mathbf{R}^n)$.
[5] Giuseppe Peano. "Sur la définition de la dérivée". In : *Mathesis.* 2e sér. 2 (1892), p. 12-14. URL : http://resolver.sub.uni-goettingen.de/purl?PPN599218835_0012.

Exercice. Soit la fonction $f\colon \mathbf{R} \to \mathbf{R}$ définie ainsi :

$$f(x) = \begin{cases} x^2 & \text{si } x \in \mathbf{Q}, \\ 0 & \text{sinon.} \end{cases}$$

(1) Montrer que f est différentiable en 0.

(2) Montrer que f n'est continue qu'en 0.

Proposition. *Soient $A \subset \mathbf{R}$ et $f\colon A \to \mathbf{R}^n$. Soit $I \subset A$ un intervalle ouvert. Alors f est fortement différentiable en tout point de I si et seulement si f est différentiable en tout point de I et f' est continue dans I.*

Exercice. Prouver cette proposition.

Exercice. Soit la fonction $f\colon \mathbf{R} \to \mathbf{R}$ définie ainsi :

$$f(x) = \begin{cases} 0 & \text{si } x = 0, \\ x^2 \sin \dfrac{1}{x} & \text{sinon.} \end{cases}$$

(1) Montrer que f est différentiable partout dans $\mathbf{R}$.

(2) Montrer que f' n'est pas continue en 0.

II.1.5. Diverses appellations de la dérivée totale

L'application linéaire $f'(X) = Df(X)\colon \mathbf{R}^m \to \mathbf{R}^n$, qui est la valeur de la dérivée totale de $f\colon A \to \mathbf{R}^n$ en $X \in A$, où $A \subset \mathbf{R}^m$, est aussi couramment connue sous deux autres noms et parfois est notée autrement ainsi :

(1) comme la *différentielle* de f en X, notée « $\mathrm{d}f(X)$ » ou « $(\mathrm{d}f)_X$ »,

(2) comme l'*application tangente* de f en X, notée « $T_X f$ ».

On pourrais aussi l'appeler raisonnablement la *linéarisation* de f en X et noter, par exemple, « $L_X f$ », mais cette appellation n'est pas courante et, malheureusement, le terme « linéarisation » est parfois utilisé pour l'approximation *affine* de f au voisinage de X.

Les trois appellations sont issues des approches différentes et ne doivent pas être confondues dans tous les contextes. Par exemple, même si on peut confondre la *dérivée première* $f'(X) = Df(X)$ avec la *différentielle première* $\mathrm{d}f(X)$, cela n'est plus le cas[6] pour la *dérivée seconde*[7] $f''(X) = D^2 f(X)$ et la *différentielle seconde*[8] $\mathrm{d}^2 f(X)$.

[6] Sauf que différents auteurs définissent leurs termes différemment, et ce que certains appellent *dérivée seconde*, d'autre appellent *différentielle seconde*.

[7] La *dérivée seconde* $f''(X) = D^2 f(X)$ (de f en X) peut être identifiée à une application *bilinéaire* $\mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}^n$.

[8] La *différentielle seconde* $\mathrm{d}^2 f(X)$ (de f en X) est d'habitude définie comme l'application *quadratique* $\mathbf{R}^m \to \mathbf{R}^n$ associée à l'application bilinéaire $f''(X) = D^2 f(X)$ ainsi : $\mathrm{d}^2 f(X)(\boldsymbol{v}) = D^2(X)(\boldsymbol{v}, \boldsymbol{v})$.

En plus, il n'y a pas de consensus général sur les définitions exactes ou sur l'usage des termes « la dérivée », « la différentielle », « l'application tangente ».

En plus, il ne faut pas confondre la *différentielle* d'une *fonction* (d'une *application*) avec la *différentielle* d'une *variable*, où la différentielle d'une variable « x » est une nouvelle variable « dx ». Par exemple, ici « x », « y », « dx », « dy » sont quatre variables réelles :

$$d(x^2 + xy) = 2xdx + xdy + ydx.$$

Quant à l'expression « $d(x^2 + xy)$ », c'est une façon de noter $f'(x,y)(dx,dy)$ lorsque $f(x,y) = x^2 + xy$.

En plus, il y a une opération dite *différentiation extérieure*,[9] qui, elle aussi, est notée avec « d », et qui peut être considérée comme une certaine généralisation de l'opération de dérivation de fonctions.

La prolifération des différentes terminologies et notations, parfois incompatibles entre elles, est en une partie due à des raisons historiques : on tente de préserver le langage traditionnel en l'appliquant à des concepts qui évoluent, et différents auteurs parfois le font différemment.

II.1.6. Dérivées totales de quelques fonctions particulières

On va considérer les dérivées totales de quelque fonctions usuelles à valeurs réelles.

Pour une fonction $f\colon A \to \mathbf{R}$ avec $A \subset \mathbf{R}$, si on connaît une expression de $f'(x)$ au sens d'un nombre réel, on connaît une expression de $f'(x)$ au sens une application linéaire $\mathbf{R} \to \mathbf{R}$.

Par exemple, si

$$f(x) = x^3,$$

alors, au sens d'un nombre réel, $f'(x) \in \mathbf{R}$ est donné ainsi :

$$f'(x) = 3x^2.$$

Au sens d'une application linéaire, $f'(x)\colon \mathbf{R} \to \mathbf{R}$ est donnée ainsi :

$$f'(x)(v) = 3x^2 v \qquad \text{pour tout } v \in \mathbf{R}.$$

Si on utilise « dx » au lieu de « v » pour nommer l'argument de l'application $f'(x)$, on a :

$$f'(x)(dx) = 3x^2 dx \qquad \text{pour tout } dx \in \mathbf{R}.$$

Les tableaux suivants donnent les dérivées en x des fonctions exp, ln, sin et cos comme applications linéaires :

[9] La différentiation extérieure intervient notamment dans le *théorème de Stokes* sur l'intégration.

$f(x)$	$f'(x)(dx)$
e^x	$e^x dx$
$\ln x$	$\dfrac{dx}{x}$

$f(x)$	$f'(x)(dx)$
$\sin x$	$(\cos x)dx$
$\cos x$	$-(\sin x)dx$

Considérons maintenant les opérations d'addition, de soustraction, de multiplication et de division comme fonctions de deux arguments réels. Les tableaux suivants donnent leurs dérivées totales :

$f(x,y)$	$f'(x,y)(dx,dy)$
$x+y$	$dx+dy$
$x-y$	$dx-dy$

$f(x,y)$	$f'(x,y)(dx,dy)$
xy	$y\,dx+x\,dy$
$\dfrac{x}{y}$	$\dfrac{y\,dx-x\,dy}{y^2}$

Exercice. Vérifier si les formules données dans ces tableaux sont justes.

II.1.7. Notation classique (traditionnelle)

Une grande partie des notations utilisées en analyse sont apparues en XVII–XVIII siècles. À l'époque, les termes « variable » et « fonction » voulaient dire autre chose que ce qu'ils signifient maintenant en mathématique « moderne ». Malheureusement, l'ancien et le nouveau sens de ces termes sont souvent entremêlés ou confondus. Dans les domaines appliquées, comme dans la mécanique et dans la physique, on utilise ces termes couramment dans leurs sens original, qui n'est pas formalisé en mathématique moderne.

Pour donner une idée de ces deux sens différents, considérons la formule suivante qu'on peut rencontrer dans des textes sur l'analyse mathématique :

$$y = f(x).$$

En utilisant l'interprétation et la terminologie modernes, on dirait qu'ici f qui est une *fonction* (une *fonction* tout court, pas une fonction *de quelque chose*), et que x et y sont deux *objets mathématiques* (par exemple, deux nombres) tels que y est l'*image* de x par f, et que x est un *antécédent* de y par f. Les symboles « x », « y », « f » eux-mêmes sont dits *variables*. La valeur de la variable « f » ici est la fonction f, et les valeurs des variables « x » et « y » sont les objets mathématiques x et y, connus ou inconnus.

En utilisant l'interprétation et la terminologie des XVII–XVIII siècles, on dirait plutôt qu'ici x et y sont deux *quantités variables*, et que y est une *fonction* de x. Quant à « f », Lagrange l'appelait une *caractéristique*.[10] Il faut souligner que le fait que la quantité y soit une fonction de la quantité x n'empêche pas que x puisse au même temps être une fonction de y, ainsi que d'autres quantités variables.

`[...]`

Avec cette méta-notation abrégée, on a :

$$\begin{aligned} d(e^x) &= e^x dx, & d(\ln x) &= \frac{dx}{x},\\ d(\sin x) &= (\cos x)dx, & d(\cos x) &= -(\sin x)dx,\\ d(x+y) &= dx+dy, & d(x-y) &= dx-dy,\\ d(xy) &= y\,dx + x\,dy, & d\frac{x}{y} &= \frac{y\,dx - x\,dy}{y^2}. \end{aligned}$$

II.1.8. Dérivée totale d'une fonction composée

Théorème. *Soient $A \subset \mathbf{R}^p$, $B \subset \mathbf{R}^q$, $f\colon A \to \mathbf{R}^q$ et $g\colon B \to \mathbf{R}^r$. Soient X un point à l'intérieur de A et Y un point à l'intérieur de B tels que $Y = f(X)$. Supposons que f est différentiable en X et que g est différentiable en Y. Alors la fonction composée $g \circ f$ est différentiable en X et*

$$(g \circ f)'(X) = g'(Y) \circ f'(X).$$

Démonstration. Observons d'abord que si $L\colon \mathbf{R}^p \to \mathbf{R}^q$ est une application linéaire, alors $L(\boldsymbol{h}) = O(\|\boldsymbol{h}\|)$ lorsque $\boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^p}$.

Alors, lorsque $\boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^p}$, on a :

$$\begin{aligned} &(g\circ f)(X+\boldsymbol{h})\\ &\quad= g(f(X+\boldsymbol{h}))\\ &\quad= g(f(X) + f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|))\\ &\quad= g(Y + f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|))\\ &\quad= g(Y) + g'(Y)(f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|)) + o(\|f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|)\|)\\ &\quad= g(Y) + g'(Y)(f'(X)(\boldsymbol{h})) + g'(Y)(o(\|\boldsymbol{h}\|))\\ &\qquad\qquad + o(\|f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|)\|) \end{aligned}$$

Or, toujours lorsque $\boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^p}$,

$$g'(Y)(o(\|\boldsymbol{h}\|)) = O(\|o(\|\boldsymbol{h}\|)\|) = o(\|\boldsymbol{h}\|),$$

et

$$o(\|f'(X)(\boldsymbol{h}) + o(\|\boldsymbol{h}\|)\|) = o(\|O(\|\boldsymbol{h}\|) + o(\|\boldsymbol{h}\|)\|) = o(\|O(\|\boldsymbol{h}\|)\|) = o(\|\boldsymbol{h}\|).$$

[10] Le symbole « d » qui apparaît dans la notation de Leibniz pour les différentielles totales et pour les dérivées totales, ainsi que le symbole « ∂ » qui apparaît dans la notation de Legendre et de Jacobi pour les dérivées partielles, étaient aussi dits *caractéristiques*.

Ainsi,

$$\begin{aligned}(g\circ f)(X+\boldsymbol{h}) &= g(Y)+g'(Y)(f'(X)(\boldsymbol{h}))+o(\|\boldsymbol{h}\|)\\ &= (g\circ f)(X)+(g'(Y)\circ f'(X))(\boldsymbol{h})+o(\|\boldsymbol{h}\|)\end{aligned}$$

lorsque $\boldsymbol{h}\to\mathbf{0}_{\mathbf{R}^p}$. Donc, $g\circ f$ est différentiable en X, et

$$(g\circ f)'(X)=g'(Y)\circ f'(X). \qquad \square$$

En utilisant la méta-notation abrégée, le résultat de ce théorème peut être présenté ainsi :

$$d(g(f(X))=g'(f(X))(d(f(X)))=g'(f(X))(f'(X)(dX)),$$

ou ainsi :

$$\begin{aligned}d(g(f(X)) &= d(g(Y))|_{Y=f(X),\,dY=d(f(X))}\\ &= g'(Y)(dY)|_{Y=f(X),\,dY=d(f(X))}\\ &= g'(f(X))(d(f(X)))=g'(f(X))(f'(X)(dX)).\end{aligned}$$

II.2. 🚧 Dérivées directionnelles issues de la dérivée totale

Dans cette section, on va définir une espèce de *dérivées directionnelles* à partir de la dérivée totale. Plus tard, dans la section II.6, on va définir une espèce de *dérivées directionnelles* plus générales, au sens qu'elle prolongent les dérivées directionnelles définies ici.

Si on aura besoin de distinguer entre les deux notions de dérivées directionnelles, on va dire que celles définies dans cette section sont *issues de la dérivée totale*, et que celles définies dans la section II.6 sont *issues des restrictions sur les droites*. La plupart du temps on n'aura pas besoin de distinguer entre les deux notions car soit la différence n'aura pas d'importance dans le contexte donné, soit le contexte permettra d'éliminer toute ambiguïté. En plus, on va utiliser la même notation dans les deux cas en espérant que cela ne doit pas causer une confusion.

Définition. Soient $A\subset\mathbf{R}^m$ et $f\colon A\to\mathbf{R}^n$. Soit B la partie de A où f est différentiable. Soit $\boldsymbol{v}\in\mathbf{R}^m$. On va appeler la *dérivée directionnelle* de f suivant $\boldsymbol{v}$ la fonction $g\colon B\to\mathbf{R}^n$ définie par la formule :

$$g(X)=f'(X)(\boldsymbol{v}).$$

Notation. La dérivée directionnelle de f suivant un vecteur $\boldsymbol{v}$ sera notée « $D_{\boldsymbol{v}}f$ ».

Ainsi :

$$D_{\boldsymbol{v}}f(X)=f'(X)(\boldsymbol{v}).$$

Notation. On va écrire « $D^2_{\boldsymbol{v}}f$ » pour désigner $D_{\boldsymbol{v}}D_{\boldsymbol{v}}f$, « $D^3_{\boldsymbol{v}}f$ » pour désigner $D_{\boldsymbol{v}}D_{\boldsymbol{v}}D_{\boldsymbol{v}}f$, et ainsi de suite. En général :

$$D^0_{\boldsymbol{v}}f=f \qquad \text{et} \qquad D^{k+1}_{\boldsymbol{v}}f=D_{\boldsymbol{v}}D^k_{\boldsymbol{v}}f \qquad \text{pour tout } k\in\mathbf{N}.$$

Exercice. Soient $\boldsymbol{v}\in\mathbf{R}^m$ et $f,g\colon\mathbf{R}^m\to\mathbf{R}$. Supposons que f et g sont différentiables en $X\in\mathbf{R}^m$.

(1) Montrer que $f+g$ est différentiable en X, et que

$$D_{\boldsymbol{v}}(f+g)(X)=D_{\boldsymbol{v}}f(X)+D_{\boldsymbol{v}}g(X).$$

(2) Montrer que fg est différentiable en X, et que

$$D_{\boldsymbol{v}}(fg)(X)=D_{\boldsymbol{v}}f(X)g(X)+f(X)D_{\boldsymbol{v}}g(X).$$

Proposition. *Soient $A\subset\mathbf{R}^m$, $f\colon A\to\mathbf{R}^n$, et $f_1,\dots,f_n\colon A\to\mathbf{R}$ tels que*

$$f(X)=(f_1(X),\,\dots,\,f_n(X)) \qquad \textit{pour tout } X\in A.$$

Alors, pour tout $\boldsymbol{v}\in\mathbf{R}^m$ et pour tout $X\in A$ tel que f est différentiable en X,

$$D_{\boldsymbol{v}}f(X)=(D_{\boldsymbol{v}}f_1(X),\,\dots,\,D_{\boldsymbol{v}}f_n(X)).$$

Exercice. Prouver cette proposition.

La proposition suivante est un corollaire plus ou moins immédiat des définitions.

Proposition. *Soient $A\subset\mathbf{R}^m$ et $f\colon A\to\mathbf{R}^n$. Soit X un point à l'intérieur de A tel que f est différentiable en X. Alors*

$$D_{\boldsymbol{u}+\boldsymbol{v}}f(X)=D_{\boldsymbol{u}}f(X)+D_{\boldsymbol{v}}f(X) \qquad \textit{pour tous } \boldsymbol{u},\boldsymbol{v}\in\mathbf{R}^m,$$

et

$$D_{\alpha\boldsymbol{v}}f(X)=\alpha D_{\boldsymbol{v}}f(X) \qquad \textit{pour tous } \alpha\in\mathbf{R} \textit{ et } \boldsymbol{v}\in\mathbf{R}^m.$$

Exercice. Prouver cette proposition.

Corollaire. *Soient $A\subset\mathbf{R}^m$ et $f\colon A\to\mathbf{R}^n$. Soient $\boldsymbol{v}\in\mathbf{R}^m$, $k\in\mathbf{N}$ et $X\in A$ tels que la fonction $D^k_{\boldsymbol{v}}f$ soit définie en X. Alors pour tout $\alpha\in\mathbf{R}$, la fonction $D^k_{\alpha\boldsymbol{v}}f$ est définie en X, et*

$$D^k_{\alpha\boldsymbol{v}}f(X)=\alpha^kD^k_{\boldsymbol{v}}f(X).$$

Exercice. Prouver ce corollaire.

Rappelons-nous une proposition de II.1.2, en utilisant la nouvelle notation :

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A tel que f est différentiable en X. Soient $\boldsymbol{v} \in \mathbf{R}^m$, $I \subset \mathbf{R}$ un intervalle, et $\gamma\colon I \to \mathbf{R}^m$ tels que*

$$\gamma(0) = X, \qquad \gamma'(0) = \boldsymbol{v}.$$

Alors

$$D_{\boldsymbol{v}} f(X) = (f \circ \gamma)'(0) = \lim_{h \to 0} \frac{f(\gamma(h)) - f(X)}{h} = \lim_{h \to 0} \frac{f(X) - f(\gamma(-h))}{h}.$$

Cette proposition peut étendue ainsi :

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soient $\boldsymbol{v} \in \mathbf{R}^m$, $K \in \mathbf{N}$ et $X \in A$ tels que la fonction $D_{\boldsymbol{v}}^K f$ soit définie en X. Soient $I \subset \mathbf{R}$ un intervalle et $\gamma\colon I \to \mathbf{R}^m$ une fonction telle que*

$$\gamma(0) = X, \qquad \gamma'(0) = \boldsymbol{v}, \qquad \textit{et} \qquad \gamma^{(k)}(0) = \mathbf{0}_{\mathbf{R}^m} \qquad \textit{pour } 2 \leqslant k \leqslant K.$$

Alors

$$D_{\boldsymbol{v}}^k f(X) = (f \circ \gamma)^{(k)}(0) \qquad \textit{pour } 0 \leqslant k \leqslant K.$$

Exercice. Prouver cette proposition.

II.3. Dérivées partielles

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit B la partie de A où f est différentiable. Alors les *dérivées partielles* de f sont les fonctions $g_1, \dots, g_m\colon B \to \mathbf{R}^n$ définies par l'identité :

$$f'(X)(dX) = g_1(X)dx_1 + \cdots + g_m(X)dx_m$$

pour tout $dX = (dx_1, \dots, dx_m) \in \mathbf{R}^m$.

Notation. On pourra noter les dérivées partielles de f comme « $D_1 f$ », …, « $D_m f$ ».

Avec cette notation, les dérivées partielles sont définies par l'identité :

$$Df(X)(dx_1, \dots, dx_m) = D_1 f(X)dx_1 + \cdots + D_m f(X)dx_m.$$

La notation traditionnelle pour les valeurs des dérivées partielles $D_1 f$, …, $D_m f$ en $X = (x_1, \dots, x_m)$ est « $\dfrac{\partial f(X)}{\partial x_1}$ », …, « $\dfrac{\partial f(X)}{\partial x_m}$ » :

$$\frac{\partial f(X)}{\partial x_1} = D_1 f(X), \qquad \dots, \qquad \frac{\partial f(X)}{\partial x_m} = D_m f(X),$$

où il est entendu que $X = (x_1, \dots, x_m)$. Il est entendu ici que les parenthèses implicites sont placées ainsi :

$$\frac{\partial(f(X))}{\partial x_1} = (D_1 f)(X), \qquad \dots, \qquad \frac{\partial(f(X))}{\partial x_m} = (D_m f)(X).$$

La notation traditionnelle pour $f'(X)(dX)$ est $d(f(X))$:

$$d(f(X)) = f'(X)(dX).$$

Avec la notation traditionnelle, les dérivées partielles sont définies par l'identité :

$$d(f(X)) = \frac{\partial(f(X))}{\partial x_1}dx_1 + \cdots + \frac{\partial(f(X))}{\partial x_m}dx_m,$$

si $X = (x_1, \dots, x_m)$ et $dX = (dx_1, \dots, dx_m)$.

Les dérivées partielles coincident avec les dérivées directionnelles suivant les vecteurs de la base canonique :

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ la base canonique de $\mathbf{R}^m$. Alors*

$$D_1 f(X) = D_{\boldsymbol{e}_1} f(X), \qquad \dots, \qquad D_m f(X) = D_{\boldsymbol{e}_m} f(X).$$

Exercice. Prouver cette proposition.

II.4. Matrice jacobienne

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A où f est différentiable. Alors définissons la *matrice jacobienne* de f en X comme la matrice de l'application linéaire $f'(X)\colon \mathbf{R}^m \to \mathbf{R}^n$ par rapport aux bases canoniques de $\mathbf{R}^m$ et de $\mathbf{R}^n$.

Notation. La matrice jacobienne f en X sera notée « $J_f(X)$ ».

Si, par exemple, on utilise la notation « $[L]$ » pour la matrice d'une application linéaire $L\colon \mathbf{R}^m \to \mathbf{R}^n$ par rapport aux bases canoniques de $\mathbf{R}^m$ et de $\mathbf{R}^n$, alors on peut donner la définition suivante de la matrice jacobienne :

$$J_f(X) = [f'(X)].$$

Soient $A \subset \mathbf{R}^m$, $f\colon A \to \mathbf{R}^n$, et $f_1, \dots, f_n\colon A \to \mathbf{R}$ tels que

$$f(x) = (f_1(x), \dots, f_n(x)) \qquad \text{pour tout } x \in A.$$

Soit $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ la base canonique de $\mathbf{R}^m$. Alors, pour tout $X \in A$ tel que f est différentiable en X, on a :

$$J_f(X) = \begin{pmatrix} f_1'(X)(\boldsymbol{e}_1) & \cdots & f_1'(X)(\boldsymbol{e}_m) \\ \vdots & \ddots & \vdots \\ f_n'(X)(\boldsymbol{e}_1) & \cdots & f_n'(X)(\boldsymbol{e}_m) \end{pmatrix} = \begin{pmatrix} D_{\boldsymbol{e}_1} f_1(X) & \cdots & D_{\boldsymbol{e}_m} f_1(X) \\ \vdots & \ddots & \vdots \\ D_{\boldsymbol{e}_1} f_n(X) & \cdots & D_{\boldsymbol{e}_m} f_n(X) \end{pmatrix}.$$

En termes des dérivées partielles :

$$J_f(X) = \begin{pmatrix} D_1 f_1(X) & \cdots & D_m f_1(X) \\ \vdots & \ddots & \vdots \\ D_1 f_n(X) & \cdots & D_m f_n(X) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1(X)}{\partial x_1} & \cdots & \dfrac{\partial f_1(X)}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n(X)}{\partial x_1} & \cdots & \dfrac{\partial f_n(X)}{\partial x_m} \end{pmatrix},$$

où $X = (x_1, \ldots, x_m)$.

Définition. Si $A \subset \mathbf{R}^n$, $f\colon A \to \mathbf{R}^n$, et X est un point à l'intérieur de A où f est différentiable, alors le déterminant de la matrice jacobienne $J_f(X)$ (qui est carrée dans ce cas) est dit le *déterminant jacobien*, ou le *jacobien*, de f en X.

II.5. Gradient

Rappelons-nous que le *produit scalaire canonique* dans $\mathbf{R}^m$ est définie par la formule :

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle \overset{\text{déf}}{=} u_1 v_1 + \cdots + u_m v_m,$$

où $\boldsymbol{u}, \boldsymbol{v} \in \mathbf{R}^m$, $\boldsymbol{u} = (u_1, \ldots, u_m)$ et $\boldsymbol{v} = (v_1, \ldots, v_m)$.

Définition. Soient $A \subset \mathbf{R}$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A où f est différentiable. Alors le *gradient* de f en X est le vecteur $\boldsymbol{u} \in \mathbf{R}^m$ tel que

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = f'(X)(\boldsymbol{v}) \qquad \text{pour tout vecteur } \boldsymbol{v} \in \mathbf{R}^m.$$

Notation. On peut noter le gradient de f comme « $\operatorname{grad} f$ ».

Ainsi :

$$\langle (\operatorname{grad} f)(X), \boldsymbol{v} \rangle = f'(X)(\boldsymbol{v}).$$

Exemple. Soit f la fonction définie par :

$$f(x, y, z) = \frac{1}{x^2 + y^2 + z^2}, \qquad x, y, z \in \mathbf{R}.$$

Alors,

$$\begin{aligned} f'(x, y, z)(dx, dy, dz) = d(f(x, y, z)) &= \frac{-2x\,dx - 2y\,dy - 2z\,dz}{(x^2 + y^2 + z^2)^2} \\ &= \frac{-2\,\langle (x, y, z), (dx, dy, dz) \rangle}{(x^2 + y^2 + z^2)^2}. \end{aligned}$$

D'où,

$$(\operatorname{grad} f)(x, y, z) = \frac{-2(x, y, z)}{(x^2 + y^2 + z^2)^2}.$$

II.6. Dérivées directionnelles issues des restrictions sur les droites

SECTION-BROUILLON

Soient $A \subset \mathbf{R}^m$, $f\colon A \to \mathbf{R}^n$, et $\boldsymbol{v} \in \mathbf{R}^m$. Dans la section II.2, on a définit la *dérivée directionnelle* $D_{\boldsymbol{v}} f$ par la formule :

$$D_{\boldsymbol{v}} f(X) = f'(X)(\boldsymbol{v}).$$

D'après une proposition de II.1.2, si f est différentiable en X et que $\gamma\colon I \to \mathbf{R}^m$ est une n'importe quelle fonction définie sur un intervalle $I \subset \mathbf{R}$ telle que

$$\gamma(0) = X \qquad \text{et} \qquad \gamma'(0) = \boldsymbol{v},$$

alors

$$D_{\boldsymbol{v}} f(X) = (f \circ \gamma)'(0).$$

Comme le choix « naturel » d'une telle fonction γ, on peut prendre la fonction $\gamma\colon \mathbf{R} \to \mathbf{R}^m$ définie ainsi :

$$\gamma(t) = X + \boldsymbol{v}t.$$

(Ceci est une paramétrisation de la droite passant par X avec un vecteur directeur $\boldsymbol{v}$.) Avec cette γ,

$$D_{\boldsymbol{v}} f(X) = (f \circ \gamma)'(0) = \lim_{h \to 0} \frac{f(X + \boldsymbol{v}h) - f(X)}{h} = \lim_{h \to 0} \frac{f(X) - f(X - \boldsymbol{v}h)}{h}.$$

On va adopter cette formule pour re-définir la *dérivée directionnelle* de f suivant le vecteur $\boldsymbol{v}$.

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. La *dérivée directionnelle* de f suivant un vecteur $\boldsymbol{v} \in \mathbf{R}^m$ est la fonction g définie par la formule :

$$g(X) \overset{\text{déf}}{=} \lim_{h \to 0} \frac{f(X + \boldsymbol{v}h) - f(X)}{h} = \lim_{h \to 0} \frac{f(X) - f(X - \boldsymbol{v}h)}{h}.$$

Remarque. Certains auteurs utilisent le terme « dérivée directionnelle de f suivant $\boldsymbol{v}$ » pour la fonction g définie ainsi :

$$g(X) = \lim_{h \to 0,\, h > 0} \frac{f(X + \boldsymbol{v}h) - f(X)}{h} = \lim_{h \to 0,\, h < 0} \frac{f(X) - f(X - \boldsymbol{v}h)}{h}.$$

Remarque. Par contre, il serait très difficile de trouver quelqu'un qui utilise l'appellation « dérivée directionnelle de f suivant $\boldsymbol{v}$ » pour la fonction g définie ainsi :

$$g(X) = \lim_{h \to 0,\, h < 0} \frac{f(X + \boldsymbol{v}h) - f(X)}{h} = \lim_{h \to 0,\, h > 0} \frac{f(X) - f(X - \boldsymbol{v}h)}{h}.$$

Peut-être l'origine de cette discrimination est dans le préjugé selon lequel le vecteur vitesse d'un point mobile est d'habitude dessiné comme une flèche « partant » du point, plutôt qu'y « arrivant ».

Notation. La dérivée directionnelle de f suivant un vecteur $\boldsymbol{v}$ sera notée « $D_{\boldsymbol{v}} f$ ».

Observons que si le domaine de définition de f est A, le domaine de définition de f' est B, et que le domaine de définition de $D_{\boldsymbol{v}} f$ est $B_{\boldsymbol{v}}$, alors $B \subset B_{\boldsymbol{v}} \subset A$.

Exemple. Considérons la fonction $f\colon \mathbf{R}^2 \to \mathbf{R}$ définie ainsi :

$$f(x,y) = |x-y|\,.$$

Alors le domaine de définition de $f' = Df$ est $\{\,(x,y) \in \mathbf{R}^2 \mid x \neq y\,\}$, et

$$f'(x,y)(dx,dy) = \begin{cases} dx - dy & \text{si } x > y, \\ dy - dx & \text{si } y > x. \end{cases}$$

Posons $\boldsymbol{v} = (1,1) \in \mathbf{R}^2$. Alors $D_{\boldsymbol{v}} f$ est la fonction nulle $\mathbf{R}^2 \to \mathbf{R}$, définie sur $\mathbf{R}^2$ tout entier.

Notation. On va écrire « $D^2_{\boldsymbol{v}} f$ » pour désigner $D_{\boldsymbol{v}} D_{\boldsymbol{v}} f$, « $D^3_{\boldsymbol{v}} f$ » pour désigner $D_{\boldsymbol{v}} D_{\boldsymbol{v}} D_{\boldsymbol{v}} f$, et ainsi de suite. En général :

$$D^0_{\boldsymbol{v}} f = f \qquad \text{et} \qquad D^{k+1}_{\boldsymbol{v}} f = D_{\boldsymbol{v}} D^k_{\boldsymbol{v}} f \qquad \text{pour tout } k \in \mathbf{N}.$$

Exercice. Soient $\boldsymbol{v} \in \mathbf{R}^m$ et $f, g\colon \mathbf{R}^m \to \mathbf{R}$. Supposons que les dérivées directionnelles $D_{\boldsymbol{v}} f$ et $D_{\boldsymbol{v}} g$ sont définies en $X \in \mathbf{R}^m$.

(1) Montrer que $D_{\boldsymbol{v}}(f+g)$ est définie en X, et que

$$D_{\boldsymbol{v}}(f+g)(X) = D_{\boldsymbol{v}} f(X) + D_{\boldsymbol{v}} g(X).$$

(2) Montrer que $D_{\boldsymbol{v}}(fg)$ est définie en X, et que

$$D_{\boldsymbol{v}}(fg)(X) = D_{\boldsymbol{v}} f(X) g(X) + f(X) D_{\boldsymbol{v}} g(X).$$

Proposition. *Soient $A \subset \mathbf{R}^m$, $f\colon A \to \mathbf{R}^n$, et $f_1, \ldots, f_n\colon A \to \mathbf{R}$ tels que*

$$f(X) = (f_1(X), \ldots, f_n(X)) \qquad \textit{pour tout } X \in A.$$

Alors, pour tout $\boldsymbol{v} \in \mathbf{R}^m$ et pour tout $X \in A$,

(1) *la dérivée directionnelle $D_{\boldsymbol{v}} f$ est définie en X si et seulement si les dérivées directionnelles $D_{\boldsymbol{v}} f_1, \ldots, D_{\boldsymbol{v}} f_n$ sont toutes définies en X, et*

(2) *si la dérivée directionnelle $D_{\boldsymbol{v}} f$ est définie en X, alors*

$$D_{\boldsymbol{v}} f(X) = (D_{\boldsymbol{v}} f_1(X), \ldots, D_{\boldsymbol{v}} f_n(X)).$$

Exercice. Prouver cette proposition.

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A tel que f est différentiable en X. Alors pour tout $\boldsymbol{v} \in \mathbf{R}^m$, la dérivée directionnelle $D_{\boldsymbol{v}} f$ est définie en X, et*

$$D_{\boldsymbol{v}} f(X) = f'(X)(\boldsymbol{v}).$$

Exercice. Prouver cette proposition.

Corollaire. *Sous les hypothèses de la proposition précédente,*

$$D_{\boldsymbol{u}+\boldsymbol{v}} f(X) = D_{\boldsymbol{u}} f(X) + D_{\boldsymbol{v}} f(X) \qquad \textit{pour tous } \boldsymbol{u}, \boldsymbol{v} \in \mathbf{R}^m,$$

et

$$D_{\alpha\boldsymbol{v}} f(X) = \alpha D_{\boldsymbol{v}} f(X) \qquad \textit{pour tous } \alpha \in \mathbf{R} \textit{ et } \boldsymbol{v} \in \mathbf{R}^m.$$

En fait, la deuxième partie de ce corollaire peut être démontrée sans l'hypothèse de la dérivabilité totale de f en X.

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soient $\boldsymbol{v} \in \mathbf{R}^m$ et $X \in A$ tels que la dérivée directionnelle $D_{\boldsymbol{v}} f$ soit définie en X. Alors pour tout $\alpha \in \mathbf{R}$, la dérivée directionnelle $D_{\alpha\boldsymbol{v}} f$ est définie en X, et*

$$D_{\alpha\boldsymbol{v}} f(X) = \alpha D_{\boldsymbol{v}} f(X).$$

Exercice. Prouver cette proposition.

Corollaire. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soient $\boldsymbol{v} \in \mathbf{R}^m$, $k \in \mathbf{N}$ et $X \in A$ tels que la fonction $D^k_{\boldsymbol{v}} f$ soit définie en X. Alors pour tout $\alpha \in \mathbf{R}$, la fonction $D^k_{\alpha\boldsymbol{v}} f$ est définie en X, et*

$$D^k_{\alpha\boldsymbol{v}} f(X) = \alpha^k D^k_{\boldsymbol{v}} f(X).$$

Exercice. Prouver ce corollaire.

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soient $\boldsymbol{v} \in \mathbf{R}^m$, $X \in A$, et $\gamma\colon \mathbf{R} \to \mathbf{R}$ définie par :*

$$\gamma(t) = X + \boldsymbol{v}t.$$

Alors, pour tout $k \in \mathbf{N}$, $D^k_{\boldsymbol{v}} f$ est définie en X si et seulement si $(f \circ \gamma)^{(k)}$ est définie en 0, et dans ce cas

$$D^k_{\boldsymbol{v}} f(X) = (f \circ \gamma)^{(k)}(0).$$

Exercice. Prouver cette proposition.

Théorème. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ une base de $\mathbf{R}^m$. Soit U une partie ouverte de $\mathbf{R}^m$ telle que toutes les dérivées directionnelles $D_{\boldsymbol{v}_1} f, \ldots, D_{\boldsymbol{v}_m} f$ soient définies et continues dans U. Alors f est continue et différentiable dans U.*

II.7. Classes de régularité

Si $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$, on dit que f est (totalement) *différentiable* ou *totalement dérivable* en X si et seulement si la dérivée totale de f est définie en X, et on dit que f est (totalement) *différentiable* ou *totalement dérivable* dans un ensemble U si et seulement si f est différentiable en tout point de U.

Définition. Soit U une partie ouverte de $\mathbf{R}^m$. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$.

(1) La fonction f est dite être 1 *fois différentiable* ou 1 *fois totalement dérivable* dans U si et seulement si f est définie et différentiable dans U.

(2) Pour $k \in \mathbf{N} \setminus \{0\}$, la fonction f est dite être $k+1$ *fois différentiable* ou $k+1$ *fois totalement dérivable* dans U si et seulement si
- f est définie et différentiable dans U, et
- pour tout vecteur $\boldsymbol{v} \in \mathbf{R}^m$, la dérivée directionnelle $D_{\boldsymbol{v}} f$ est k fois différentiable dans U.

(3) La fonction f est dite être *indéfiniment différentiable* ou *indéfiniment totalement dérivable* dans U si et seulement si elle y est k fois différentiable pour tout $k \in \mathbf{N} \setminus \{0\}$.

Définition. Soient $A \subset \mathbf{R}^m$, $f\colon A \to \mathbf{R}^n$, et $X \in A$ un point à l'intérieur de A.

(1) La fonction f est dite être 1 *fois différentiable* ou 1 *fois totalement dérivable* en X si et seulement si f est différentiable en X.

(2) Pour $k \in \mathbf{N} \setminus \{0\}$, la fonction f est dite être $k+1$ *fois différentiable* ou $k+1$ *fois totalement dérivable* en X si et seulement si
- f est différentiable dans un voisinage de X, et
- pour tout vecteur $\boldsymbol{v} \in \mathbf{R}^m$, la dérivée directionnelle $D_{\boldsymbol{v}} f$ est k fois différentiable en X.

(3) La fonction f est dite être *indéfiniment différentiable* ou *indéfiniment totalement dérivable* en X si et seulement si elle y est k fois différentiable pour tout $k \in \mathbf{N}\setminus\{0\}$.

Définition. Soit U une partie ouverte de $\mathbf{R}^m$. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$.

(1) La fonction f est dite être de *classe* C^0 ou de *classe* C dans U si et seulement si f est définie et continue dans U.

(2) Pour $k \in \mathbf{N}$, la fonction f est dite être de *classe* C^{k+1} ou $k+1$ *fois continûment différentiable* dans U si et seulement si
- f est définie et différentiable dans U, et
- pour tout vecteur $\boldsymbol{v} \in \mathbf{R}^m$, la dérivée directionnelle $D_{\boldsymbol{v}} f$ est de classe C^k dans U.

(3) La fonction f est dite être de *classe* C^∞ dans U si et seulement si elle y est de classe C^k pour tout $k \in \mathbf{N}$.

Notation. Si U est une partie ouverte de $\mathbf{R}^m$, alors l'ensemble de fonctions $U \to \mathbf{R}^n$ de classe C^k dans U sera noté « $C^k(U, \mathbf{R}^n)$ ».

Observons que :

(1) si f est k fois différentiable dans une partie ouverte U de $\mathbf{R}^m$, alors f est de classe C^{k-1} dans U et k fois différentiable en tout point de U ;

(2) si f est indéfiniment différentiable dans une partie ouverte U de $\mathbf{R}^m$, alors f est de classe C^∞ dans U ;

(3) si f est de classe C^k dans une partie ouverte U de $\mathbf{R}^m$, alors f est k fois différentiable dans U ;

(4) si f est de classe C^∞ dans une partie ouverte U de $\mathbf{R}^m$, alors f est indéfiniment différentiable différentiable dans U ;

(5) si f est k fois différentiable en un point X à l'intérieur de son domaine de définition, et que $k \geqslant 2$, alors f est $k-1$ fois différentiable dans un voisinage de X.

Proposition. *Soit U une partie ouverte de $\mathbf{R}^m$. Soient $A \subset \mathbf{R}^m$, $f\colon A \to \mathbf{R}^n$, et $f_1, \dots, f_n\colon A \to \mathbf{R}$ tels que*

$$f(X) = (f_1(X), \dots, f_n(X)) \qquad \textit{pour tout } X \in A.$$

Alors, pour tout $k \in \mathbf{N} \cup \{\infty\}$, f est de classe C^k dans U si et seulement si $f_1, \dots, f_n$ sont toutes de classe C^k dans U.

Théorème. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit U une partie ouverte de $\mathbf{R}^m$ telle que f est différentiable dans U. Soit $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ une base de $\mathbf{R}^m$. Alors :*

(1) *pour tout $k \in \mathbf{N} \setminus \{0\}$, f est $k+1$ fois différentiable dans U si et seulement si les dérivées directionnelles $D_{\boldsymbol{v}_1} f, \dots, D_{\boldsymbol{v}_m} f$ sont toutes k fois différentiables dans U,*

(2) *f est indéfiniment différentiable dans U si et seulement si les dérivées directionnelles $D_{\boldsymbol{v}_1} f, \dots, D_{\boldsymbol{v}_m} f$ sont toutes indéfiniment différentiables dans U,*

(3) *pour tout $k \in \mathbf{N}$, f est de classe C^{k+1} dans U si et seulement si les dérivées directionnelles $D_{\boldsymbol{v}_1} f, \dots, D_{\boldsymbol{v}_m} f$ sont toutes de classe C^k dans U.*

Corollaire. *Sous les hypothèses du théorème précédent,*

(1) *pour tout $k \in \mathbf{N} \setminus \{0\}$, f est $k+1$ fois différentiable dans U si et seulement si les dérivées partielles $D_1 f, \dots, D_m f$ sont toutes k fois différentiables dans U,*

(2) *f est indéfiniment différentiable dans U si et seulement si les dérivées partielles $D_1 f, \dots, D_m f$ sont toutes indéfiniment différentiables dans U,*

(3) *pour tout $k \in \mathbf{N}$, f est de classe C^{k+1} dans U si et seulement si les dérivées partielles $D_1 f, \ldots, D_m f$ sont toutes de classe C^k dans U.*

Théorème. *Soient $A \subset \mathbf{R}^p$, $B \subset \mathbf{R}^q$, $f\colon A \to \mathbf{R}^q$ et $g\colon B \to \mathbf{R}^r$. Soient U une partie ouverte de $\mathbf{R}^p$ et V une partie ouverte de $\mathbf{R}^q$ telles que $U \subset A$, $V \subset B$ et $f(U) \subset V$ (c'est-à-dire, $f(X) \in V$ pour tout $X \in U$). Alors :*

(1) *pour tout $k \in \mathbf{N} \setminus \{0\}$, si f est k fois différentiable dans U et que g est est k fois différentiable dans V, alors la fonction composée $g \circ f$ est k fois différentiable dans U,*

(2) *si f est indéfiniment différentiable dans U et que g est est indéfiniment différentiable dans V, alors la fonction composée $g \circ f$ est indéfiniment différentiable dans U,*

(3) *pour tout $k \in \mathbf{N}$, si f est de classe C^k dans U et que g est est de classe C^k dans V, alors la fonction composée $g \circ f$ est de classe C^k dans U.*

Théorème. *Soient $A \subset \mathbf{R}^p$, $B \subset \mathbf{R}^q$, $f\colon A \to \mathbf{R}^q$ et $g\colon B \to \mathbf{R}^r$. Soient X un point à l'intérieur de A et Y un point à l'intérieur de B tels que $f(X) = Y$. Alors :*

(1) *pour tout $k \in \mathbf{N} \setminus \{0\}$, si f est k fois différentiable en X et que g est est k fois différentiable en Y, alors la fonction composée $g \circ f$ est k fois différentiable en X,*

(2) *si f est indéfiniment différentiable en X et que g est est indéfiniment différentiable en Y, alors la fonction composée $g \circ f$ est indéfiniment différentiable en X.*

II.8. Dérivée seconde et dérivées d'ordres supérieurs

Théorème. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A tel que f est deux fois différentiable en X. Alors pour tous $\boldsymbol{u}, \boldsymbol{v} \in \mathbf{R}^m$,*

$$D_{\boldsymbol{v}} D_{\boldsymbol{u}} f(X) = D_{\boldsymbol{u}} D_{\boldsymbol{v}} f(X).$$

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A tel que f est deux fois différentiable en X. Alors définissons la valeur de la *dérivée totale seconde* de f en X comme l'application bilinéaire symétrique $B\colon \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}^n$ définie par la formule :

$$B(\boldsymbol{u}, \boldsymbol{v}) \stackrel{\text{déf}}{=} D_{\boldsymbol{u}} D_{\boldsymbol{v}} f(X) = D_{\boldsymbol{v}} D_{\boldsymbol{u}} f(X).$$

Notation. La dérivée totale seconde de f sera notée « f'' » ou « $D^2 f$ ».

Ainsi,

$$f''(X)(\boldsymbol{u}, \boldsymbol{v}) = D_{\boldsymbol{u}} D_{\boldsymbol{v}} f(X) = D_{\boldsymbol{v}} D_{\boldsymbol{u}} f(X)$$

lorsque f est deux fois différentiable en X.

Remarque. La définition la plus courante du terme « dérivée totale seconde » est un peu différente de celle donnée ci-dessus. Étant donnés $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$, d'habitude on définit la *dérivée totale seconde* de f comme la dérivée totale de la dérivée totale de f. Or, pour faire ainsi, il faut d'abord donner un sens au « dérivée totale » d'une fonction dont les valeurs sont des applications linéaires, vu que les valeurs de f' sont des applications linéaires $\mathbf{R}^m \to \mathbf{R}^n$. Cependant, dans un certain sens précis, on peut « identifier » la valeur de la dérivé totale de la dérivée totale de f en X avec l'application bilinéaire symétrique $B\colon \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}^n$ définie par $B(\boldsymbol{u}, \boldsymbol{v}) = D_{\boldsymbol{v}} D_{\boldsymbol{u}} f(X) = D_{\boldsymbol{u}} D_{\boldsymbol{v}} f(X)$. Cela justifie le choix de la définition donnée.

Lorsque f est k fois différentiable en X, on peut, de manière analogique, définir la valeur de la *dérivée totale d'ordre k* de f en X, qui sera une application multilinéaire (k-linéaire) symétrique $(\mathbf{R}^m)^k \to \mathbf{R}^n$.

II.9. Dérivée seconde et matrice hessienne d'une fonction à valeurs réelles

SECTION-BROUILLON

Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit X un point à l'intérieur de A où f est deux fois différentiable. Alors, d'après la définition donnée dans la section II.8,

$$f''(X)(\boldsymbol{u}, \boldsymbol{v}) = D_{\boldsymbol{u}} D_{\boldsymbol{v}} f(X) = D_{\boldsymbol{v}} D_{\boldsymbol{u}} f(X).$$

Comme l'application $f''(X)\colon \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}$ est bilinéaire et symétrique, les propriétés des applications bilinéaires symétriques s'y appliquent.

Voici donc quelques théorèmes pertinents, qui peuvent être traités plus en détails dans un cours d'algèbre linéaire.

Théorème. *Soit $B\colon \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}$ une application bilinéaire symétrique. Alors il existe une base $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ de $\mathbf{R}^m$ telle que*

$$B(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0 \qquad \text{pour tous } i \text{ et } j \text{ tels que } i \neq j.$$

Ce théorème peut être démontré à l'aide de l'algèbre linéaire assez basique, et un algorithme assez simple pour trouver une base comme dans la conclusion peut être donné.

Observons que l'énoncé du dernier théorème ne dit pas que la base dont le théorème affirme l'existence soit orthonormée, ni que les vecteurs de la base soient orthogonaux entre eux.

Le théorème suivant est plus difficile à démontrer, et une base dont il affirme l'existence est plus compliquée à calculer.

Théorème. *Soit $B\colon \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}$ une application bilinéaire symétrique. Alors il existe une base orthonormée $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ de $\mathbf{R}^m$ telle que*

$$B(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0 \qquad \text{pour tous } i \text{ et } j \text{ tels que } i \neq j.$$

Toute démonstration habituelle de ce théorème utilise le calcul différentiel (en plus de l'algèbre linéaire).

Corollaire. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit X un point à l'intérieur de A où f est deux fois différentiable. Alors il existe une base orthonormée $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ de $\mathbf{R}^m$ telle que*

$$D_{\boldsymbol{u}_i} D_{\boldsymbol{u}_j} f(X) = f''(X)(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0 \qquad \text{pour tous } i \text{ et } j \text{ tels que } i \neq j.$$

Dans les hypothèses de ce corollaire, soit $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ une base de $\mathbf{R}^m$ comme dans sa conclusion. Pour tout $i = 1, \dots, m$, posons

$$\lambda_i = D^2_{\boldsymbol{u}_i} f(X).$$

Soit

$$\boldsymbol{v} = \alpha_1 \boldsymbol{u}_1 + \dots + \alpha_m \boldsymbol{u}_m$$

un vecteur arbitraire dans $\mathbf{R}^m$. Alors

$$D^2_{\boldsymbol{v}} f = \alpha_1^2 D^2_{\boldsymbol{u}_1} f(X) + \dots + \alpha_m^2 D^2_{\boldsymbol{u}_m} f(X) = \lambda_1 \alpha_1^2 + \dots + \lambda_m \alpha_m^2.$$

[...]

Définition. Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit X un point à l'intérieur de A où f est deux fois différentiable. Alors définissons la *matrice hessienne* de f en X comme la matrice de l'application bilinéaire $f''(X)\colon \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}$ par rapport à la base canonique de $\mathbf{R}^m$.

Notation. La matrice hessienne f en X sera notée « $H_f(X)$ ».

Ainsi, $H_f(X)$ est définie par l'identité :

$$f''(X)(\boldsymbol{u}, \boldsymbol{v}) = {}^{\mathbf{t}}\!\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} H_f(X) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} u_1 & \cdots & u_m \end{pmatrix} H_f(X) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix},$$

où $\boldsymbol{u} = (u_1, \dots u_m)$ et $\boldsymbol{v} = (v_1, \dots v_m)$. En particulier,

$$D^2_{\boldsymbol{v}} f(X) = {}^{\mathbf{t}}\!\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} H_f(X) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} v_1 & \cdots & v_m \end{pmatrix} H_f(X) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix},$$

où $\boldsymbol{v} = (v_1, \dots v_m)$.

Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ la base canonique de $\mathbf{R}^m$. Alors, pour tout $X \in A$ tel que f est deux fois différentiable en X, on a :

$$\begin{aligned} H_f(X) &= \begin{pmatrix} f_1''(X)(\boldsymbol{e}_1, \boldsymbol{e}_1) & \cdots & f_1''(X)(\boldsymbol{e}_1, \boldsymbol{e}_m) \\ \vdots & \ddots & \vdots \\ f_n''(X)(\boldsymbol{e}_m, \boldsymbol{e}_1) & \cdots & f_n''(X)(\boldsymbol{e}_m, \boldsymbol{e}_m) \end{pmatrix} \\ &= \begin{pmatrix} D_{\boldsymbol{e}_1} D_{\boldsymbol{e}_1} f_1(X) & \cdots & D_{\boldsymbol{e}_1} D_{\boldsymbol{e}_m} f_1(X) \\ \vdots & \ddots & \vdots \\ D_{\boldsymbol{e}_m} D_{\boldsymbol{e}_1} f_n(X) & \cdots & D_{\boldsymbol{e}_m} D_{\boldsymbol{e}_m} f_n(X) \end{pmatrix}. \end{aligned}$$

En termes des dérivées partielles :

$$H_f(X) = \begin{pmatrix} D_1 D_1 f(X) & \cdots & D_1 D_m f(X) \\ \vdots & \ddots & \vdots \\ D_m D_1 f(X) & \cdots & D_m D_m f(X) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 f(X)}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f(X)}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(X)}{\partial x_m \partial x_1} & \cdots & \dfrac{\partial^2 f(X)}{\partial x_m \partial x_m} \end{pmatrix},$$

où $X = (x_1, \dots, x_m)$.

II.10. Formule de Taylor-Young

Théorème. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}^n$. Soit X un point à l'intérieur de A. Supposons que f est K fois différentiable en X, $K \geqslant 1$. Alors*

$$f(X + \boldsymbol{h}) = \sum_{k=0}^{K} \frac{D^k_{\boldsymbol{h}} f(X)}{k!} + o(\|\boldsymbol{h}\|^K) \qquad \text{lorsque } \boldsymbol{h} \to \mathbf{0}_{\mathbf{R}^m}.$$

II.11. Étude des extrema

Proposition. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit X un point à l'intérieur de A tel que f est différentiable en X. Si f atteint en X un extremum local (strict ou pas), alors $D_{\boldsymbol{h}} f(X) = 0$ pour tout $\boldsymbol{h} \in \mathbf{R}^m$ (autrement dit, $f'(X)$ est l'application linéaire nulle).*

Exercice. Prouver cette proposition.

La dernière proposition peut être facilement déduite du théorème suivant.

Théorème. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit X un point à l'intérieur de A. Supposons que f est K fois différentiable en X, $K \geqslant 1$. Définissons la fonction $\phi\colon \mathbf{R}^m \to \mathbf{R}$ par la formule :*

$$\phi(\boldsymbol{h}) = \sum_{k=0}^{K} \frac{D^k_{\boldsymbol{h}} f(X)}{k!}.$$

(1) *Si f atteint en X un maximum local (strict ou pas), alors ϕ atteint en $\mathbf{0}_{\mathbf{R}^m}$ un maximum local (strict ou pas).*

(2) *Il en est de même pour un minimum (au lieu d'un maximum).*

(3) *Si ϕ atteint en $\mathbf{0}_{\mathbf{R}^m}$ un maximum local strict, alors f atteint en X un maximum local strict.*

(4) *Il en est de même pour un minimum (au lieu d'un maximum).*

Esquisse d'une démonstration. Appliquer la formule de Taylor-Young. □

La proposition précédente est un corollaire facile de ce théorème. En voici un autre.

Corollaire. *Soient $A \subset \mathbf{R}^m$ et $f\colon A \to \mathbf{R}$. Soit X un point à l'intérieur de A tel que f est 2 fois différentiable en X et que $D_{\boldsymbol{h}} f(X) = 0$ pour tout $\boldsymbol{h} \in \mathbf{R}^m$.*

(1) *Si $D^2_{\boldsymbol{h}} f(X) > 0$ pour tout $\boldsymbol{h} \in \mathbf{R}^m \setminus \{\mathbf{0}_{\mathbf{R}^m}\}$, alors f atteint en X un minimum local strict.*

(2) *Si $D^2_{\boldsymbol{h}} f(X) < 0$ pour tout $\boldsymbol{h} \in \mathbf{R}^m \setminus \{\mathbf{0}_{\mathbf{R}^m}\}$, alors f atteint en X un maximum local strict.*

(3) *S'il existe $\boldsymbol{h} \in \mathbf{R}^m$ tel que $D^2_{\boldsymbol{h}} f(X) > 0$ et qu'il existe $\boldsymbol{h} \in \mathbf{R}^m$ tel que $D^2_{\boldsymbol{h}} f(X) < 0$, alors f n'a pas d'extremum en X.*

II.12. Différentielle totale

SECTION-BROUILLON
[...]

$$(\mathrm{d}f)(X) = f'(X).$$

[...]

$$d(f(X)) = (\mathrm{d}f)(X)(dX) = f'(X)(dX).$$

[...]

II.13. Différentielles partielles

SECTION-BROUILLON
[...]

$$(\mathrm{d}f)(X)(\boldsymbol{v}) = (\partial_1 f)(X)(v_1) + \cdots + (\partial_m f)(X)(v_m),$$

où $\boldsymbol{v} = (v_1, \ldots, v_m)$. Ici,

$$(\partial_1 f)(X)(v_1) = D_1 f(X) v_1, \qquad \ldots, \qquad (\partial_m f)(X)(v_m) = D_m f(X) v_m.$$

[...]

$$\partial_{x_1}(f(X)) = (\partial_1 f)(X)(dx_1) = \frac{\partial_{x_1}(f(X))}{\partial_{x_1} x_1} dx_1 = \frac{\partial(f(X))}{\partial x_1} dx_1,$$
$$\ldots,$$
$$\partial_{x_m}(f(X)) = (\partial_m f)(X)(dx_m) = \frac{\partial_{x_m}(f(X))}{\partial_{x_m} x_m} dx_m = \frac{\partial(f(X))}{\partial x_m} dx_m,$$

où $X = (x_1, \ldots, x_m)$, $dX = (dx_1, \ldots, dx_m)$.
[...]

II.14. Différentielles d'ordres supérieurs

SECTION-BROUILLON
[...]
Si f est 2 fois différentiable en X,

$$(\mathrm{d}^2 f)(X)(\boldsymbol{v}) = f''(X)(\boldsymbol{v}, \boldsymbol{v}) = D^2_{\boldsymbol{v}} f(X).$$

[...]

$$d^2(f(X)) = (\mathrm{d}^2 f)(X)(dX) = f''(X)(dX, dX) = D^2_{dX} f(X).$$

[...]