Analyse Mathématique pour l'Ingénieur

Paris, 2009 - 2016

Cours 06

Introduction au calcul différentiel

François Dubois

Analyse Mathématique pour l'Ingénieur

Cours 6
Introduction au calcul différentiel *

a) Fonctions de deux variables réelles

- Continuité
- Dérivées partielles
- Application linéaire tangente
- Différentielle d'une application composée

b) Compléments de calcul différentiel

- Matrices jacobiennes
- Dérivée seconde et matrice Hessienne
- Forme bilinéaire symétrique
- Conditions d'extremum
- Formule de Taylor

* François Dubois, 2009, édition septembre 2015, 34 pages.

AMI

ch(5) Fonctions de deux variables réelles

• Continuité

On considère une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ (f est à valeurs numériques, $f(x) \in \mathbb{R}$) définie "au voisinage" de $x_0 \in \mathbb{R}^2$; il existe donc un voisinage $V(x_0)$ du point x_0, c'est à dire un sous ensemble de $\mathbb{R}^2$ qui contient une "boule ouverte" centrée en x_0. Nous notons $|x|$ la norme euclidienne de $x \in \mathbb{R}^2$:

(1) $$|x| = \sqrt{x_1^2 + x_2^2}, \qquad x = (x_1, x_2) \in \mathbb{R}^2,$$

$B(x_0, r)$ la boule ouverte de centre x_0 et rayon r :

(2) $$B(x_0, r) = \{x \in \mathbb{R}^2, \ |x - x_0| < r\}, \quad x_0 \in \mathbb{R}^2, r > 0, .$$

* Un voisinage $V(x_0)$ contient (par définition) une telle boule ouverte

(3) $$\exists r > 0, \ B(x_0, r) \subset V(x_0).$$

Nous supposons que f est une application définie de $V(x_0)$ et à valeurs dans $\mathbb{R}$:

(4) $\quad \mathbb{R}^2 \supset V(x_0) \ni x \longmapsto f(x) \in \mathbb{R}$

o Attention aux notations !

La fonction f pourra être notée comme en (4) mais aussi sous la forme $V(x_0) \ni (x,y) \mapsto f(x,y) \in \mathbb{R}$ ou bien $V(x_0) \ni (x_1, x_2) \mapsto f(x_1, x_2) \in \mathbb{R}$, la lettre "$x$" désignant tantôt un point de $\mathbb{R}^2$ et tantôt un nombre réel. Le contexte permettra de trancher.

(def) Continuité en un point.

Soit f définie au voisinage de $x_0 \in \mathbb{R}^2$ et à valeurs réelles. On dit que f est continue en x_0 si et seulement si

(5) $\forall \varepsilon > 0, \ \exists \eta > 0, \ |x - x_0| < \eta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$

• exemples.

Les fonctions polynomiales, composées de ces fonctions avec les fonctions transcendantes classiques telles que l'exponentielle et les fonctions trigonométriques, etc. Parmi celles-ci les coordonnées "x" et "y", avec par définition ici

$$(6) \begin{cases} \mathbb{R}^2 \ni (x_1, x_2) \longmapsto x(x_1, x_2) = x_1 \in \mathbb{R} \\ \mathbb{R}^2 \ni (x_1, x_2) \longmapsto y(x_1, x_2) = x_2 \in \mathbb{R} \end{cases}$$

On note que l'application "distance" est continue (ie $\mathbb{R}^2 \ni x \mapsto |x| \in \mathbb{R}$), grâce à l'inégalité triangulaire sous la forme

(7) $\big|\,|x|-|y|\,\big| \leq |x-y|, \quad x, y \in \mathbb{R}^2$.

- Contre exemple (dû à Laurent Schwartz, Cours d'Analyse (tome 1), Hermann, Paris, 1967, page 196)

$$(8) \quad f(x,y) = \begin{cases} \dfrac{x^5}{(y-x^2)^2 + x^8}, & (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0 \end{cases}$$

En se plaçant sur la parabole d'équation $y = x^2$, on a pour $x \neq 0$ $f(x, x^2) = 1/x^3$ qui ne tend pas vers $f(0,0) = 0$ si x tend vers 0. La fonction f n'est pas continue en $x_0 = (0,0)$. Pourtant, elle a d'autres propriétés étudiées plus loin.

• Dérivées partielles

L'idée est de fixer une des variables et de considérer la fonction d'une seule variable ainsi obtenue, qui est une classique fonction numérique à valeurs réelles (cf le chapitre précédent!)

On pose

$f_1(x) = f(x, x_2)$ (x_2 est fixé, $x \in \mathbb{R}$) et
$f_2(y) = f(x_1, y)$ (x_1 fixé, y varie dans $\mathbb{R}$).
On dérive (on suppose qu'on peut le faire !)
f_1 par rapport à x ; le résultat obtenu est par définition la <u>dérivée partielle</u> $\frac{\partial f}{\partial x_1}$ par rapport à la première variable. De même si on dérive f_2 par rapport à y ; on définit de cette façon la dérivée partielle $\frac{\partial f}{\partial x_2}$ par rapport à la seconde variable :

$$(9) \begin{cases} \dfrac{\partial f}{\partial x_1}(x_1, x_2) = \left(\dfrac{d}{dx} f_1(x)\right)_{x = x_1} \\ \qquad = \lim\limits_{\theta \to 0} \dfrac{1}{\theta}\left(f(x_1+\theta, x_2) - f(x_1, x_2)\right) \end{cases}$$

$$(10) \begin{cases} \dfrac{\partial f}{\partial x_2}(x_1, x_2) = \left(\dfrac{d}{dy} f_2(y)\right)_{y = x_2} \\ \qquad = \lim\limits_{\theta \to 0} \dfrac{1}{\theta}\left(f(x_1, x_2+\theta) - f(x_1, x_2)\right) \end{cases}$$

- Une fonction de deux variables réelles peut admettre deux dérivées partielles $\frac{\partial f}{\partial x}$ (ou $\frac{\partial f}{\partial x_1}$) et $\frac{\partial f}{\partial y}$ (ou $\frac{\partial f}{\partial x_2}$) en un point $x_0 \in \mathbb{R}^2$
et ne <u>pas</u> être continue en ce point !

* Par exemple l'exemple proposé en (8). On a facilement $f(x, 0) = \dfrac{x}{1+x^4}$ si $x \neq 0$, donc $\frac{\partial f}{\partial x}(0,0) = 1$ et f admet une

dérivée partielle par rapport à la première variable en $x_0 = (0,0)$. Par ailleurs, $f(0,y) = 0$ si $y \neq 0$. Donc $\frac{\partial f}{\partial y}(0,0) = 0$ et f admet aussi une dérivée partielle par rapport à la seconde variable en $x_0 = (0,0)$. Mais f n'est pas pour autant continue en x_0, comme nous l'avons vu plus haut ! Le lien entre dérivation et continuité est donc plus délicat pour les fonctions de deux variables que pour les fonctions d'une seule variable.

- Application linéaire tangente.

On se donne $x_0 = (x_1, x_2) \in \mathbb{R}^2$ et un "petit" vecteur $h = (h_1, h_2) \in \mathbb{R}^2$. On a le calcul suivant

$$
\begin{aligned}
f(x_0+h) &= f(x_1+h_1,\ x_2+h_2) \\
&= f(x_0) + \big(f(x_1+h_1,\ x_2) - f(x_0)\big) \\
&\qquad + \big(f(x_1+h_1,\ x_2+h_2) - f(x_1+h_1,\ x_2)\big) \\
&= f(x_0) + \frac{\partial f}{\partial x}(x_1, x_2)\, h_1 + h_1\, \varepsilon_1(h_1; x_1, x_2) \\
&\qquad + \frac{\partial f}{\partial y}(x_1+h_1, x_2)\, h_2 + h_2\, \varepsilon_2(h_2; x_1, x_2, h_1)
\end{aligned}
$$

en supposant f avec des dérivées partielles $\frac{\partial f}{\partial x}$

et $\frac{\partial f}{\partial y}$ au voisinage de x_0. Dans la relation précédente $\varepsilon_1(h_1; x_1, x_2)$ tend vers 0 si h_1 tend vers 0 et $\varepsilon_2(h_2; x_1, x_2, h_1)$ tend vers 0 si h_2 tend vers 0.

* On fait l'hypothèse supplémentaire que $\frac{\partial f}{\partial y}$ est <u>continue</u> au voisinage de x_0. On a donc :

$$\frac{\partial f}{\partial y}(x_1+h_1, x_2) = \frac{\partial f}{\partial y}(x_1, x_2) + \varepsilon_3(h_1; x_1, x_2)$$

En injectant cette expression dans le développement précédent, on trouve

$$(11)\quad f(x_0+h) = f(x_0) + \frac{\partial f}{\partial x}(x_0)\,h_1 + \frac{\partial f}{\partial y}(x_0)\,h_2 + |h|\,\varepsilon(h)$$

avec $\varepsilon(h)$ qui dépend aussi de (x_1, x_2) et tend vers 0 pour $|h|$ tendant vers 0.

* Le développement (11) met en évidence deux nombres réels, les dérivées partielles $\frac{\partial f}{\partial x}(x_0) \in \mathbb{R}$ et $\frac{\partial f}{\partial y}(x_0) \in \mathbb{R}$ et l'<u>application linéaire</u> de $\mathbb{R}^2$ dans $\mathbb{R}$ notée $df(x_0)$:

$$(12)\qquad \mathbb{R}^2 \ni h = (h_1, h_2) \mapsto df(x_0)\cdot h = \frac{\partial f}{\partial x} h_1 + \frac{\partial f}{\partial y} h_2 \in \mathbb{R}$$

est appelée <u>application linéaire tangente</u> à f au point x_0.

- On peut représenter l'application linéaire tangente $h \mapsto df(x_0).h$ à l'aide d'une <u>matrice ligne</u>, notée traditionnellement $J_f(x_0)$ et appelée <u>matrice jacobienne de f</u> au point x_0 :

(13) $$J_f(x_0) = \left(\frac{\partial f}{\partial x}(x_0) \quad \frac{\partial f}{\partial y}(x_0) \right) \in \mathcal{M}_{1,2}(\mathbb{R})$$

qui est une matrice à une ligne et deux colonnes. Si $h \in \mathbb{R}^2$, on pose

(14) $$h^t = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \quad h = (h_1, h_2) \in \mathbb{R}^2$$

la matrice colonne correspondante, et on a, compte tenu des relations (12) et (13) :

(15) $$df(x_0).h = J_f(x_0).h^t, \quad h \in \mathbb{R}^2.$$

on note que $df(x_0).h \in \mathbb{R}$ est l'image de $h \in \mathbb{R}^2$ par l'application linéaire tangente $df(x_0) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$. La matrice de $df(x_0)$ relativement à la base canonique de $\mathbb{R}^2$ est tout simplement la matrice jacobienne $J_f(x_0)$.

def <u>Application différentiable en un point.</u>

Soit f définie au voisinage de x_0 (i.e. $\mathbb{R}^2 \supset V(x_0) \ni x \mapsto f(x) \in \mathbb{R}$). On dit que f est différentiable en x_0 si et seulement si il existe $df(x_0) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ et $V(0) \ni h \mapsto \varepsilon(h) \in \mathbb{R}$ continue en 0 et telle que $\varepsilon(0) = 0$ (donc $\varepsilon(h)$ tend vers 0 si h tend vers 0) de sorte que

$$(16) \quad f(x_0 + h) = f(x_0) + df(x_0) \cdot h + |h| \, \varepsilon(h)$$

(Th) <u>Différentiabilité implique continuité.</u>

Si f est différentiable en x_0, alors elle est continue en ce point.

- la <u>preuve</u> est claire. La différence $f(x_0 + h) - f(x_0)$ vaut $df(x_0) \cdot h + |h| \, \varepsilon(h)$ si $f(\cdot)$ est différentiable en x_0; cette expression tend vers 0 si h tend vers 0 dans $\mathbb{R}^2$. Donc f est continue au point x_0.
- La réciproque <u>est fausse</u>, comme le montre l'exemple $f(x, y) = \sqrt{x^2 + y^2}$ $(x, y \in \mathbb{R})$ qui est continue en 0 (cf (7)) mais n'est pas différentiable puisque l'application partielle $\mathbb{R} \ni x \mapsto f(x, 0) = |x| \in \mathbb{R}$ n'est pas dérivable en $x_0 = 0$.

(Prop) Si f est différentiable en x_0, elle a des dérivées partielles $\frac{\partial f}{\partial x}(x_0)$ et $\frac{\partial f}{\partial y}(x_0)$. De plus, $df(x_0) \cdot h$ peut être calculé par l'expression (12).

- Calculons la différentielle de l'application "x" (de première composante) définie à la relation (6). On a clairement $dx(x_1,x_2)\cdot h = h_1$ et de même, $dy(x_1,x_2)\cdot h = h_2$. Nous pouvons donc réécrire la relation (12) sous la forme

$$df(x_0)\cdot h = \frac{\partial f}{\partial x}(x_0)\left(dx(x_0)\cdot h\right) + \frac{\partial f}{\partial y}(x_0)\left(dy(x_0)\cdot h\right)$$
$$= \left(\frac{\partial f}{\partial x}(x_0)\,dx(x_0) + \frac{\partial f}{\partial y}(x_0)\,dy(x_0)\right)\cdot h .$$

Cette relation est vraie $\forall h \in \mathbb{R}^2$; nous avons donc l'égalité suivante, entre <u>formes linéaires</u> :

$$df(x_0) = \frac{\partial f}{\partial x}(x_0)\,dx + \frac{\partial f}{\partial y_0}(x_0)\,dy$$

puisque $dx(x_0)$ et $dy(x_0)$ ne dépendent pas du point $x_0 \in \mathbb{R}^2$ de référence. Si on omet maintenant le point où l'on est en train de dériver, on peut écrire aussi

$$(17) \quad df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy$$

ce qui constitue une façon d'écrire tout à fait classique.

- <u>Exemple</u> : coordonnées polaires dans le plan.

Si on pose $x = r\cos\theta$ et $y = r\sin\theta$ avec $\theta \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[$ pour fixer les idées, on a $r = \sqrt{x^2+y^2}$ et $\theta = \operatorname{arctg}(y/x)$. De $r^2 = x^2+y^2$ et $\operatorname{tg}\theta = \frac{y}{x}$, on déduit sans difficulté

$$(18)\quad \begin{cases} dr = \dfrac{x}{r}\,dx + \dfrac{y}{r}\,dy \\ d\theta = -\dfrac{y}{r^2}\,dx + \dfrac{x}{r^2}\,dy . \end{cases}$$

def Dérivée selon un vecteur

Soit $f : V(x_0) \to \mathbb{R}$ défini au voisinage de $x_0 \in \mathbb{R}^2$ et $v \in \mathbb{R}^2$ un vecteur <u>non nul</u> de $\mathbb{R}^2$. On dit que f est dérivable en x_0 dans la direction v si et seulement si l'application $\mathbb{R} \ni \theta \mapsto \varphi(\theta) = f(x_0 + \theta v) \in \mathbb{R}$ est dérivable en $\theta = 0$ (comme fonction de $\mathbb{R}$ dans $\mathbb{R}$)

On voit sans difficulté que pour $v = e_1$ (premier vecteur de la base canonique de $\mathbb{R}^2$), on a

$$(19)\quad df(x_0) \cdot e_1 = \frac{\partial f}{\partial x_0}(x_0) \; ; \; df(x_0) \cdot e_2 = \frac{\partial f}{\partial y}(x_0)$$

et une relation analogue pour $v = e_2$.

Prop Si f est différentiable en x_0, alors elle est dérivable selon tout vecteur $v \neq 0$ et

$$(20)\quad \left(\frac{d}{d\theta} \big(f(x_0 + \theta v) \big) \right)_{\theta = 0} = df(x_0) \cdot v .$$

- La <u>preuve</u> consiste à revenir à la définition (16) de la différentiabilité, en prenant $h = \theta v$ avec $\theta \in \mathbb{R}$ dans un voisinage de 0. Il vient

$$f(x_0+\theta v) = f(x_0) + df(x_0)\cdot(\theta v) + |\theta v|\,\varepsilon(\theta v)$$
$$= f(x_0) + \theta\, df(x_0)\cdot v + |\theta|\,\tilde{\varepsilon}(\theta)$$
car $df(x_0)$ est linéaire et $v \in \mathbb{R}^2$ est fixé.
On en déduit, avec la notation $\varphi(\theta) = f(x_0+\theta v)$:
$$\varphi(\theta) = \varphi(0) + \theta\, df(x_0)\cdot v + |\theta|\,\tilde{\varepsilon}(\theta),$$
ce qui montre que φ est différentiable en $\theta = 0$, et que $\varphi'(0) = df(x_0)\cdot v$. La propriété est établie. $\square$

- Attention, la réciproque est fausse. Une fonction peut avoir une dérivée partielle selon <u>tout</u> vecteur non nul v sans pour autant être différentiable au point considéré ! C'est le cas en particulier pour la fonction $f(x,y)$ introduite à la relation (8). Vérifions qu'elle possède une dérivée partielle en $x_0 = 0$ selon tout vecteur $v \equiv (X,Y) \neq (0,0)$.

1) Si $Y \neq 0$, alors
$$\frac{1}{\theta}\big(f(\theta X, \theta Y) - f(0)\big) = \frac{1}{\theta}\,\frac{\theta^5 X^5}{\theta^2(Y-\theta X^2)^2 + \theta^8 X^8}$$
$$= \theta^2\,\frac{X^5}{(Y-\theta X^2)^2 + \theta^6 X^8}$$
qui tend vers 0 si $\theta \to 0$. On a donc $\varphi'(0) = 0$ avec $\varphi(\theta) \equiv f(\theta X, \theta Y)$ et f introduite à la relation (8).

2) Si $Y=0$ et $X\neq 0$, on a

$$\frac{1}{\theta}\left(f(\theta X,0)-f(0)\right)=\frac{(\theta X)^5}{\theta\left[(\theta X)^4+(\theta X)^8\right]}=\frac{X}{1+\theta^4X^4}$$

qui tend vers X si θ tend vers 0. Donc $\varphi'(0)=X$ avec $\varphi(\theta)=f(\theta X,0)$, $X\neq 0$. Par contre la fonction $f(\cdot)$ définie en (8) est clairement non différentiable en $x_0=0$ puisqu'elle n'y est même pas continue !

(Th) Continuité des dérivées partielles implique la différentiabilité.

Soit f définie au voisinage de x_0. On suppose les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ exister au voisinage de x_0 et on les suppose aussi <u>continues</u> au point x_0. Alors f est différentiable en x_0 et $df(x_0)\cdot h$ est donné par l'expression (12) en fonction des dérivées partielles $\frac{\partial f}{\partial x}(x_0)$ et $\frac{\partial f}{\partial y}(x_0)$ au point x_0.

- La <u>preuve</u> a été essentiellement présentée aux pages 5 et 6 de ce document pour introduire l'application linéaire tangente. □

• Différentielle d'une application composée

Afin de ne pas avoir à traiter de multiples cas particuliers, nous abandonnons dans ce paragraphe le cas $f : \mathbb{R}^2 \to \mathbb{R}$ pour présenter le cas général d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}^m$.

(dif) Différentiabilité

Soit $x_0 \in \mathbb{R}^n$, $V(x_0)$ un voisinage de x_0 ($V(x_0)$ contient une boule ouverte centrée en x_0), $f : V(x_0) \to \mathbb{R}^m$ une application définie au voisinage de x_0. Cette application est différentiable en x_0 si et seulement si il existe une application linéaire tangente $df(x_0)$, linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ ($df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$) et une application ε définie au voisinage de 0 (dans $\mathbb{R}^n$) et à valeurs dans $\mathbb{R}^m$ telle que $\varepsilon(0) = 0$ et ε continue en $0 \in \mathbb{R}^n$ de sorte que

(21) $f(x_0+h) = f(x_0) + df(x_0).h + \|h\| \, \varepsilon(h)$,

où $\|h\|$ désigne la norme euclidienne de $h \in \mathbb{R}^n$.

- L'application $df(x_0)$ est représentée par la matrice jacobienne $J_f(x_0)$ à m lignes et n colonnes ($J_f(x_0) \in \mathcal{M}_{m,n}(\mathbb{R})$) de sorte que

$$(22) \qquad (df(x_0) \cdot h)^t = J_f(x_0) \cdot h^t \quad , h \in \mathbb{R}^n$$

en respectant scrupuleusement les vecteurs lignes (de $\mathbb{R}^n$) et les vecteurs colonnes (de $\mathbb{R}^{tn}$). La matrice $J_f(x_0)$ a m lignes et n colonnes ; son élément générique (i,j) ($1 \leq i \leq m$; $1 \leq j \leq n$) est donné par

$$(23) \qquad (J_f(x_0))_{ij} = \frac{\partial f_i}{\partial x_j}(x_0), \quad x_0 \in \mathbb{R}^n, \quad 1 \leq i \leq m, \ 1 \leq j \leq n.$$

* On écrit le vecteur $f(x) \in \mathbb{R}^m$ à l'aide de ses composantes $f_i(x) \in \mathbb{R}$ et de la base canonique e_i de $\mathbb{R}^m$:

$$(24) \qquad f = \sum_{i=1}^{m} f_i \, e_i \, , \quad e_i \in \mathbb{R}^m, \ f_i(x) \in \mathbb{R}.$$

Le nombre $\frac{\partial f_i}{\partial x_j}(x_0)$ est la j^o dérivée partielle de la i^o composante du vecteur $f_i(x_0)$.
On peut aussi écrire la relation (23) sous la forme

$$J_f(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_j} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_i}{\partial x_1} & \cdots & \frac{\partial f_i}{\partial x_j} & \cdots & \frac{\partial f_i}{\partial x_n} \\ \vdots & & & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_j} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

- Dans l'exemple où $n = 2$ et $m = 1$, la matrice jacobienne est donnée par la matrice ligne décrite à la relation (13).

(th) Différentiabilité d'une application composée

Soit n, m, p trois entiers ≥ 1, $x_0 \in \mathbb{R}^n$, $V(x_0)$ un voisinage de x_0, $f : V(x_0) \to \mathbb{R}^m$, $y_0 = f(x_0)$, $W(y_0)$ un voisinage de y_0 de sorte que $f(V(x_0)) \subset W(y_0)$, $g : W(y_0) \to \mathbb{R}^p$ définie au voisinage de y_0. Alors l'application $g \circ f : V(x_0) \to \mathbb{R}^p$ est bien définie pour $x \in V(x_0)$ par la relation $(g \circ f)(x) = g(f(x))$.

On suppose f différentiable au point x_0 et g différentiable en y_0. On note $df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ et $dg(y_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$ les applications linéaires tangentes correspondantes. Alors $g \circ f$ est différentiable au point x_0 et l'on a

$$(25) \quad d(g\circ f)(x_0) \cdot h = dg(y_0) \cdot df(x_0) \cdot h, \quad h \in \mathbb{R}^n.$$

- Avant la preuve, on vérifie que les êtres mathématiques introduits de part et d'autre de l'égalité existent bel et bien. En effet, pour $h \in \mathbb{R}^n$, $df(x_0) \cdot h \in \mathbb{R}^m$. Par ailleurs, si $k \in \mathbb{R}^m$, $dg(y_0) \cdot k \in \mathbb{R}^p$ et si on pose $k = df(x_0) \cdot h$, le produit $dg(y_0) \cdot df(x_0) \cdot h$ vaut simplement $dg(y_0) \cdot k \in \mathbb{R}^p$. Par ailleurs, $g \circ f$ est définie au voisinage de $x_0 \in \mathbb{R}^n$ et est à valeurs dans $\mathbb{R}^p$, donc $d(g\circ f)(x_0)$ est a priori une application linéaire de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Donc $d(g\circ f)(x_0) \cdot h \in \mathbb{R}^p$ et on tient un bout de cohérence. On peut aussi voir le membre de droite de (25) sous la forme $(dg(y_0) \circ df(x_0)) \cdot h$, où l'on compose $df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ avec $dg(y_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$, la composée de ces deux différentielles ayant bien un sens. La relation (25) peut alors s'exprimer sans l'emploi de $h \in \mathbb{R}^n$, avec l'égalité entre différentielles suivante

$$(26) \qquad d(g \circ f)(x_0) = dg(y_0) \circ df(x_0), \quad y_0 = f(x_0)$$

et la différentielle de la composée en x_0 est égale à la composée des différentielles. On peut aussi omettre la référence à x_0 et on a

$$(27) \qquad d(g \circ f) = (dg \circ f) \circ df$$

égalité entre champs d'applications linéaires.

- La preuve de la relation (25) consiste à écrire (21) puis l'analogue pour g :

$$g(y_0 + k) = g(y_0) + dg(y_0) \cdot k + \|k\| \, \varepsilon_g(k)$$

avec $k \in \mathbb{R}^m$ et $y_0 = f(x_0)$. Alors avec $k = df(x_0) \cdot h + \|h\| \varepsilon(h)$ donné par (21), soit $k = f(x_0+h) - f(x_0)$, il vient

$$\begin{aligned} g(f(x_0+h)) = g(y_0+k) &= g(y_0) + dg(y_0) \cdot k + \|k\| \varepsilon_g(k) \\ &= g(y_0) + dg(y_0) \cdot [df(x_0) \cdot h + \|h\| \varepsilon(h)] + \|k\| \varepsilon_g(k) \\ &= g(y_0) + dg(y_0) \cdot df(x_0) \cdot h + \|h\| \widetilde{\varepsilon}(h), \text{ avec} \end{aligned}$$

$\widetilde{\varepsilon}(h) = dg(y_0) \cdot \varepsilon(h) + \left(df(x_0) \cdot \frac{h}{\|h\|} + \varepsilon(h) \right) \varepsilon_g\big(df(x_0) \cdot h + \|h\| \varepsilon(h) \big)$ qui tend clairement vers 0 si $h \to 0$ dans $\mathbb{R}^n$. La relation (25) en résulte et la proposition est établie. $\square$

Dubois 25 nov 2009

Corrections de coquilles, 5 juin 2011.

AMI

ch 6 Compléments de calcul différentiel

• Matrices jacobiennes

Nous avons vu que si $f: \mathbb{R}^n \to \mathbb{R}^m$ est définie au voisinage de x_0 ($x_0 \in \mathbb{R}^n$), si on pose $y_0 = f(x_0)$ (donc $y_0 \in \mathbb{R}^m$), si $g: \mathbb{R}^m \to \mathbb{R}^p$ est définie au voisinage de y_0, alors $g \circ f: \mathbb{R}^n \to \mathbb{R}^p$ est définie au voisinage de x_0. Si f est différentiable en x_0 et g différentiable en y_0, alors $g \circ f$ est différentiable en x_0 et on a

$$(1) \quad d(g \circ f)(x_0) = dg(y_0) \circ df(x_0)$$

et cette égalité a lieu dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$.

• Introduisons la matrice jacobienne $J_f(x_0)$ à n colonnes et m lignes ($J_f(x_0) \in \mathcal{M}_{m,n}(\mathbb{R})$) et la matrice jacobienne $J_g(y_0)$ à m colonnes et p lignes ($J_g(y_0) \in \mathcal{M}_{p,m}(\mathbb{R})$) grâce aux relations

$$(2) \begin{cases} (df(x_0)\cdot h)^t = J_f(x_0)\cdot h^t, & h \in \mathbb{R}^n \\ (dg(y_0)\cdot k)^t = J_g(y_0)\cdot k^t, & k \in \mathbb{R}^m. \end{cases}$$

La relation (1) permet de calculer la matrice jacobienne $J_{g\circ f}(x_0)$ de la composée $g\circ f$ au point x_0. On a d'une part

$$\left(\left(d(g\circ f)\right)(x_0)\cdot h\right)^t = J_{g\circ f}(x_0)\cdot h^t$$

par définition même d'une matrice jacobienne et d'autre part, compte tenu de (2) :

$$\begin{aligned}(dg(y_0)\cdot df(x_0)\cdot h)^t &= J_g(y_0)\cdot (df(x_0)\cdot h)^t \\ &= J_g(y_0)\cdot (J_f(x_0)\cdot h^t) \\ &= (J_g(y_0)\cdot J_f(x_0))\cdot h^t .\end{aligned}$$

L'égalité de ces deux expressions (compte tenu de la relation (1)) montre que l'on a

$$(3) \quad J_{g\circ f}(x_0) = J_g(y_0)\cdot J_f(x_0) \ ; \ y_0 = f(x_0).$$

- Rappelons aussi que l'élément générique "i,j" (i indice de ligne, $1\leq i\leq m$ et j indice de colonne, $1\leq j\leq n$) de la matrice $J_f(x_0)$ vaut la dérivée partielle $\frac{\partial f_i}{\partial x_j}(x_0)$ de la i° composante de la fonction f :

$$(4) \quad \left(J_f(x_0)\right)_{ij} = \frac{\partial f_i}{\partial x_j}(x_0) \quad , \ 1\leq i\leq m,\ 1\leq j\leq n .$$

• Nous pouvons donner un (double) cas particulier à partir de $f : \mathbb{R}^2 \to \mathbb{R}$ et $\varphi : \mathbb{R} \to \mathbb{R}^2$; la fonction f est une fonction numérique de deux variables et φ une courbe paramétrée de composantes $X(t)$, $Y(t)$:

(5) $\quad \varphi(t) = (X(t), Y(t)) \in \mathbb{R}^2, \; t \in \mathbb{R}$.

* on peut d'une part former la composée $g = f \circ \varphi$. alors g est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ (la valeur de la fonction numérique f sur la courbe φ) et la relation (3) peut s'écrire dans ce cas :

$$g'(t) = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} dX/dt \\ dY/dt \end{pmatrix} = \frac{\partial f}{\partial x}\frac{dX}{dt} + \frac{\partial f}{\partial y}\frac{dY}{dt}.$$

on peut ne pas utiliser les matrices jacobiennes et écrire formellement

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy, \quad dx = \frac{dX}{dt} dt, \; dy = \frac{dY}{dt} dt$$

ce qui conduit à $\quad df = \left(\frac{\partial f}{\partial x}\frac{dX}{dt} + \frac{\partial f}{\partial y}\frac{dY}{dt}\right) dt$ sur la courbe $(x, y) = \varphi(t)$, soit

(6) $$\frac{d}{dt} f(\varphi(t)) = \frac{\partial f}{\partial x}\frac{dX}{dt} + \frac{\partial f}{\partial y}\frac{dY}{dt}.$$

* on peut aussi former la seconde composée $\beta(x,y) = \varphi(f(x,y))$, soit $\beta = \varphi \circ f$ qui est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$. La différentielle de β peut se calculer grâce à la relation

$d\beta(x_0) = d\varphi(f(x_0)) \cdot df(x_0)$. Si on pose $\theta = f(x_0)$, on a (grâce à la relation (3)) :

$$J_\beta(x_0) = \begin{pmatrix} dX/dt(\theta) \\ dY/dt(\theta) \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f}{\partial x}(x_0) & \frac{\partial f}{\partial y}(x_0) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{dX}{dt}(\theta)\,\frac{\partial f}{\partial x}(x_0) & \frac{dX}{dt}(\theta)\,\frac{\partial f}{\partial y}(x_0) \\ \frac{dY}{dt}(\theta)\,\frac{\partial f}{\partial x}(x_0) & \frac{dY}{dt}(\theta)\,\frac{\partial f}{\partial y}(x_0) \end{pmatrix}$$

qui est bien une matrice deux par deux. On remarque qu'elle est de rang 1 en général (ses deux colonnes sont proportionnelles), ce qui est une conséquence facile de la relation (3) [preuve laissée au lecteur ; il s'agit d'un exercice classique d'algèbre linéaire].

- On retient des résultats précédents le calcul usuel de la dérivée d'une fonction composée, où l'on confond (dans les notations uniquement !) la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ et les valeurs prises $y \in \mathbb{R}^m$ via cette fonction : $f(x) = y$. On a donc $y = f(x)$ donc $dy_k = \sum_j \frac{\partial y_k}{\partial x_j} dx_j$

et de façon analogue $z = g(y)$, soit

$dz_\ell = \sum_k \frac{\partial z_\ell}{\partial y_k} dy_k$. Si on suppose que y est fonction de x (si on s'intéresse à la composée $g \circ f$), on injecte la valeur dy_k tirée de $y = f(x)$ au sein de la différentielle de la fonction g. Il vient

$$dz_\ell = \sum_k \frac{\partial z_\ell}{\partial y_k} \left(\sum_j \frac{\partial y_k}{\partial x_j} dx_j \right)$$

$$= \sum_j \left(\sum_k \frac{\partial z_\ell}{\partial y_k} \frac{\partial y_k}{\partial x_j} \right) dx_j$$

$$\equiv \sum_j \frac{\partial z_\ell}{\partial x_j} dx_j$$

ce qui montre, en identifiant ces deux dernières expressions, que

$$(7) \qquad \frac{\partial z_\ell}{\partial x_j} = \sum_k \frac{\partial z_\ell}{\partial y_k} \frac{\partial y_k}{\partial x_j}.$$

La relation (7) constitue la "règle de calcul" usuelle de manipulation des fonctions composées et de leurs différentielles. Elle est particulièrement simple à retenir !

- Dérivée seconde et matrice Hessienne.

Soit f une fonction numérique de deux variables réelles, c'est à dire $f : \mathbb{R}^2 \to \mathbb{R}$. Supposons la différentiable au voisinage de $x_0 \in \mathbb{R}^2$. On dispose donc de $df(x) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ si x est au voisinage de x_0, donc des dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$, qui sont elles-mêmes des

fonctions $\mathbb{R}^2 \to \mathbb{R}$ définies au voisinage de x_0.
On peut donc former l'application $\phi(x)$ de composantes $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$:

$$(8) \quad \mathbb{R}^2 \ni x \mapsto \phi(x) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \in \mathbb{R}^2 .$$

Cette application ϕ qui "regroupe" les dérivées partielles premières est donc une fonction $\mathbb{R}^2 \to \mathbb{R}^2$ définie au voisinage de $x_0 \in \mathbb{R}^2$.

- On peut différentier cette fonction vectorielle. Donc $d\phi(x) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ a pour matrice jacobienne $J_\phi \equiv H(x)$ qui est une matrice à deux lignes et deux colonnes. On a simplement

$$(9) \quad H(x) \equiv J_\phi(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \\ \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Le théorème de Schwarz affirme que la matrice H (la matrice Hessienne) est en fait symétrique:

$$(10) \quad \left[\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right](x_0) = \left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right](x_0).$$

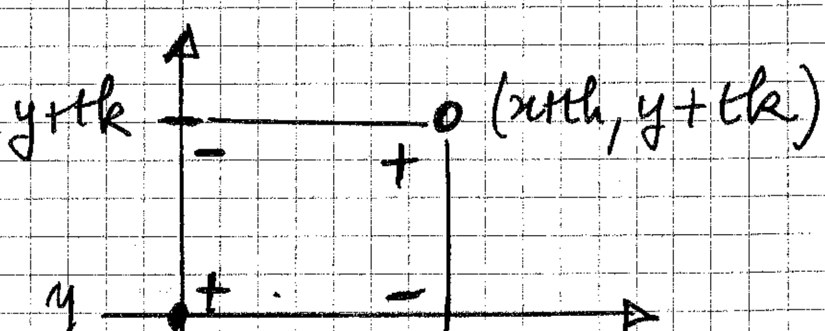
(Th) d'interversion des dérivées partielles (Schwarz)

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie au voisinage de $x_0 \in \mathbb{R}^2$. On suppose que les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont définies et continues au voisinage de x_0. On suppose de plus que $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont différentiables en x_0 alors la relation (10) est satisfaite.

- Nous proposons ici une <u>preuve simplifiée</u> qui consiste à supposer des hypothèses un peu plus fortes que celles du théorème. On se donne $(x, y) \in \mathbb{R}^2$ et un vecteur incrément $(h, k) \in \mathbb{R}^2$. On suppose que $f : \mathbb{R}^2 \to \mathbb{R}$ est de classe $\mathcal{C}^2$, c'est à dire deux fois continûment dérivable au voisinage de (x, y). Pour $t \in \mathbb{R}$, on forme l'expression

$$(11) \left\{ \begin{aligned} \varphi(t) = {} & f(x+th, y+tk) - f(x+th, y) \\ & - f(x, y+tk) + f(x, y) \end{aligned} \right.$$

illustrée à la figure ci-contre.



* Nous allons exprimer $\varphi(t)$ de deux façons.
D'une part, à t fixé, nous posons
$\alpha_t(h,k) \equiv f(x+th, y+tk) - f(x, y+tk)$.
alors $\varphi(t) = \alpha_t(h,k) - \alpha_t(h,0)$,
expression qui se transforme à l'aide du théorème des accroissements finis :

$$\varphi(t) = \frac{\partial \alpha_t}{\partial y}(h, \theta_1 k)\, k.$$

Nous calculons $\frac{\partial \alpha_t}{\partial y}(h,k)$ à l'aide de la relation (6). Il vient :

$$\frac{\partial \alpha_t}{\partial y} = \left(\frac{\partial f}{\partial y}(x+th, y+tk) - \frac{\partial f}{\partial y}(x, y+tk)\right) t$$

et cette expression peut à nouveau être exprimée avec le théorème des accroissements finis :

$$\left(\frac{\partial}{\partial y}\alpha_t\right)(h,k) = \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right)(x+\theta_2 th, y+tk)\, t^2 h$$

on en déduit :

$$(12) \quad \varphi(t) = \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right)(x+t\theta_2 h, y+t\theta_1 k)\, t^2 hk$$

avec $0 < \theta_1, \theta_2 < 1$.

* On introduit maintenant

$$\beta_t(h,k) = f(x+th, y+tk) - f(x+th, y).$$

alors
$$\varphi(t) = \beta_t(h,k) - \beta_t(0,k)$$
$$= \frac{\partial}{\partial x}\beta_t(\theta_3 h, k)\, h \quad , 0<\theta_3<1$$
grâce au théorème des accroissements finis. or

$$\left(\frac{\partial}{\partial x}\beta_t\right)(h,k) = \left[\frac{\partial f}{\partial x}(x+th, y+tk) - \frac{\partial f}{\partial x}(x+th, y)\right] t$$

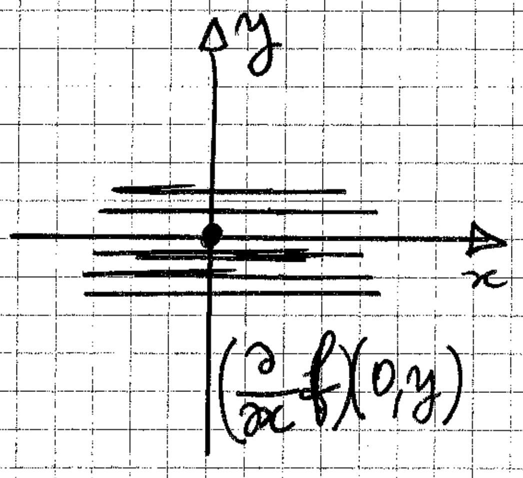
$$= \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)(x+t\ \ h, y+t\theta_4 k)\ t^2 k$$

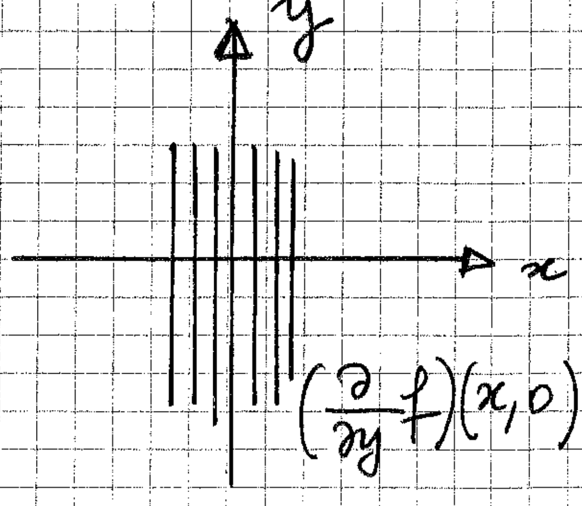
En appliquant cette relation avec h remplacé par $\theta_3 h$, on en déduit

(13) $$\varphi(t) = \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)(x+t\theta_3 h, y+t\theta_4 k)\ t^2 hk,$$
$$0<\theta_3, \theta_4<1$$

* On fait maintenant tendre t vers 0. Alors $\psi(t) \equiv \frac{1}{t^2}\varphi(t)$ tend (puisque $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$ est supposée continue en (x,y)) vers $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x,y)\,hk$ compte tenu de (13) et vers $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x,y)\,hk$ puisque $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ est continue en (x,y), en utilisant la relation (12). La relation (10) est alors conséquence de l'unicité de la limite. □

o Nous pouvons illustrer le théorème d'interversion des dérivées partielles avec les deux dessins ci-dessous.





Pour calculer $\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)(0,0)$, on a besoin des $\frac{\partial f}{\partial x}(0,y)$ (figure de gauche) alors que pour calculer $\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right)(0,0)$, on a besoin des valeurs $\left(\frac{\partial f}{\partial y}\right)(x,0)$ (figure de droite). En quelque sorte, le théorème de Schwarz exprime que "l'espace n'est pas une pâte feuilletée", si f est assez régulière.

- Bien entendu, le théorème d'interversion des dérivées partielles peut être en défaut si ses hypothèses ne sont pas satisfaites, ainsi que l'illustre l'exemple suivant. On pose

$$(14) \qquad f(x,y) = xy\,\frac{x^2-y^2}{x^2+y^2}\,, \quad (x,y)\neq(0,0)$$

et $f(0,0)=0$. On a sans difficulté

$$\frac{\partial f}{\partial x} = y\,\frac{x^2-y^2}{x^2+y^2} + xy\,\frac{2x}{x^2+y^2} + xy\,(x^2-y^2)\,\frac{(-2x)}{(x^2+y^2)^2}$$

donc $\left(\frac{\partial f}{\partial x}\right)(0,y) = -y$, d'où

$\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)(0,0) = -1$. Par ailleurs, on a

$$\frac{\partial f}{\partial y} = x\,\frac{x^2-y^2}{x^2+y^2} + xy\,\frac{-2y}{x^2+y^2} + xy(x^2-y^2)\,\frac{-2y}{(x^2+y^2)^2}$$

d'où $\left(\frac{\partial}{\partial y}f\right)(x,0) = x$ et $\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right)(0,0) = +1$ et la relation (10) n'est <u>pas</u> satisfaite.

• Forme bilinéaire symétrique.

La différentielle seconde $d^2f(x_0)$ de f définie de $\mathbb{R}^2$ dans $\mathbb{R}$ au voisinage de x_0 est en fait une forme bilinéaire sur $\mathbb{R}^2$:

(15) $$d^2f(x_0) \in \mathcal{L}_2(\mathbb{R}^2, \mathbb{R}) .$$

Elle opère sur $(h,k) \in \mathbb{R}^2$: $(d^2f)(x_0)\cdot(h,k) \in \mathbb{R}$ et elle est linéaire à la fois en h ($\bar{a}$ k fixé) et en k ($\bar{a}$ h fixé). En effet, la différentielle première $\mathbb{R}^2 \ni x \mapsto df(x) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ définit pour h fixé dans $\mathbb{R}^2$ une nouvelle fonction de $\mathbb{R}^2$ dans $\mathbb{R}$: $\mathbb{R}^2 \ni x \mapsto df(x)\cdot h \in \mathbb{R}$.
Si cette fonction (de $x \in \mathbb{R}^2$) est différentiable, sa différentielle (notée $d^2f(x)$) est ($\bar{a}$ h fixé) une forme linéaire $\mathbb{R}^2 \ni k \mapsto (d^2f(x)\cdot h)\cdot k \in \mathbb{R}$
Comme $\mathbb{R}^2 \ni h \mapsto df(x)\cdot h \in \mathbb{R}$ est linéaire, il est facile de constater que la fonction $\mathbb{R}^2 \ni h \mapsto (d^2f(x)\cdot h)\cdot k \in \mathbb{R}$ est également linéaire :

Prop La différentielle seconde $(d^2f)(x_0)$ est une forme bilinéaire qui opère sur $(h,k) \in \mathbb{R}^2 \times \mathbb{R}^2$:

(16) $$(d^2f)(x_0)\cdot(h,k) \in \mathbb{R}, \quad h \in \mathbb{R}^2, k \in \mathbb{R}^2.$$

Si e_i désigne un vecteur de base de $\mathbb{R}^2$,

on a (17) $d^2f(x_0)\cdot(e_i,e_j) = \left(\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)\right)(x_0)$;

le théorème de Schwarz exprime donc la propriété suivante

Prop La différentielle seconde est une forme bilinéaire symétrique.

Soit f définie $\mathbb{R}^2 \to \mathbb{R}$ au voisinage de x_0, différentiable au voisinage de x_0 avec des dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ continues au voisinage de x_0. Si ces deux applications sont différentiables en x_0, alors la différentielle seconde $d^2f(x_0) \in \mathcal{L}_2(\mathbb{R}^2,\mathbb{R})$ est une forme bilinéaire symétrique.

• Nous pouvons exprimer $(d^2f)(x_0)\cdot(h,k)$ à l'aide de la matrice Hessienne :

$$(18) \quad (d^2f)(x_0)\cdot(h,k) = h\cdot H(x_0)\cdot k^t .$$

En effet, il suffit d'exprimer la bilinéarité :

$$(d^2f)(x_0)\cdot(h,k) = (d^2f)(x_0)\left(\sum_i h_i e_i, \sum_j k_j e_j\right)$$

$$= \sum_{i,j} h_i k_j \,(d^2f)(x_0)(e_i,e_j)$$

$$= \sum_{i,j} h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = h\cdot H(x_0)\cdot k^t .$$

Prop) La Hessienne est positive.

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ différentiable au voisinage de x_0, deux fois différentiable en x_0. On suppose f minimale au voisinage de x_0 (relation (19)). Alors la matrice Hessienne symétrique $H(x_0)$ est positive :

$$(22) \qquad h \cdot H(x_0) \cdot h^t \geqslant 0, \quad \forall h \in \mathbb{R}^2.$$

- La preuve consiste à étudier $\varphi(\theta)$ introduite en (21) au voisinage de $\theta = 0$:

$$\varphi(\theta) = \varphi(0) + \frac{1}{2}\varphi''(0)\,\theta^2 + \theta^2\,\varepsilon(\theta)$$

puisque $\varphi'(0) = 0$ compte tenu de la proposition précédente. De $\varphi(\theta) \geqslant \varphi(0)$ au voisinage de $\theta = 0$ on déduit $\varphi''(0) \geqslant 0$. Or $\varphi''(0) = d^2f(x_0) \cdot (h, h)$, ce qui montre (22) compte tenu de la relation (18).

- On peut rappeler que la relation (22) entraîne que les valeurs propres (réelles) de $H(x_0)$ sont alors toutes positives ou nulles et que la relation (22) s'exprime à l'aide des dérivées partielles sous la forme

$$(23) \qquad \frac{\partial^2 f}{\partial x^2} \geqslant 0, \qquad \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 \geqslant 0.$$

- Conditions d'extremum.

Soit $f: \mathbb{R}^2 \to \mathbb{R}$ définie au voisinage de x_0, on la suppose différentiable en x_0 et minimale au voisinage de x_0 :

$$(19) \quad \forall x \in \mathcal{V}(x_0), \quad f(x) \geq f(x_0)$$

Prop Nullité de la différentielle

Sous les hypothèses précédentes,

$$(20) \quad \frac{\partial f}{\partial x}(x_0) = \frac{\partial f}{\partial y}(x_0) = 0, \quad \text{ie} \quad df(x_0) = 0 .$$

- La preuve consiste à se donner h arbitraire dans $\mathbb{R}^2$ et à introduire, pour $\theta \in \mathbb{R}$ au voisinage de 0, la fonction $\mathbb{R} \to \mathbb{R}$ suivante

$$(21) \quad \varphi(\theta) = f(x + \theta h).$$

alors l'hypothèse (19) exprime que $\varphi(\theta) \geq \varphi(0)$ pour θ au voisinage de 0. Donc (c'est classique, et c'est un exercice laissé au lecteur !) $\varphi'(0) = 0$. Or $\varphi'(0) = df(x_0) \cdot h$. En particulier pour $h = e_1$ et $h = e_2$, ce qui montre la relation (20) et la propriété. $\square$

- Le cas où f est maximale en x_0 (remplacer $\geq$ par $\leq$ dans la relation (19)) conduit au même résultat [Exercice!].

• Les relations précédentes relèvent d'un cours d'algèbre linéaire classique. Toutefois, dans le cas de $f: \mathbb{R}^2 \to \mathbb{R}$, la preuve se réduit à l'étude du signe du trinome. On pose

$$(24) \qquad H(x_0) = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$$

et $h \cdot H(x_0) \cdot h^t = \alpha x^2 + 2\beta xy + \gamma y^2$ pour $h = (x, y)$. En prenant par exemple $y = 1$, le trinome $x \mapsto \alpha x^2 + 2\beta x + \gamma$ est toujours ≥ 0 si et seulement s'il ne s'annule pas (donc son discriminant $\beta^2 - \alpha\gamma$ est ≤ 0, ce qui exprime la seconde condition de (23)) et si le coefficient α de x^2 est ≥ 0 (ce qui correspond à la première condition de (23)).

Les valeurs propres λ de $H(x_0)$ annulent le déterminant $\det(H(x_0) - \lambda I) = \begin{vmatrix} \alpha-\lambda & \beta \\ \beta & \gamma-\lambda \end{vmatrix}$
$= \lambda^2 - (\alpha+\gamma)\lambda + \alpha\gamma - \beta^2$
qui a lui toujours deux racines réelles

$$\lambda_\pm = \frac{1}{2}\left[(\alpha+\gamma) \pm \sqrt{4\beta^2 + (\alpha-\gamma)^2}\,\right].$$

Leur produit $\lambda_+\lambda_- = \alpha\gamma - \beta^2$ est ≥ 0 à cause de (23) donc leur signe est celui de leur somme $\lambda_+ + \lambda_- = \alpha + \gamma$ qui est positive compte-tenu de (23) à nouveau [pourquoi exactement?]

• Formule de Taylor.

On procède comme pour les fonctions d'une seule variable réelle. Avec $\theta \mapsto \varphi(\theta)$ définie à la relation (21), on a $\varphi'(\theta) = df(x+\theta h) \cdot h$ et

$$\varphi(1) - \varphi(0) = f(x+h) - f(x),$$
$$= \int_0^1 \varphi'(\theta)\, d\theta = \int_0^1 df(x+\theta h) \cdot h \, d\theta .$$

$$(25) \quad f(x+h) = f(x) + \int_0^1 df(x+\theta h) \cdot h \, d\theta$$

∘ On peut intégrer par parties la relation

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(\theta)\, d\theta$$
$$= \Big[(\theta-1)\varphi'(\theta)\Big]_0^1 - \int_0^1 (\theta-1)\varphi''(\theta)\, d\theta$$
$$= \varphi'(0) + \int_0^1 (1-\theta)\varphi''(\theta)\, d\theta .$$

Comme $\varphi''(\theta) = (d^2 f)(x+\theta h) \cdot (h,h)$, on en déduit

Prop Une première <u>formule de Taylor</u>

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ deux fois continuement dérivable au voisinage du segment $[x_0, x_0+h] = \{(1-\theta)\, x_0 + \theta(x_0+h),\ 0 \leq \theta \leq 1\}$. Alors on a la formule de Taylor (avec reste intégral) suivante

$$(26) \quad f(x_0+h) = f(x_0) + df(x_0) \cdot h + \int_0^1 (1-\theta)\, d^2 f(x_0+\theta h) \cdot (h,h)\, d\theta$$

(Th) Formule de Taylor avec reste intégral.

On suppose f $(n+1)$ fois continuement différentiable au voisinage du segment $[x_0, x_0+h]$. On a la formule de Taylor (avec reste intégral) :

$$(27)\left\{\begin{aligned} f(x_0+h) &= f(x_0) + df(x_0)\cdot h + \cdots + \frac{1}{k!} d^k f(x_0)\cdot(h,\dots,h) \\ &+ \cdots + \frac{1}{n!} (d^n f)(x_0)(h,\dots,h) + \frac{1}{n!}\int_0^1 (1-\theta)^n d^{n+1} f(x_0)(h,\dots,h)\, d\theta \end{aligned}\right.$$

- La preuve est un exercice qui consiste à reprendre ce qui a été fait au chapitre 4, à l'appliquer à $\varphi(\theta)$ définie à la relation (21) en remarquant que $\varphi^{(k)}(\theta) = d^k f(x_0+\theta h)\cdot\underbrace{(h,\dots,h)}_{k \text{ facteurs}}$ $\square$

- Nous n'avons pas développé le fait que la différentielle k^e $d^k f(x_0)$ est une forme k-linéaire symétrique de ses arguments. Le lecteur pourra consulter (par exemple) l'ouvrage d'Henri Cartan "Calcul différentiel" (Hermann, Paris, 1977).

Dubois

01 décembre 2009

Suite à la lecture de ces Notes par Claude Durand, auditeur CNAM en 2010-2011, correction de coquilles.

Dubois, 4 juin 2011.