5 Observer le voisinage : les différentielles d'ordres supérieurs et fonctions de classe C^k

Dans cette partie, on se restreindra aux fonctions à valeurs réelles. On généralisera aisément aux fonctions à valeurs dans $\mathbb{R}^p$ en décomposant $f = (f_1, ..., f_p)$ où les f_i sont à valeurs réelles.

5.1 Fonctions de classe C^1

Définition 5.1 (Fonction de classe C^1). *Soit $\Omega \subset \mathbb{R}^n$ un ouvert et $f : \Omega \mapsto \mathbb{R}$. On dit que f est de classe C^1 sur Ω, noté $f \in C^1(\Omega)$, si f admet des dérivées partielles premières continues sur Ω.*

Exemple 5.2. (Fonction de classe C^1) Soit $f : \mathbb{R}^3 \setminus \{(0,0,0)\} \to \mathbb{R}$ définie par

$$\forall (x, y, z) \in \mathbb{R}^3 \setminus \{(0,0,0)\}, \quad f(x, y, z) = \ln(x^2 + y^2 + z^2).$$

Alors f est de classe C^1 sur $\mathbb{R}^3 \setminus \{(0,0,0)\}$. En effet, on a

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

et ces trois fonctions sont continues sur $\mathbb{R}^3 \setminus \{(0,0,0)\}$ comme quotient de fonctions continues dont le dénominateur ne s'annule pas.

Proposition 5.3 (C^1 implique différentiable). *Si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ est de classe C^1 sur un ouvert Ω et $x \in \Omega$, alors f est différentiable en x.*

Démonstration. Pour alléger les notations, on suppose que $n = 2$ (le cas général se montre de même). Pour $h = (h_1, h_2)$ tel que $x + h \in \Omega$, on définit

$$r(h) = f(x+h) - f(x) - h_1 \frac{\partial f}{\partial x_1}(x) - h_2 \frac{\partial f}{\partial x_2}(x).$$

Montrons que, quand $h \to 0$, $r(h) = o(\|h\|_2)$. En effet, on a

$$\begin{aligned}
f(x+h) - f(x) &= f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) + f(x_1, x_2 + h_2) - f(x_1, x_2) \\
&= \int_0^{h_1} \frac{\partial f}{\partial x_1}(x_1 + t, x_2 + h_2)dt + \int_0^{h_2} \frac{\partial f}{\partial x_2}(x_1, x_2 + t)dt \\
&= h_1 \int_0^1 \frac{\partial f}{\partial x_1}(x_1 + sh_1, x_2 + h_2)ds + h_2 \int_0^1 \frac{\partial f}{\partial x_2}(x_1, x_2 + sh_2)ds,
\end{aligned}$$

où on a posé $t = sh_1$ dans la première intégrale et $t = sh_2$ dans la deuxième. On peut donc écrire

$$r(h) = h_1 \int_0^1 \left(\frac{\partial f}{\partial x_1}(x_1 + sh_1, x_2 + h_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right) ds + h_2 \int_0^1 \left(\frac{\partial f}{\partial x_2}(x_1, x_2 + sh_2) - \frac{\partial f}{\partial x_2}(x_1, x_2) \right) ds.$$

Soit $\varepsilon > 0$, alors par continuité des dérivées partielles de f en x, il existe $\delta > 0$ tel que pour tout $h_1, h_2 \in]-\delta, \delta[$ et $s \in [0,1]$,

$$\left| \frac{\partial f}{\partial x_1}(x_1 + sh_1, x_2 + h_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right| < \varepsilon$$

et

$$\left|\frac{\partial f}{\partial x_2}(x_1, x_2 + sh_2) - \frac{\partial f}{\partial x_2}(x_1, x_2)\right| < \varepsilon.$$

On en déduit que $|r(h)| \le \varepsilon(|h_1| + |h_2|)$ et donc que $r(h) = o(\|h\|_2)$ par équivalence des normes.

□

Proposition 5.4 (Opérations et fonctions de classe C^1). *Les sommes, produits, quotients (avec dénominateur non-nul) et composées de fonctions de classe C^1 sont de classe C^1.*

Démonstration. C'est une conséquence immédiate des résultats sur les opérations de fonctions continues et celles qui sont différentiables. □

5.2 Différentielle d'ordre 2, fonctions de classe C^2 et formule de Taylor-Young

Remarquons tout d'abord que si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ est différentiable, alors $h \mapsto D_x f(h)$ (pour x fixé) est linéaire mais $x \mapsto D_x f(h)$ (pour h fixé) ne l'est pas nécessairement ! On peut donc considérer la différentiabilité de l'application $x \mapsto D_x f$ sans que cela soit trivial.

Si f est différentiable, on sait que pour tout $h = (h_1, ..., h_n) \in \mathbb{R}^n$, pour tout $x \in \Omega$,

$$D_x f(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i.$$

Ainsi, supposer la différentiabilité de $x \mapsto D_x f$, c'est supposer celle de $x \mapsto \frac{\partial f}{\partial x_i}(x)$ pour tout $i \in \{1, ..., n\}$. Dans le cas, on calcule aisément cette différentielle comme suit : pour tout $i \in \{1, ..., n\}$, pour tout $k = (k_1, ..., k_n) \in \mathbb{R}^n$,

$$D_x\left(\frac{\partial f}{\partial x_i}\right)(k) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x) k_j,$$

ce qui implique, par linéarité, que

$$D_x(D_x f(h))(k) = \sum_{i=1}^{n} D_x\left(\frac{\partial f}{\partial x_i}\right)(k) h_i = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x) h_i k_j.$$

On notera :

- $D_x^2 f(h,k) := D_x(D_x f(h))(k)$, $D_x^2 f$ étant donc une application bilinéaire ;
- $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(x) := \dfrac{\partial}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)(x)$ les n^2 dérivées partielles secondes de f au point x.

Définition 5.5 (Différentielle seconde). *Soit $\Omega \subset \mathbb{R}^n$ un ouvert et $f : \Omega \to \mathbb{R}$ une application différentiable. Si, pour tout $1 \le i \le p$, l'application $\frac{\partial f}{\partial x_i}$ est différentiable sur Ω, alors on dit que f est deux fois différentiable sur Ω et on note $D_x^2 f = D_x(Df)$ la différentielle seconde de f au point $x \in \Omega$. De plus, on sait que pour tout $1 \le i \le p$, $\frac{\partial f}{\partial x_i}$ admet des dérivées partielles en tout point $x \in \Omega$, appelées dérivées partielles secondes, et données par*

$$\forall 1 \le i \le j, \quad \forall 1 \le j \le n, \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(x) := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x) = \partial_{ji} f(x).$$

La différentielle seconde est donnée par la matrice $n \times n$ dont le coefficient (i, j) est $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$, c'est-à-dire

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix},$$

appelée Matrice Hessienne de f au point x. De la même façon que la forme linéaire $\mathbb{R}^n \to \mathbb{R}$, $h \mapsto \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i$ est associée à la différentielle première de f en x, la forme bilinéaire associée à la différentielle seconde est donnée par

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad (h, k) \mapsto D_x^2 f(h, k) = \langle H_f(x) h, k \rangle = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i k_j.$$

Remarque 5.6. Il est important de bien comprendre les domaines et espaces d'arrivée des différentielles secondes. L'application Df est définie sur un voisinage V de x contenu dans $\Omega \subset \mathbb{R}^n$ et $Df : V \to \mathscr{L}(\mathbb{R}^n, \mathbb{R})$. Ainsi, $D_x^2 f = D_x(Df) \in \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R}))$. Donc $D_x f$ s'applique linéairement à des éléments $h \in \mathbb{R}^n$ et $D_x^2 f(h)$ s'applique linéairement à des éléments $k \in \mathbb{R}^n$ en étant à valeurs réelles. C'est pour cela que l'on peut donc identifier $D_x^2 f$ à une forme bilinéaire en les deux arguments h et k. On pourra donc noter $D_x^2 f(h)(k) = D_x^2 f(h, k)$.

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathscr{L}(\mathbb{R}^n, \mathbb{R}) \quad \longrightarrow \mathbb{R} \\ h &\longmapsto D_x^2 f(h) : k \longmapsto D_x^2 f(h)(k) = (D_x(Df)(h))(k) = D_x^2 f(h, k). \end{aligned}$$

Exemple 5.7. (Différentielle seconde)
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie pour tout $(x, y) \in \mathbb{R}^2$ par

$$f(x, y) = x^4 y + 2x^2 - 3y^3.$$

Alors les dérivées partielles premières sont

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 y + 4x, \quad \frac{\partial f}{\partial y}(x, y) = x^4 - 9y^2.$$

Les applications $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont différentiables car polynomiales. On en déduit que f est deux fois différentiable sur $\mathbb{R}^2$. Ses dérivées partielles secondes sont données par

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 12x^2 y + 4, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = -18y, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = 4x^3.$$

La hessienne de f au point (x,y) est donc

$$H_f(x,y) = \begin{pmatrix} 12x^2y+4 & 4x^3 \\ 4x^3 & -18y \end{pmatrix},$$

et en particulier, au point $(0,1)$,

$$H_f(0,1) = \begin{pmatrix} 4 & 0 \\ 0 & -18 \end{pmatrix}.$$

Dans le cas général, la forme bilinéaire associée à la différentielle seconde est, pour $h = (h_1, h_2) \in \mathbb{R}^2$ et $k = (k_1, k_2) \in \mathbb{R}^2$,

$$D^2_{(x,y)} f(h,k) = (12x^2y+4)h_1k_1 + 4x^3h_2k_1 + 4x^3h_1k_2 - 18yh_2k_2,$$

de forme quadratique correspondante

$$\langle H_f(x,y)h, h\rangle = D^2_{(x,y)} f(h,h) = (12x^2y+4)h_1^2 + 8x^3h_1h_2 - 18h_2^2.$$

et donc, au point $(0,1)$, cette forme vaut

$$D^2_{(0,1)} f(h,k) = 4h_1k_1 - 18h_2k_2,$$

et sa forme quadratique associée est

$$D^2_{(0,1)} f(h,h) = 4h_1^2 - 18h_2^2.$$

Remarque 5.8. (Développement au premier ordre de la différentielle) - (Partie distribuée en CM) Au sens de la définition précédente, être deux fois différentiable pour f est équivalent avec le fait que Df soit différentiable, car

$$\forall x_0 \in \Omega, \forall h \in \mathbb{R}^n, \quad D_{x_0} f(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) h_i.$$

En effet, la différentiabilité (par rapport au point) de toutes les dérivées partielles équivaut à celle de l'application $x_0 \mapsto D_{x_0} f(h)$ pour tout h fixé. Ainsi, on peut écrire, dans l'espace $\mathscr{L}(\mathbb{R}^n, \mathbb{R})$, pour tout h' suffisamment petit (tendant vers 0), pour tout $x_0 \in \Omega$,

$$D_{x_0+h'} f = D_{x_0} f + D_{x_0}(Df)(h') + o(\|h'\|),$$

qui peut se réécrire, avec $\varepsilon : \mathbb{R}^n \to \mathscr{L}(\mathbb{R}^n, \mathbb{R})$ tendant vers l'application linéaire nulle quand $h' \to 0$:

$$D_{x_0+h'} f = D_{x_0} f + D_{x_0}(Df)(h') + \|h'\|\varepsilon(h').$$

On voit ici qu'un problème conceptuel/technique survient car nous n'avons pas défini la norme d'une application linéaire, donc la notion $o(\|h'\|)$ pourrait paraître ici un peu ambigu. Une autre façon d'écrire cette égalité est de l'appliquer à un vecteur $h \in \mathbb{R}^n$:

$$D_{x_0+h'} f(h) = D_{x_0} f(h) + D_{x_0}[(Df)(h')](h) + \|h'\|\varepsilon(h').$$

où maintenant $\varepsilon : \mathbb{R}^n \to \mathbb{R}$ dépend de h et tend vers 0, uniformément par rapport à h, quand $h' \to 0$. On rappelle ce que cette égalité signifie : $\forall h \in \mathbb{R}^n$, $\forall \varepsilon > 0$, $\exists \delta > 0$, $\forall h' \in \mathbb{R}^n$,

$$\|h'\| < \delta \Rightarrow |D_{x_0+h'} f(h) - D_{x_0} f(h) - D_{x_0}[(Df)(h')](h)| < \varepsilon \|h'\|.$$

Comme l'application

$$\varphi : h \mapsto D_{x_0+h'} f(h) - D_{x_0} f(h) - D_{x_0}[(Df)(h')](h)$$

est linéaire, on peut appliquer l'égalité précédente au vecteur $\frac{h}{\|h\|}$, $h \neq 0$ (le cas $h = 0$ sera évident à la fin du raisonnement) et on obtient (pour un nouveau ε qui tend toujours vers 0, puisque l'on a seulement échangé h par $h/\|h\|$ et que la limite était uniforme en h) :

$$\varphi\left(\frac{h}{\|h\|}\right) = \|h'\|\varepsilon(h'),$$

ce qui implique que $\varphi(h) = \|h\|\|h'\|\varepsilon(h')$ et finalement, comme $D_{x_0}[(Df)(h')](h) = D^2_{x_0} f(h', h)$, que, pour tout $x_0 \in \Omega$, pour tout h' suffisamment petit et pour tout $h \in \mathbb{R}^n$,

$$D_{x_0+h'} f(h) = D_{x_0} f(h) + D^2_{x_0} f(h', h) + \|h\|\|h'\|\varepsilon(h').$$

Encore une fois, cela se traduit par : $\forall h \in \mathbb{R}^n$, $\forall \varepsilon > 0$, $\exists \delta > 0$, $\forall h' \in \mathbb{R}^n$,

$$\|h'\| < \delta \Rightarrow |D_{x_0+h'} f(h) - D_{x_0} f(h) - D^2_{x_0} f(h', h)| < \varepsilon\|h\|\|h'\|.$$

Proposition 5.9 (**Opérations et fonctions deux fois différentiables**). *Les sommes, produits, quotients (avec dénominateur non-nul) et composées de fonctions deux fois différentiables sont deux fois différentiables.*

Démonstration. C'est une conséquence immédiate des résultats sur les opérations de fonctions continues et celles qui sont différentiables. □

Définition 5.10 (**Fonction de classe C^2**). *Soit $\Omega \subset \mathbb{R}^n$ un ouvert et $f : \Omega \mapsto \mathbb{R}$. On dit que f est de classe C^2 sur Ω, noté $f \in C^2(\Omega)$, si f admet des dérivées partielles premières et secondes continues sur Ω, c'est-à-dire si $f \in C^1(\Omega)$ et admet des dérivées partielles secondes continues sur Ω.*

Exemple 5.11. (Fonction de classe C^2)
Il est clair que la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie pour tout $(x, y) \in \mathbb{R}^2$ par

$$f(x, y) = x^4 y + 2x^2 - 3y^3$$

est de classe C^2, ses dérivées partielles secondes étant polynomiales.

Proposition 5.12 (**Opérations et fonctions de classe C^2**). *Les sommes, produits, quotients (avec dénominateur non-nul) et composées de fonctions de classe C^2 sont de classe C^2.*

Démonstration. C'est une conséquence immédiate des résultats sur les opérations de fonctions continues et celles qui sont différentiables. □

Proposition 5.13 (**C^2 implique deux fois différentiable**). *Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $x \in \Omega$ et $f : \Omega \mapsto \mathbb{R}$. Si f est de classe C^2, alors f est deux fois différentiable en x.*

Démonstration. Admise. La preuve est similaire à celle du cas C^1. □