Passer de $\mathbb{R}^n$ à $\mathbb{R}$: la fonction auxiliaire

– complément au Chapitre 7. Fonctions de classe C^k –

NB Ce texte détaille les calculs de dérivées pour la fonction auxiliaire $t \mapsto f((1-t)x+ty)$.

Une méthode systématique pour « transférer » une propriété connue pour des fonctions $g : I \to \mathbb{R}$, avec $I \subset \mathbb{R}$ intervalle, à des fonctions $f : U \to \mathbb{R}$, avec $U \subset \mathbb{R}^n$ ouvert, consiste à considérer la fonction auxiliaire

$$J \ni t \mapsto g(t) := f((1-t)x+ty),$$

où $x, y \in \mathbb{R}^n$ sont fixés et J est l'ensemble ouvert $J := \{t \in \mathbb{R}\,;\ (1-t)x+ty \in U\}$.
Pour que le transfert soit possible, il faut traduire les propriétés de différentiabilité de f en propriété de dérivabilité de g, et être en mesure de calculer les dérivées de g d'ordre $k \geq 1$ à partir des dérivées partielles de f. Ceci est l'objet des résultats qui suivent. Nous allons donner deux formules de calcul, une, « étalée », dont la preuve est plus facile à suivre, l'autre, plus « compacte », dont la validité repose sur le théorème de symétrie de Schwarz et les « coefficients multinomiaux ». Par ailleurs, nous allons détailler ces formules aux ordres 1, 2, 3. Les ordres 1 et 2 sont les plus utilisés.

Proposition. (Formules « étalées ») Si f est k fois différentiable, alors g est k fois dérivable, et nous avons, pour $t \in J$,

$$g'(t) = \sum_{j=1}^{n} [(\partial_{x_j} f)((1-t)x+ty)](y_j - x_j) \tag{1}$$

$$g''(t) = \sum_{p=1}^{n}\sum_{j=1}^{n} [(\partial_{x_p}\partial_{x_j} f)((1-t)x+ty)](y_j - x_j)(y_p - x_p), \tag{2}$$

$$g'''(t) = \sum_{r=1}^{n}\sum_{p=1}^{n}\sum_{j=1}^{n} [(\partial_{x_r}\partial_{x_p}\partial_{x_j} f)((1-t)x+ty)](y_j - x_j)(y_p - x_p)(y_r - x_r), \tag{3}$$

$$\ldots$$

$$\begin{aligned} g^{(k)}(t) = \sum_{j_k=1}^{n}\sum_{j_{k-1}=1}^{n}\ldots\sum_{j_1=1}^{n} & [(\partial_{x_{j_k}}\partial_{x_{j_{k-1}}}\cdots\partial_{x_{j_1}} f)((1-t)x+ty)] \\ & \times (y_{j_1} - x_{j_1})\cdots(y_{j_{k-1}} - x_{j_{k-1}})(y_{j_k} - x_{j_k}). \end{aligned} \tag{4}$$

Si, de plus, $f \in C^k$, alors $g \in C^k$.

Les formules (1)–(3) sont les déclinaisons de la formule générale (4) dans les cas particuliers $k = 1, 2, 3$. Pour appréhender cette formule, il est important de comprendre comment on passe de $k = 1$ à $k = 2$, respectivement de $k = 2$ à $k = 3$: on ajoute une nouvelle somme, par exemple dans (3) la somme sur r, et pour chaque expression de la somme précédente (dans cet exemple, de (2)), on dérive par rapport à x_r et on multiplie par $(y_r - x_r)$.

Démonstration. Pour simplifier les notations, posons $F(t) := (1-t)x + ty, \forall\, t \in J$, de sorte que : (i) F est dérivable ; (ii) $F'(t) = y - x, \forall\, t \in J$; (iii) $g = f \circ F$.

La preuve de (4) se fait par récurrence sur k. Si $k = 1$, f est différentiable, et la règle de la chaîne appliquée à $g = f \circ F$ donne que g est dérivable et

$$g'(t) = F'(t) \cdot [(\nabla f)(F(t))] = (y - x) \cdot [\nabla f(F(t))] = \sum_{j=1}^{n} [(\partial_{x_j} f)(F(t))](y_j - x_j), \quad (5)$$

ce qui donne (1) et le cas où $k = 1$.

Supposons (4) vraie, et montrons la formule analogue pour $k + 1$. f étant $(k + 1)$ fois différentiable, elle est k fois différentiable (donc (4) est valide) et ses dérivées partielles d'ordre k sont différentiables. En appliquant la règle de la chaîne à $[\partial_{x_{j_k}} \partial_{x_{j_{k-1}}} \ldots \partial_{x_{j_1}} f] \circ F$, nous obtenons (comme dans la preuve de (5))

$$\begin{aligned}\Big([(\partial_{x_{j_k}} \partial_{x_{j_{k-1}}} \ldots \partial_{x_{j_1}} f] \circ F\Big)'(t) = \sum_{j_{k+1}=1}^{n} &(\partial_{x_{j_{k+1}}} \partial_{x_{j_k}} \partial_{x_{j_{k-1}}} \ldots \partial_{x_{j_1}} f)(F(t)) \\ &\times (y_{j_{k+1}} - x_{j_{k+1}}).\end{aligned} \quad (6)$$

En combinant (4) et (6), nous obtenons que g est $(k + 1)$ fois dérivable et que

$$\begin{aligned}g^{(k+1)}(t) = \sum_{j_{k+1}=1}^{n} \sum_{j_k=1}^{n} \sum_{j_{k-1}=1}^{n} \ldots \sum_{j_1=1}^{n} &[(\partial_{x_{j_{k+1}}} \partial_{x_{j_k}} \partial_{x_{j_{k-1}}} \cdots \partial_{x_{j_1}} f)(F(t))] \\ &\times (y_{j_1} - x_{j_1}) \cdots (y_{j_k} - x_{j_k})(y_{j_{k+1}} - x_{j_{k+1}}),\end{aligned}$$

ce qui achève la preuve par récurrence.

Enfin, si $f \in C^k$, le membre de droite de (4) est continu par rapport à t, et donc $g \in C^k$. □

Nous allons maintenant donner quelques formules plus courte pour l'expression de $g^{(k)}(t)$.

Formule courte si $k = 1$. C'est celle de (5), c'est-à-dire

$$g'(t) = [(\nabla f)(F(t))] \cdot (y - x). \quad (7)$$

Formule courte si $k = 2$. Nous avons

$$g''(t) = [H_{F(t)} f(y - x)] \cdot (y - x), \quad (8)$$

où $H_a f$ désigne la matrice hessienne de f évaluée en $a \in U$. Vérifions la validité de (8). L'élément en p^e ligne et j^e colonne de $(H_{F(t)} f)$ est $(\partial_{x_p} \partial_{x_j} f)(F(t))$, et donc $H_{F(t)} f(y - x)$ est le vecteur dont la p^e coordonnée est

$$\sum_{j=1}^{n} [(\partial_{x_p} \partial_{x_j} f)(F(t))](y_j - x_j), \; \forall\, 1 \le p \le n.$$

Il s'ensuit que

$$[H_{F(t)}f(y-x)]\cdot(y-x)=\sum_{p=1}^{n}\sum_{j=1}^{n}[(\partial_{x_p}\partial_{x_j}f)(F(t))](y_j-x_j)(y_p-x_p),$$

et l'expression trouvée est précisément le membre de droite de (2).

Passons au cas général.

Proposition. (Formule « compacte ») Si f est k fois différentiable, nous avons

$$g^{(k)}(t)=k!\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=k}}\frac{1}{\alpha!}[(\partial^\alpha f)(F(t))](y-x)^\alpha. \tag{9}$$

Rappelons que, si $\alpha=(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}^n$ et $z=(z_1,\ldots,z_n)\in\mathbb{R}^n$, alors : (i) $\alpha!:=\alpha_1!\alpha_2!\cdots\alpha_n!$; (ii) $z^\alpha:=(z_1)^{\alpha_1}(z_2)^{\alpha_2}\cdots(z_n)^{\alpha_n}$; (iii) $\partial^\alpha f:=\partial_{x_n}^{\alpha_n}\partial_{x_{n-1}}^{\alpha_{n-1}}\ldots\partial_{x_1}^{\alpha_1}f$.

Démonstration. La preuve de cette égalité repose sur : (i) le théorème de symétrie de Schwarz ; (ii) la formule des coefficients multinomiaux. Plus précisément :

(i) Le théorème de Schwarz montre que la valeur de la dérivée partielle $\partial_{x_{j_k}}\partial_{x_{j_{k-1}}}\cdots\partial_{x_{j_1}}f$ ne change pas si les variables $x_{j_1},x_{j_2},\ldots,x_{j_k}$ sont permutées. Il s'ensuit que la valeur de cette dérivée partielle ne dépend que de $\alpha_1,\ldots,\alpha_n$, où α_ℓ est le nombre d'indices j tels que $j=\ell$. Une autre façon d'exprimer cette symétrie est la suivante. Posons $\alpha:=(\alpha_1,\ldots,\alpha_n)$, de sorte que : (j) $|\alpha|=k$ (car nous avons une dérivée d'ordre k) ; (jj) $\partial_{x_{j_k}}\partial_{x_{j_{k-1}}}\cdots\partial_{x_{j_1}}f=\partial^\alpha f$ (grâce au théorème de Schwarz) ; (jjj) $(y_{j_1}-x_{j_1})\cdots(y_{j_{k-1}}-x_{j_{k-1}})(y_{j_k}-x_{j_k})=(y-x)^\alpha$.

(ii) De ce qui précède, le terme $[(\partial^\alpha f)(F(t))](y-x)^\alpha$ apparaît autant de fois que le nombre de listes $(j_1,\ldots,j_k)$ faisant apparaître α_1 fois 1, α_2 fois $2,\ldots,\alpha_n$ fois n. Ce nombre est appelé coefficient multinomial, et vaut $\dfrac{k!}{\alpha!}$. Pour comprendre où ce nombre apparaît dans la pratique et sa formule, voir la formule du multinôme

`https://fr.wikipedia.org/wiki/Formule_du_multinôme_de_Newton`

Si $n=2$, nous retrouvons les coefficients binomiaux $\binom{k}{\alpha_1}$.

Nous obtenons (9) de (4), (i), (ii) et la formule du coefficient multinomial. $\square$