Chapitre I
Calcul Différentiel dans $\mathbb{R}^n$

La dérivée d'une fonction réelle d'une variable réelle se révèle être un outil essentiel pour l'étude de cette fonction. Dans ce chapitre, nous introduisons son extension aux applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$.

Après avoir donné la définition et les premières propriétés de la différentielle, nous montrons comment la calculer à partir des dérivées partielles premières : la matrice formée des dérivées partielles premières, appelée matrice jacobienne, joue le rôle de la dérivée d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Nous relions ensuite continuité, différentiabilité et existence de dérivées partielles premières. Nous montrons également comment la différentielle se comporte à la composition de deux applications différentiables g et f : la matrice jacobienne de $g \circ f$ est le produit des matrices jacobiennes de g et de f. Enfin, nous étudions les dérivées partielles d'ordre supérieur et établissons les formules de Taylor-Lagrange et Taylor-Young pour les applications de $\mathbb{R}^2$ dans $\mathbb{R}$.

Les définitions et les propriétés, liées à la topologie de $\mathbb{R}^n$ et utilisées dans ce chapitre, figurent dans l'Annexe.

I.1 Différentiabilité.

Commençons en rappelant qu'une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ est *dérivable* en $a \in \mathbb{R}$ si la limite de $\dfrac{f(a+h)-f(a)}{h}$, lorsque h tend vers zéro, existe. On note cette limite $f'(a)$ et on l'appelle dérivée de f en a. Cette définition équivaut à :

$$f(a+h) - f(a) = hf'(a) + |h|\,\varepsilon(h)\,, \text{ avec } \lim_{h\to 0}\varepsilon(h) = 0.$$

Constatons que l'application linéaire $h \mapsto hf'(a)$ est une valeur approchée de l'accroissement $h \mapsto \Delta(h) = f(a+h) - f(a)$ lorsque h est petit. Sous cette forme, la définition de dérivabilité pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ se laisse généraliser aux applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$.

Définition I.1. Une application $f : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ est *différentiable en* $\mathbf{a} \in U$ si f est définie au voisinage de $\mathbf{a}$ et si, pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $\mathbf{a} + \mathbf{h} \in U$, on a :

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\,\boldsymbol{\varepsilon}(\mathbf{h}),$$

où l'application $df(\mathbf{a})\colon \mathbb{R}^n \to \mathbb{R}^p$ est linéaire et l'application $\boldsymbol{\varepsilon}\colon \mathbb{R}^n \to \mathbb{R}^p$ a pour limite 0 quand $\mathbf{h}$ tend vers 0.

Définition I.2. L'application $df(\mathbf{a})$ est appelée *différentielle* de f au point $\mathbf{a}$. L'application f est dite *différentiable sur* un ouvert U si f est différentiable en tout point de U.

Directement à partir des définitions, nous remarquons :
— une application *différentiable en* $\mathbf{a}$ *est continue en* $\mathbf{a}$ (cf. Définition A.18 et Proposition A.25) ;
— *une application constante est différentiable* en tout point et a pour différentielle l'application constante sur 0.

Proposition I.3. *La différentielle, si elle existe, est unique.*

Démonstration. Supposons $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \Phi_{\mathbf{a}}(\mathbf{h}) + ||\mathbf{h}||\,\varepsilon(\mathbf{h}) = \tilde{\Phi}_{\mathbf{a}}(\mathbf{h}) + ||\mathbf{h}||\,\tilde{\varepsilon}(\mathbf{h})$, où $\Phi_{\mathbf{a}}$ et $\tilde{\Phi}_{\mathbf{a}}$ sont des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$ et $\lim_{\mathbf{h}\to 0} \varepsilon(\mathbf{h}) = \lim_{\mathbf{h}\to 0} \tilde{\varepsilon}(\mathbf{h}) = 0$. On en déduit :

$$\lim_{\mathbf{h}\to 0} \frac{\Phi_{\mathbf{a}}(\mathbf{h}) - \tilde{\Phi}_{\mathbf{a}}(\mathbf{h})}{||\mathbf{h}||} = 0\,.$$

Fixons maintenant un élément, non nul et quelconque, $\mathbf{k} \in \mathbb{R}^n$. Si $t \in \mathbb{R}$, on a évidemment $\lim_{t\to 0} t\,\mathbf{k} = 0$, d'où, en utilisant la linéarité de $\Phi_{\mathbf{a}}$ et $\tilde{\Phi}_{\mathbf{a}}$:

$$0 = \lim_{t\to 0} \frac{\left\|\Phi_{\mathbf{a}}(t\mathbf{k}) - \tilde{\Phi}_{\mathbf{a}}(t\mathbf{k})\right\|}{||t\,\mathbf{k}||} = \frac{\left\|\Phi_{\mathbf{a}}(\mathbf{k}) - \tilde{\Phi}_{\mathbf{a}}(\mathbf{k})\right\|}{||\mathbf{k}||}\,.$$

Ce dernier terme étant constant, on a $\Phi_{\mathbf{a}}(\mathbf{k}) = \tilde{\Phi}_{\mathbf{a}}(\mathbf{k})$. L'élément $\mathbf{k}$ ayant été choisi quelconque, les deux applications linéaires $\Phi_{\mathbf{a}}$ et $\tilde{\Phi}_{\mathbf{a}}$ coïncident sur $\mathbb{R}^n$. □

Nous pouvons donc parler de *la différentielle d'une application différentiable*. Pour les applications linéaires, l'*approximation linéaire* coïncide avec l'application de départ.

Proposition I.4. *Si* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ *est linéaire, alors* f *est différentiable en tout point* $\mathbf{a} \in \mathbb{R}^n$ *et* $df(\mathbf{a}) = f$.

Démonstration. Ceci provient de la définition et de l'égalité $f(\mathbf{a}+\mathbf{h})-f(\mathbf{a}) = f(\mathbf{h})+||\mathbf{h}||\cdot 0$, avec $\mathbf{h} \mapsto f(\mathbf{h})$ linéaire. □

Parmi les applications linéaires, relevons les projections canoniques $dx_i : \mathbb{R}^n \to \mathbb{R}$, définies par $dx_i(a_1, \ldots, a_n) = a_i$.

Illustrons par un exemple concret le fait que la différentielle soit une *approximation au premier ordre* de l'accroissement.

Exemple I.5. Considérons une plaque rectangulaire d'aire xy, soumise à la dilatation $(\Delta x\,,\ \Delta y)$. L'accroissement de son aire s'écrit : $(x+\Delta x)(y+\Delta y) - xy = x\,\Delta y + y\,\Delta x + \Delta x\,\Delta y$ et la quantité $x\,\Delta y + y\,\Delta x$ est une approximation au premier ordre de cet accroissement.
En utilisant des notations similaires à celles de la Définition I.1, on a une fonction $S : \mathbb{R}^2 \to \mathbb{R}$, $S(x, y) = xy$. L'égalité précédente s'écrit, en notant $(\Delta x, \Delta y) = (h_1, h_2) = \mathbf{h}$,

$$S((x,y) + (h_1, h_2)) - S(x,y) = x\,h_2 + y\,h_1 + ||\mathbf{h}|| \cdot \frac{h_1\,h_2}{||\mathbf{h}||}.$$

Notons $d\,S(x,y)\,(\mathbf{h}) = x\,h_2 + y\,h_1$ et $\varepsilon(\mathbf{h}) = \dfrac{h_1 h_2}{||h||}$. Alors :

• l'application $d\,S(x,y) \colon \mathbb{R}^2 \to \mathbb{R}$ est linéaire ; c'est la différentielle de la fonction S ;

• de l'inégalité $|\varepsilon(h_1, h_2)| = \left|\dfrac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}}\right| \leq \dfrac{||\mathbf{h}||^2}{||\mathbf{h}||} = ||\mathbf{h}||$, on déduit que la fonction ε a pour limite 0 quand $\mathbf{h}$ tend vers 0. □

Montrons maintenant que l'étude de la différentiabilité des applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$ se ramène à celle des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$.

Proposition I.6. *Soit $f: U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application définie sur un ouvert U de $\mathbb{R}^n$. Notons $f_j : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $j = 1, \ldots, p$, les composantes de f, i.e. $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_p(\mathbf{x}))$ avec $\mathbf{x} \in U$.*
Alors, f est différentiable en $\mathbf{a} = (a_1, \ldots, a_n) \in U$ si, et seulement si, f_j est différentiable en $\mathbf{a}$ pour tout $j = 1, \ldots, p$. Dans ce cas, on a $df(\mathbf{a}) = (df_1(\mathbf{a}), df_2(\mathbf{a}), \ldots, df_p(\mathbf{a}))$.

Démonstration. Supposons f différentiable en $\mathbf{a}$. La projection de l'égalité :
$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\,\varepsilon(\mathbf{h})$$
sur la $j^{\text{ème}}$-composante fournit $f_j(\mathbf{a}+\mathbf{h}) - f_j(\mathbf{a}) = (df(\mathbf{a})(\mathbf{h}))_j + \|\mathbf{h}\|\,(\varepsilon(\mathbf{h}))_j$. L'application $\mathbf{h} \mapsto (df(\mathbf{a})(\mathbf{h}))_j$ est linéaire comme composée des applications linéaires $df(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^p$ et $dx_j : \mathbb{R}^p \to \mathbb{R}$. La fonction $\mathbf{h} \mapsto (\varepsilon(\mathbf{h}))_j$ a pour limite 0 quand $\mathbf{h}$ tend vers 0. Donc la fonction f_j est différentiable en $\mathbf{a}$, de différentielle l'application linéaire $\mathbf{h} \mapsto (df(\mathbf{a})(\mathbf{h}))_j$.
Réciproquement, supposons la fonction f_j différentiable en $\mathbf{a}$ pour tout $j = 1, \ldots, p$. Nous pouvons donc écrire :
$$f_j(\mathbf{a}+\mathbf{h}) - f_j(\mathbf{a}) = df_j(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\,\varepsilon_j(\mathbf{h}),$$
avec $\mathbf{h} \mapsto df_j(\mathbf{a})(\mathbf{h})$ linéaire et $\lim_{\mathbf{h}\to 0} \varepsilon_j(\mathbf{h}) = 0$. Posons :
$$df(\mathbf{a})(\mathbf{h}) = \big(df_1(\mathbf{a})(\mathbf{h}), \ldots, df_p(\mathbf{a})(\mathbf{h})\big) \text{ et } \varepsilon(\mathbf{h}) = \big(\varepsilon_1(\mathbf{h}), \ldots, \varepsilon_p(\mathbf{h})\big).$$
La conclusion provient alors des remarques suivantes :
• l'application $\mathbf{h} \mapsto df(\mathbf{a})(\mathbf{h})$ est linéaire car :
$df(\mathbf{a})(\mathbf{h}+\mathbf{h}') = (df_1(\mathbf{a})(\mathbf{h}+\mathbf{h}'), \ldots, df_p(\mathbf{a})(\mathbf{h}+\mathbf{h}')) = (df_1(\mathbf{a})(\mathbf{h}), \ldots, df_p(\mathbf{a})(\mathbf{h})) +$
$(df_1(\mathbf{a})(\mathbf{h}'), \ldots, df_p(\mathbf{a})(\mathbf{h}')) = df(\mathbf{a})(\mathbf{h}) + df(\mathbf{a})(\mathbf{h}')$
et, de même, $df(\mathbf{a})(\lambda\mathbf{h}) = \lambda df(\mathbf{a})(\mathbf{h})$, pour $\lambda \in \mathbb{R}$,
• l'application $\mathbf{h} \mapsto \varepsilon(\mathbf{h})$ a pour limite 0 quand $\mathbf{h}$ tend vers 0. □

Le lien entre la différentiabilité d'une application f et la différentiabilité de ses applications partielles fait l'objet des deux sections suivantes. Abordons maintenant le comportement de la différentielle à la composition de deux applications.

Théorème I.7. *Considérons une application $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$, différentiable en $\mathbf{a}$, et une application $g : U' \subseteq \mathbb{R}^p \to \mathbb{R}^q$, différentiable en $\mathbf{b} = f(\mathbf{a})$, où U et U' sont des ouverts avec $f(U) \subseteq U'$. Alors, l'application composée $g \circ f : U \subseteq \mathbb{R}^n \to \mathbb{R}^q$ est différentiable en $\mathbf{a}$ et*
$$d(g \circ f)(\mathbf{a}) = dg(\mathbf{b}) \circ df(\mathbf{a}).$$

Démonstration. L'application f étant différentiable en $\mathbf{a}$, nous avons par hypothèse :
$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\,\varepsilon(\mathbf{h}),$$
avec $df(\mathbf{a})$ linéaire et ε de limite 0 lorsque $\mathbf{h}$ tend vers 0. Notons $\mathbf{k} = f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\,\varepsilon(\mathbf{h})$ et remarquons que $\mathbf{k}$ tend vers 0 avec $\mathbf{h}$.
Calculons maintenant $g(f(\mathbf{a}+\mathbf{h})) - g(f(\mathbf{a}))$, en utilisant la différentiabilité de g en $f(\mathbf{a})$ et celle de f en $\mathbf{a}$:
$$\begin{aligned} g(f(\mathbf{a}+\mathbf{h})) - g(f(\mathbf{a})) &= g(f(\mathbf{a})+\mathbf{k}) - g(f(\mathbf{a})) = dg(f(\mathbf{a}))(\mathbf{k}) + \|\mathbf{k}\|\,\varepsilon'(\mathbf{k}) \\ &= dg(f(\mathbf{a}))(df(\mathbf{a})(\mathbf{h})) + dg(f(\mathbf{a}))(\|\mathbf{h}\|\,\varepsilon(\mathbf{h})) + \|\mathbf{k}\|\,\varepsilon'(\mathbf{k}). \end{aligned}$$
Nous voyons apparaître le terme $dg(f(\mathbf{a}))(df(\mathbf{a})(\mathbf{h})) = (dg(f(\mathbf{a})) \circ df(\mathbf{a}))(\mathbf{h})$ annoncé. Il reste à étudier le terme résiduel :
$$\begin{aligned} \mathbf{R} &= dg(f(\mathbf{a}))(\|\mathbf{h}\|\,\varepsilon(\mathbf{h})) + \|df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\,\varepsilon(\mathbf{h})\|\,\varepsilon'(\mathbf{k}) \\ &= \|\mathbf{h}\|\left(dg(f(\mathbf{a}))(\varepsilon(\mathbf{h})) + \left(\left\|\frac{df(\mathbf{a})(\mathbf{h})}{\|\mathbf{h}\|} + \varepsilon(\mathbf{h})\right\|\right)\varepsilon'(\mathbf{k})\right). \end{aligned}$$

D'après la Proposition A.26, la fonction $\mathbf{h} \mapsto \left\| \dfrac{df(\mathbf{a})(\mathbf{h})}{\|\mathbf{h}\|} \right\|$ est bornée et nous pouvons conclure $\lim_{\mathbf{h}\to 0} \dfrac{\mathbf{R}}{\|\mathbf{h}\|} = 0$. □

Les propriétés suivantes se démontrent directement à partir de la définition et de l'unicité de la différentielle. Les preuves sont similaires à celles des résultats correspondants pour les fonctions réelles d'une variable réelle et nous laissons au lecteur l'écriture de leur adaptation.

Proposition I.8. *Si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ et $g : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ sont des applications différentiables définies sur un ouvert U de $\mathbb{R}^n$, alors :*

(i) l'application $f+g$ est différentiable sur U, de différentielle $d(f+g)(\mathbf{x}) = df(\mathbf{x})+dg(\mathbf{x})$;

(ii) pour tout nombre réel λ, l'application λf est différentiable sur U, de différentielle $d(\lambda f)(\mathbf{x}) = \lambda\, df(\mathbf{x})$;

(iii) si $p = 1$, l'application $f \cdot g$ est différentiable sur U, de différentielle $d(f \cdot g)(\mathbf{x}) = f(\mathbf{x}) \cdot dg(\mathbf{x}) + g(\mathbf{x}) \cdot df(\mathbf{x})$;

(iv) si $p = 1$ et si $0 \notin f(U)$, alors l'application $\dfrac{1}{f}$ est différentiable sur U, de différentielle $d\left(\dfrac{1}{f}\right)(\mathbf{x}) = -\dfrac{df(\mathbf{x})}{f^2(\mathbf{x})}$.

I.2 Matrice Jacobienne.

Si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ est une application différentiable sur l'ouvert U, sa différentielle, en tout point $\mathbf{a}$ de U, est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Dans ce paragraphe, nous déterminons la matrice de cette différentielle relativement aux bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$. Elle est appelée *matrice jacobienne* de f et s'exprime en fonction des dérivées partielles premières de f.

Définition I.9. Soit $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ une application définie au voisinage d'un point $\mathbf{a} = (a_1, \ldots, a_n) \in U$. Notons U_i l'ensemble des réels t vérifiant $(a_1, \ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots, a_n) \in U$. L'application f est *dérivable en* $\mathbf{a}$ *par rapport à la variable* x_i, si l'application de $U_i \subseteq \mathbb{R}$ dans $\mathbb{R}^p$, qui à t associe $f(a_1, \ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots, a_n)$, est dérivable en a_i, c'est-à-dire si

$$\lim_{t\to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n)}{t}$$

existe. On note cette limite $\dfrac{\partial f}{\partial x_i}(\mathbf{a})$ et on l'appelle *dérivée partielle (ou DP) première de f en* $\mathbf{a}$ *par rapport à la variable x_i.*

On remarquera, avec la Proposition A.21, que la limite précédente a bien un sens. Si on décompose f suivant ses composantes, $f = (f_1, \ldots, f_p)$, on constate $\dfrac{\partial f}{\partial x_i} = \left(\dfrac{\partial f_1}{\partial x_i}, \ldots, \dfrac{\partial f_p}{\partial x_i}\right)$. Il suffit donc d'étudier les dérivées partielles premières des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. Si f est une application de $\mathbb{R}^n$ dans $\mathbb{R}$, une dérivée partielle première de f est la dérivée d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Les propriétés classiques des fonctions réelles d'une variable réelle ont donc une traduction immédiate dans ce cadre.

Proposition I.10. *Si f, $g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ admettent des dérivées partielles premières en* $\mathbf{a}$*, alors :*
1) la fonction $f+g$ admet des dérivées partielles premières en $\mathbf{a}$ *et :*

$$\frac{\partial (f+g)}{\partial x_i}(\mathbf{a}) = \frac{\partial f}{\partial x_i}(\mathbf{a}) + \frac{\partial g}{\partial x_i}(\mathbf{a})\,;$$

2) pour tout nombre réel λ, la fonction λf admet des dérivées partielles premières en $\mathbf{a}$ *et :*

$$\frac{\partial (\lambda f)}{\partial x_i}(\mathbf{a}) = \lambda \frac{\partial f}{\partial x_i}(\mathbf{a})\,;$$

3) la fonction $f \cdot g$ admet des dérivées partielles premières en $\mathbf{a}$ *et :*

$$\frac{\partial (f \cdot g)}{\partial x_i}(\mathbf{a}) = f(\mathbf{a}) \cdot \frac{\partial g}{\partial x_i}(\mathbf{a}) + \frac{\partial f}{\partial x_i}(\mathbf{a}) \cdot g(\mathbf{a})\,;$$

4) si la fonction f ne s'annule pas dans un voisinage de $\mathbf{a}$*, la fonction $\dfrac{1}{f}$ admet des dérivées partielles premières en* $\mathbf{a}$ *et :*

$$\frac{\partial}{\partial x_i}\left(\frac{1}{f}\right)(\mathbf{a}) = -\frac{1}{f^2(\mathbf{a})} \cdot \frac{\partial f}{\partial x_i}(\mathbf{a})\,.$$

Comme le montre l'exemple suivant, l'existence de dérivées partielles premières est une propriété trop faible pour entraîner la continuité et a fortiori la différentiabilité.

Exemple I.11. Définissons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ par :

$$\begin{cases} f(x,y) & = & \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0)\,; \\ f(0,0) & = & 0. \end{cases}$$

Calculons $\dfrac{\partial f}{\partial x}(0,0) = \lim\limits_{t\to 0} \dfrac{f(t,0)-f(0,0)}{t} = 0$; $\dfrac{\partial f}{\partial y}(0,0) = \lim\limits_{t\to 0} \dfrac{f(0,t)-f(0,0)}{t} = 0$.
La fonction f admet donc des dérivées partielles premières en $(0,0)$ mais n'est pas continue en $(0,0)$, cf. Exemple A.22. □

Nous pouvons maintenant aborder *la détermination de la matrice de l'application linéaire différentielle d'une application*. Soit $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable en $\mathbf{a} \in U$, de composantes $(f_1, \ldots, f_p)$. Nous rapportons les espaces vectoriels $\mathbb{R}^n$ et $\mathbb{R}^p$ à leur base canonique respective $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$.
Par définition, la matrice de l'application linéaire $df(\mathbf{a})$ a en colonnes les composantes des vecteurs $df(\mathbf{a})(\mathbf{e}_i)$ dans la base $(\mathbf{u}_j)_{1\leq j\leq p}$; il nous faut donc calculer la $j^{\text{ème}}$ composante de $df(\mathbf{a})(\mathbf{e}_i)$. On a déjà remarqué dans le paragraphe précédent que la $j^{\text{ème}}$ composante de la différentielle de f est égale à la différentielle de f_j, cf. Proposition I.6. Nous avons ainsi ramené le problème à celui d'une fonction à valeurs dans $\mathbb{R}$.
Considérons une fonction $g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ différentiable en $\mathbf{a} \in U$. Soit $t \in \mathbb{R}$ tel que $\mathbf{a} + t\mathbf{e}_i \in U$, on a :

$$g(\mathbf{a} + t\mathbf{e}_i) - g(\mathbf{a}) = dg(\mathbf{a})(t\mathbf{e}_i) + |t|\,\varepsilon(t\mathbf{e}_i) = t\,dg(\mathbf{a})(\mathbf{e}_i) + |t|\,\varepsilon(t\mathbf{e}_i),$$

avec $\lim_{t\to 0} \varepsilon(t\mathbf{e}_i) = 0$. La fonction de $\mathbb{R} - \{0\}$ dans $\mathbb{R}$, définie par $t \mapsto \dfrac{|t|}{t}$, étant bornée, on a $\lim\limits_{t\to 0} \dfrac{|t|\,\varepsilon(te_i)}{t} = 0$. On en déduit :

$$dg(\mathbf{a})(\mathbf{e}_i) = \lim_{t\to 0} \frac{g(\mathbf{a} + t\mathbf{e}_i) - g(\mathbf{a})}{t} = \frac{\partial g}{\partial x_i}(\mathbf{a})\,.$$

Nous avons ainsi déterminé la matrice associée à la différentielle.

Proposition I.12. *Soit $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable en $\mathbf{a} \in U$. Sa différentielle $df(\mathbf{a})$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$, ayant pour matrice relativement aux bases canoniques :*

$$J(f, \mathbf{a}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{a}) & \dots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(\mathbf{a}) & \dots & \dfrac{\partial f_p}{\partial x_n}(\mathbf{a}) \end{pmatrix}.$$

Cette matrice est appelée matrice jacobienne de f en $\mathbf{a}$. *Si $\mathbf{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$ et $\mathbf{k} = (k_1, \dots, k_p) \in \mathbb{R}^p$, l'égalité $df(\mathbf{a})(\mathbf{h}) = \mathbf{k}$ correspond au produit matriciel :*

$$J(f, \mathbf{a}) \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_p \end{pmatrix}.$$

Dans le cas particulier d'une fonction $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, différentiable en $\mathbf{a} \in U$, la matrice de $df(\mathbf{a})(\mathbf{h})$ est une matrice ligne :

$$J(f, \mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}) \dots \frac{\partial f}{\partial x_n}(\mathbf{a}) \right),$$

d'où :

$$df(\mathbf{a})(\mathbf{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, h_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, dx_i(\mathbf{h}).$$

L'égalité ci-dessus fournit une décomposition de $df(\mathbf{a})$ dans la base duale $(dx_1, \dots, dx_n)$ de la base canonique de $\mathbb{R}^n$:

$$df(\mathbf{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, dx_i.$$

Si f est différentiable en tout point de U, on écrit aussi :

$$\boxed{df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dx_i.}$$

En particulier, pour une fonction différentiable $f : \mathbb{R}^3 \to \mathbb{R}$, on a :

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz\,.$$

Le Théorème I.7 et la Proposition I.12 impliquent immédiatement :

Corollaire I.13. *Avec les notations et hypothèses du Théorème I.7, la matrice de la différentielle $d(g \circ f)(\mathbf{a})$ de l'application composée $g \circ f$ est le produit des deux matrices associées à $dg(f(\mathbf{a}))$ et $df(\mathbf{a})$:*

$$J(g \circ f, \mathbf{a}) = J(g, f(\mathbf{a})).J(f, \mathbf{a})\,.$$

On en déduit, en notant $g = (g_1, \dots, g_q)$, $f = (f_1, \dots, f_p)$ les composantes de g et f :

$$\frac{\partial \left(g_j \circ f \right)}{\partial x_i}(\mathbf{a}) = \sum_{s=1}^{p} \frac{\partial g_j}{\partial y_s}(f(\mathbf{a}))\ \frac{\partial f_s}{\partial x_i}(\mathbf{a})\,.$$

En particulier, si $n = p = q = 1$, on retrouve la formule connue pour les fonctions réelles d'une variable réelle : $(g \circ f)'(x) = g'(f(x))f'(x)$.

Terminons ce paragraphe par quelques calculs explicites.

Exemple I.14. Soit $U = \{(x,y) \in \mathbb{R}^2 | \ x > 0\}$. Définissons une application $f : U \to \mathbb{R}^2$ par $f(x,y) = (f_1(x,y), f_2(x,y)) = \left(x^y, \ \sin x + xy^3\right)$. Un simple calcul donne la matrice jacobienne de l'application f :

$$J(f,(x,y)) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x}(x,y) & \dfrac{\partial f_1}{\partial y}(x,y) \\ \dfrac{\partial f_2}{\partial x}(x,y) & \dfrac{\partial f_2}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} yx^{y-1} & x^y \ln x \\ \cos x + y^3 & 3xy^2 \end{pmatrix}.$$ □

Comme l'illustre l'exemple suivant, la détermination de la différentielle de f en un point $\mathbf{a}$ permet d'obtenir des valeurs approchées de f au voisinage de $\mathbf{a}$.

Exemple I.15. La différentielle de la fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = \sin x \cos y$ est :

$$df(x,y) = \cos x \ \cos y \ dx - \sin x \ \sin y \ dy.$$

Elle fournit une valeur approchée de $\sin 0{,}01 \ \cos 0{,}02$, donnée par :

$$f(0,0) + df((0,0))(0{,}01, \ 0{,}02) = 1 \times 1 \times 0{,}01 + 0 = 0{,}01.$$

Pour majorer l'erreur commise, il nous faut connaître l'importance du terme négligé. Ceci sera possible avec les formules de Taylor introduites dans la Section I.8. □

I.3 Continuité, différentiabilité et dérivées partielles.

Avec ce qui précède, nous pouvons calculer une différentielle si elle existe mais nous ne possédons pas de critère pratique pour savoir si une application est différentiable. Le théorème suivant nous en fournit un ; il compare également la différentiabilité, la continuité et l'existence de dérivées partielles premières.

Théorème I.16. *Soit $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ une application définie au voisinage de $\mathbf{a} \in U$. Les relations entre continuité, différentiabilité et existence de dérivées partielles premières se résument en :*

(1) f admet des DP premières dans un voisinage de $\mathbf{a}$, continues en $\mathbf{a}$.

$$\Downarrow$$

(2) f est différentiable en $\mathbf{a}$.

$$\Downarrow \qquad\qquad\qquad \Downarrow$$

(3) f est continue en $\mathbf{a}$. *(4) f admet des DP premières en $\mathbf{a}$.*

L'existence de dérivées partielles premières continues entraîne donc la différentiabilité. Ceci justifie la définition suivante.

Définition I.17. Une application $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ est *de classe C^1* sur l'ouvert U si les dérivées partielles premières $\dfrac{\partial f}{\partial x_i}$ existent et sont continues sur U, pour tout $i = 1, \ldots, n$. Notons $\mathcal{C}^1(U, \mathbb{R}^p)$ l'ensemble des applications de classe C^1 sur U à valeurs dans $\mathbb{R}^p$.

Démonstration du Théorème I.16. (1) $\Rightarrow$ (2). Il suffit (cf. Proposition I.6) de faire la démonstration dans le cas particulier d'une fonction $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. L'idée est d'utiliser le théorème des accroissements finis pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Ecrivons la preuve d'abord pour $n = 2$, avec $\mathbf{h} = (h_1, h_2)$, $\mathbf{a} = (a_1, a_2)$. En appliquant le théorème des accroissements finis à $f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2)$ et à $f(a_1, a_2 + h_2) - f(a_1, a_2)$, nous obtenons deux nombres réels θ_1 et θ_2, compris entre 0 et 1, tels que :

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) = h_1 \frac{\partial f}{\partial x_1}(a_1 + \theta_1 h_1, a_2 + h_2) + h_2 \frac{\partial f}{\partial x_2}(a_1, a_2 + \theta_2 h_2).$$

La continuité des dérivées partielles premières en (a_1, a_2) entraîne l'existence de fonctions ε et ε', de limite 0 en 0, telles que :

$$\begin{aligned} f(a_1+h_1, a_2+h_2) - f(a_1, a_2) &= h_1\left(\frac{\partial f}{\partial x_1}(a_1,a_2) + \varepsilon(\mathbf{h})\right) + h_2\left(\frac{\partial f}{\partial x_2}(a_1,a_2) + \varepsilon'(\mathbf{h})\right) \\ &= h_1\frac{\partial f}{\partial x_1}(a_1,a_2) + h_2\frac{\partial f}{\partial x_2}(a_1,a_2) + h_1\varepsilon(\mathbf{h}) + h_2\varepsilon'(\mathbf{h}). \end{aligned}$$

On en déduit le résultat annoncé en utilisant :

$$\frac{|h_1\varepsilon(\mathbf{h}) + h_2\varepsilon'(\mathbf{h})|}{||\mathbf{h}||} \leq \frac{|h_1|}{||\mathbf{h}||}|\varepsilon(\mathbf{h})| + \frac{|h_2|}{||\mathbf{h}||}|\varepsilon'(\mathbf{h})| \leq |\varepsilon(\mathbf{h})| + |\varepsilon'(\mathbf{h})|\,.$$

Dans le cas général, il faut considérer :

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) &= \sum_{i=1}^{n} f(a_1, \cdots, a_{i-1}, a_i + h_i, a_{i+1} + h_{i+1}, \cdots, a_n + h_n) - \\ &\qquad f(a_1, \cdots, a_{i-1}, a_i, a_{i+1} + h_{i+1}, \cdots, a_n + h_n) \\ &= \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a_1, \cdots, a_{i-1}, a_i + \theta_i h_i, a_{i+1} + h_{i+1}, \cdots, a_n + h_n). \end{aligned}$$

Le raisonnement se déroule ensuite de façon similaire à celui du cas particulier $n = 2$.

La démonstration de (2) $\Rightarrow$ (4) a été faite au cours du calcul de la matrice jacobienne, cf. Proposition I.12. L'implication (2) $\Rightarrow$ (3) est immédiate et a déjà été remarquée. □

Dans le Théorème I.16, toutes **les autres implications sont fausses** en général, comme le montrent les contre-exemples suivants.

(2) $\not\Rightarrow$ (1) : La fonction $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 \sin(1/x)$ est dérivable sur $\mathbb{R}$, à dérivée non continue en 0.

(3) $\not\Rightarrow$ (4) : (d'où (3) $\not\Rightarrow$ (2) et (3) $\not\Rightarrow$ (1)) La fonction "valeur absolue" de $\mathbb{R}$ dans $\mathbb{R}$ est continue sur $\mathbb{R}$ et non dérivable en 0.

(4) $\not\Rightarrow$ (3) : (d'où (4) $\not\Rightarrow$ (2) et (4) $\not\Rightarrow$ (1)) L'Exemple I.11 convient.

I.4 Applications à différentielle nulle.

Soit $f : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application différentiable sur un ouvert U de $\mathbb{R}^n$. À partir de la décomposition de $df(\mathbf{x})$ (cf. Proposition I.6) et de son calcul (cf. Proposition I.12), on déduit l'équivalence des trois propriétés suivantes :

(*i*) $df(\mathbf{x}) = 0, (\forall \mathbf{x} \in U)$;

(*ii*) $df_j(\mathbf{x}) = 0, (\forall \mathbf{x} \in U), (\forall j = 1, \ldots p)$;

(*iii*) $\dfrac{\partial f_j}{\partial x_i}(\mathbf{x}) = 0, (\forall \mathbf{x} \in U), (\forall j = 1, \ldots p), (\forall i = 1, \ldots, n)$.

Nous avons déjà constaté que si f est une application constante sur U, sa différentielle df est nulle sur U. Notons que la réciproque est fausse, même pour les fonctions réelles d'une variable réelle.