

Calcul différentiel

I. Différentielle et dérivées partielles

I.1. Position du problème

On étudie des fonctions f définies sur un ouvert U d'un \mathbb{R} -espace vectoriel E de dimension finie n , à valeurs dans un \mathbb{R} -espace vectoriel F de dimension finie p .

- Si on a choisi des bases $\mathcal{B} = (e_1, \dots, e_n)$ de E et $\mathcal{C} = (u_1, \dots, u_p)$ de F , on peut
- définir les fonctions coordonnées f_i de f par $f(x) = \sum_{i=1}^p f_i(x)u_i$: ce sont alors des fonctions de $U \subset E$ dans \mathbb{R} .
 - éventuellement, identifier chaque vecteur $x = \sum x_i e_i$ au n -uplet (x_1, \dots, x_n) de ses coordonnées, donc identifier f à la fonction φ de $V \subset \mathbb{R}^n$ dans \mathbb{R}^p définie par $\varphi(x_1, \dots, x_n) = (y_1, \dots, y_p)$ si et seulement si $f(\sum x_i e_i) = \sum y_j \varepsilon_j$.

On peut donc supposer que f est une application définie sur un ouvert U de \mathbb{R}^n , à valeurs dans \mathbb{R}^p .

I.2. Dérivée suivant un vecteur

Définition. Soient $a \in U$ et $h \in E$. On dit que f est **dérivable suivant le vecteur h** au point a , si l'application $\varphi : t \in \mathbb{R} \mapsto f(a + th)$ est dérivable en 0. Dans ce cas, le vecteur $\varphi'(0)$ est appelé **vecteur dérivé** suivant h de f en a , et noté $D_h f(a)$; et on a en 0 le DL_1 $f(a + th) = f(a) + tD_h f(a) + o(t)$.

Si f est dérivable suivant h en tout point $a \in V$, l'application $D_h f : a \mapsto D_h f(a)$ est appelée **application dérivée** suivant h de f .

Proposition I.1. Soit $\mathcal{C} = (u_1, \dots, u_q)$ une base de F . Pour tout $x \in U$, posons $f(x) = \sum_{i=1}^q f_i(x)u_i$. L'application f est dérivable suivant h en a si et seulement si chaque fonction coordonnée f_i l'est; on a alors $D_h f(a) = \sum_{i=1}^q D_h f_i(a)u_i$.

Définition. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Si f est dérivable en a suivant le i -ème vecteur e_i de \mathcal{B} (respectivement sur U), on dit que f est dérivable par rapport à la i -ème variable en a (respectivement sur U).

L'application $D_{e_i} f$ est alors appelée **dérivée partielle** de f par rapport à la i -ème variable; elle est notée $\partial_i f$, ou $\partial f / \partial x_i$ s'il n'y a pas d'ambiguïté sur le nom x_i donné à la i -ème variable.

Définition. Supposons E et F munis de bases \mathcal{B} et \mathcal{C} . Si f admet des dérivées partielles par rapport à chaque variable en a , alors la matrice $B = (b_{ij}) \in \mathcal{M}_{p,n}(\mathbb{R})$ définie par $b_{ij} = \frac{\partial f_i}{\partial x_j}(a)$ est appelée **matrice jacobienne** de f en a . Pour

chaque j , la colonne j de cette matrice contient les coordonnées dans \mathcal{C} du vecteur $D_{e_j} f(a) = \frac{\partial f}{\partial x_j}(a)$.

I.3. Applications différentiables

Définition. Soit $a \in U$. On dit que f est **différentiable** en a s'il existe une boule ouverte $B(a, r) \subset U$, une application **linéaire** $L \in \mathcal{L}(E, F)$ et une fonction ε de $B(0, r)$ dans F telles que

$$\forall h \in B(0, r) \quad f(a + h) = f(a) + L(h) + \|h\|\varepsilon(h) \quad \text{et} \quad \varepsilon(h) \xrightarrow{h \rightarrow 0} 0$$

On dit alors que le terme $\|h\|\varepsilon(h)$ est négligeable devant $\|h\|$ en 0; on écrit $\|h\|\varepsilon(h) = o(h)$. La relation $f(a + h) = f(a) + L(h) + o(h)$ est appelée un développement limité à l'ordre 1 de f en a .

Proposition I.2. Si f est différentiable en a , alors elle y est continue.

Proposition I.3. Si f est différentiable en a , alors elle y est dérivable suivant tout vecteur h , et, avec les notations précédentes, $D_h f(a) = L(h)$.

Corollaire I.4. Si f est différentiable en a , alors l'application linéaire L apparaissant dans son développement limité est unique; elle est appelée **différentielle** de f en a , et notée $df(a)$.

Par souci de lisibilité, l'image d'un vecteur u par $df(a)$ sera notée $df(a) \cdot u$ au lieu de $df(a)(u)$.

Si f est différentiable en a , alors $\frac{\partial f}{\partial x_j}(a) = df(a) \cdot e_j$ pour tout $j \in \llbracket 1, n \rrbracket$; et, si $h = \sum h_j e_j$, $df(a) \cdot h = \sum h_j \frac{\partial f}{\partial x_j}(a)$. La matrice jacobienne de f en a est alors la matrice de $df(a)$.

Définition. Si f est différentiable en chaque point de U , on dit que f est **différentiable** sur U ; l'application $df : a \in U \mapsto df(a) \in \mathcal{L}(E, F)$ est alors appelée **différentielle** de f .

I.4. Cas d'une fonction numérique

On suppose ici E euclidien.

Définition. Si f est différentiable en a , alors sa différentielle $df(a)$ est une forme linéaire sur E . L'unique vecteur u_0 vérifiant $df(a) \cdot x = (u_0 | x)$ pour tout $x \in \mathbb{R}^n$, est appelé **gradient** de f en a , et noté $\nabla f(a)$; ses coordonnées dans une base orthonormée sont les nombres $(\partial f / \partial x_j)(a)$.

Le développement limité à l'ordre 1 de f en a s'écrit alors

$$f(a+h) = f(a) + (\nabla f(a) | h) + o(\|h\|)$$

II. Fonctions de classe C^1

II.1. Généralités

Définition. On dit que f est de classe C^1 sur U si f est différentiable en tout point de U , et si l'application $df : U \rightarrow \mathcal{L}(E, F)$, $a \mapsto df(a)$ est continue sur U .

Théorème II.1. L'application f est de classe C^1 sur U si et seulement s'il existe une base dans laquelle :

- f admet des dérivées partielles par rapport à chaque coordonnée ;
- chaque dérivée partielle est continue sur U .

II.2. Combinaisons linéaires et produit

Proposition II.2. Soient f et g deux applications de $U \subset \mathbb{R}^n$ dans \mathbb{R}^p . Si f et g sont différentiables en un point $a \in U$, et si $(\lambda, \mu) \in \mathbb{R}^2$, alors $\lambda f + \mu g$ est différentiable en a , et $d[\lambda f + \mu g](a) = \lambda df(a) + \mu dg(a)$.

En particulier, si f et g sont à valeurs dans \mathbb{R} , alors $\nabla[\lambda f + \mu g](a) = \lambda \nabla f(a) + \mu \nabla g(a)$.

Corollaire II.3. Si f et g sont de classe C^1 sur U , alors $\lambda f + \mu g$ l'est aussi.

Proposition II.4. Soient $f : U \subset E \rightarrow F$ et $g : U \rightarrow G$; soit B une application bilinéaire de $F \times G$ dans H . Si f et g sont différentiables en $a \in U$, alors $B(f, g)$ l'est aussi, et, pour tout $h \in E$,

$$d[B(f, g)](a) \cdot h = B(df(a) \cdot h, g(a)) + B(f(a), dg(a) \cdot h)$$

En particulier, si f et g sont toutes deux à valeurs réelles, alors

$$\nabla[fg](a) = g(a)\nabla f(a) + f(a)\nabla g(a)$$

Corollaire II.5. Avec les notations précédentes, si f et g sont de classe C^1 sur U , alors $B(f, g)$ l'est aussi.

II.3. Composition

Théorème II.6. Soient $f : U \subset E \rightarrow F$ et $g : V \subset F \rightarrow G$, vérifiant $f(U) \subset V$. Si f est différentiable en $a \in U$ et g l'est en $f(a)$, alors $g \circ f$ est différentiable en a et $d[g \circ f](a) = [dg(f(a))] \circ [df(a)]$.

Corollaire II.7. Avec les notations précédentes, si f et g sont de classe C^1 sur U et V respectivement, alors $g \circ f$ est de classe C^1 sur U .

Corollaire II.8. Soit $\varphi : I \subset \mathbb{R} \rightarrow \mathbb{R}^p$, $t \mapsto \varphi(t) = (\varphi_1(t), \dots, \varphi_p(t))$; soit $g : V \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$, vérifiant $\varphi(I) \subset V$. Si φ est dérivable en $a \in I$ et si g est différentiable en $\varphi(a)$, alors $g \circ \varphi$ est dérivable en a , et

$$\begin{aligned} [g \circ \varphi]'(a) &= \frac{d}{dt} [g(\varphi_1(t), \dots, \varphi_p(t))]_{t=a} = \sum_{k=1}^p \varphi'_k(a) \frac{\partial g}{\partial x_k}(\varphi(a)) \\ &= [dg(\varphi(a))] \cdot \varphi'(a) \end{aligned}$$

Corollaire II.9. Soit U un ouvert **connexe par arcs** de E , et f une application différentiable sur U . Alors, f est constante sur U si et seulement si $df(a) = 0$ pour tout $a \in U$.

Corollaire II.10. Avec les hypothèses et notations du théorème II.6, la matrice jacobienne de $g \circ f$ en a est le produit de celle de g en $f(a)$ par celle de f en a . En particulier, en notant f_1, \dots, f_p et g_1, \dots, g_q les fonctions coordonnées de f et g respectivement, on a pour tout couple (i, j) :

$$\frac{\partial [g_i \circ f]}{\partial x_j}(a) = \sum_{k=1}^p \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)$$

III. Optimisation à l'ordre 1

III.1. Points critiques

Définition. Soit $f : A \subset E \rightarrow \mathbb{R}$, où A est une partie quelconque de E . On dit que f admet un maximum local en $a \in A$, s'il existe $r > 0$ tel que $\forall x \in B(a, r) \cap A$ $f(x) \leq f(a)$; on dit que f admet un extremum local en a , si elle y admet un maximum ou un minimum local.

Proposition III.1. Si f admet un extremum local en un point a de l'ouvert U et est différentiable en a , alors $df(a) = 0$; en particulier, si E est euclidien, alors $\nabla f(a) = 0$.

Définition. Un point en lequel f est différentiable et de gradient nul est appelé un **point critique** de f .

III.2. Vecteurs tangents à une partie

Définition. On appelle **arc paramétré** de classe C^k dans E , toute application γ définie sur un intervalle I de \mathbb{R} , à valeurs dans E , et de classe C^k sur I , c'est-à-dire toute fonction vectorielle de classe C^k .

Si $X \subset E$ et $\gamma(t) \in X$ pour tout $t \in I$, on dit que l'arc est tracé sur X .

Définition. Soient $X \subset E$, $a \in X$ et $u \in E$. On dit que u est un vecteur tangent à X en a s'il existe $\varepsilon \in \mathbb{R}_+^*$ et un arc $\gamma :]-\varepsilon, \varepsilon[\rightarrow X$, dérivable en 0, tel que $\gamma(0) = a$ et $\gamma'(0) = u$.

L'ensemble des vecteurs tangents en a à X est noté $T_a X$.

Remarque : si $u \in T_a X$, alors, pour tout $\lambda \in \mathbb{R}$, $\lambda u \in T_a X$.

Proposition III.2. Soit $g : U \rightarrow \mathbb{R}$, définie et de classe C^1 sur U . Soit $X = g^{-1}(\{0\})$; on suppose trouvé $a \in X$ tel que $dg(a) \neq 0$. Alors $T_a X = \text{Ker}(dg(a))$. En particulier, si E est euclidien, alors $T_a X = (\nabla f(a))^\perp$.

III.3. Extremums liés

Proposition III.3. Soit $f : U \rightarrow \mathbb{R}$. Soit $X \subset U$; soit $a \in X$, tel que f soit différentiable en a .

Si la restriction de f à X admet un extremum local en a , alors $T_a X \subset \text{Ker} df(a)$; si E est euclidien, on a donc $T_a X \subset (\nabla f(a))^\perp$.

Proposition III.4. Soient f et $g : U \rightarrow \mathbb{R}$, de classe C^1 sur U . Soit $X = g^{-1}(\{0\})$; soit $a \in X$, tel que $dg(a) \neq 0$.

Si la restriction de f à X admet un extremum local en a , alors $dg(a)$ et $df(a)$ sont colinéaires; si E est euclidien, les vecteurs $\nabla g(a)$ et $\nabla f(a)$ sont donc colinéaires.

IV. Fonctions de classe C^k

IV.1. Généralités

Définition. Soit $k \geq 1$. On dit que la fonction f est de classe C^k sur U si elle est de classe C^1 sur U et ses dérivées partielles sont de classe C^{k-1} sur U .

Théorème IV.1 (de Schwartz). Si f est de classe C^2 sur U , alors, pour tout couple (i, j) , $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$.

Proposition IV.2. Une combinaison linéaire, une composée de fonctions de classe C^k , est encore de classe C^k .

Si f et g sont de classe C^k et si B est bilinéaire, alors $B(f, g)$ est de classe C^k .

IV.2. Matrice hessienne

Définition. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^2 ; soit $a \in U$. On appelle **matrice hessienne**, ou simplement **hessienne** de f en a , la matrice

$H_f(a) \in \mathcal{M}_n(\mathbb{R})$ définie par $\forall (i, j) \quad [H_f(a)]_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$.

Le théorème de Schwartz montre que c'est une matrice symétrique.

Théorème IV.3 (Formule de Taylor-Young à l'ordre 2). Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^2 ; soit $a \in U$. Alors, au voisinage du vecteur nul :

$$f(a+h) = f(a) + (\nabla f(a) | h) + \frac{1}{2} (h | H_f(a) h) + o(\|h\|^2)$$

(où l'on identifie $h \in \mathbb{R}^n$ à la colonne correspondante pour le calcul de $H_f(a)h$).

IV.3. Optimisation à l'ordre 2

Théorème IV.4 (Condition nécessaire d'extremum). Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^2 ; soit $a \in U$. Si f admet en a un minimum local, alors $\nabla f(a) = 0$ et $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$; si f admet en a un maximum local, alors $\nabla f(a) = 0$ et $-H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$.

Théorème IV.5 (Condition suffisante d'extremum). Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^2 ; soit $a \in U$. Si $\nabla f(a) = 0$ et $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f admet en a un minimum local; si $\nabla f(a) = 0$ et $-H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f admet en a un maximum local.

Remarque : si $n = 2$ et a est un point critique de f , on obtient le signe des deux valeurs propres de $H_f(a)$ en examinant $\det H_f(a)$ (leur produit) et $\text{tr} H_f(a)$ (leur somme). Plus précisément :

- ▷ si $\det H_f(a) < 0$, alors f n'a pas d'extremum en a ;
- ▷ si $\det H_f(a) > 0$ et $\text{tr} H_f(a) > 0$, alors f a un minimum local en a ;
- ▷ si $\det H_f(a) > 0$ et $\text{tr} H_f(a) < 0$, alors f a un maximum local en a .

Si $\det H_f(a) = 0$, alors le DL₂ ne suffit pas pour conclure.