28

Calcul différentiel et optimisation

Dans tout le chapitre, E et F désignent des $\mathbb{R}$-espaces vectoriels normés de dimension finie nulle, $\mathscr{U}$ désigne un ouvert de E.

I Dérivées partielles

Dans cette partie, $\mathscr{U}$ est un ouvert non vide de $\mathbb{R}^p$, où $p \in \mathbb{N}^*$.

1 Compléments sur la continuité des fonctions de variable vectorielle

Définition 1 : Application partielle

Soit $f : \mathscr{U} \to \mathbb{R}$ une fonction, $a = (a_1, \dots, a_p) \in \mathscr{U}$, $j \in [\![1,p]\!]$. On appelle j^e **application partielle de f en** a l'application $t \mapsto f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_p)$.

Lemme 1 : « de partition »

Soit A une partie de E, $f : A \to F$, B_1, B_2 deux parties de A telles que $B_1 \cup B_2 = A$, $a \in \overline{B_1} \cap \overline{B_2}$, $\ell \in F$.

Si $f_{|B_1}(x) \xrightarrow[x\to a]{} \ell$ et $f_{|B_2}(x) \xrightarrow[x\to a]{} \ell$, alors $f(x) \xrightarrow[x\to a]{} \ell$.

En particulier, si $a \in A$, f est continue en a.

2 Dérivées partielles

$\mathscr{U}$ est toujours un ouvert de $\mathbb{R}^p$.

Définition 2 : Dérivées partielles

Soit $f : \mathscr{U} \to \mathbb{R}$, $a = (a_1, \dots, a_p) \in \mathscr{U}$, $j \in [\![1,p]\!]$. On appelle j^e **dérivée partielle de f en** a, lorsqu'elle existe, la dérivée de la j^e application partielle de f en a. On note $\partial_j f(a) = \dfrac{\partial f}{\partial x_j}(a)$ le nombre dérivé en ce point.

On appelle j^e **dérivée partielle** la fonction définie sur $\mathscr{U}$ par $a \mapsto \dfrac{\partial f}{\partial x_j}(a)$.

Définition 3 : Vecteur gradient

Soit $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles en $a \in \mathscr{U}$. On appelle **gradient** de f en a le vecteur

$$\nabla f(a) = \overrightarrow{\text{grad}}\, f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a)\right) \in \mathbb{R}^p.$$

Définition 4 : Extension aux fonctions à valeurs vectorielles

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^p$ et $f : \mathscr{U} \to \mathbb{R}^n$. On note $f = (f_1, \dots, f_n)$: f_i est la i^e composante de f.

On dit que f admet des dérivées partielles en $a \in \mathscr{U}$ si chacune des f_i admet une dérivée partielle en a.

On appelle alors j^e **dérivée partielle de f en** a le vecteur

$$\frac{\partial f}{\partial x_j}(a) = \left(\frac{\partial f_1}{\partial x_j}(a), \dots, \frac{\partial f_n}{\partial x_j}(a)\right).$$

Définition 5 : Matrice jacobienne

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^p$ et $f : \mathscr{U} \to \mathbb{R}^n$. On note $f = (f_1, \dots, f_n)$. On appelle, lorsque existe, **matrice jacobienne de f en** a la matrice

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} \in \mathscr{M}_{n,p}(\mathbb{R})$$

3 Fonctions de classe $\mathscr{C}^1$

Définition 6 : Classe $\mathscr{C}^1$

Soit $f : \mathscr{U} \to \mathbb{R}^n$. On dit que f est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ lorsqu'en tout point de $\mathscr{U}$, les dérivées partielles de f existent, et que ces dérivées partielles sont continues.

On note $\mathscr{C}^1(\mathscr{U}, \mathbb{R}^n)$ l'ensemble de ces fonctions.

Propriété 1 : Équivalence avec les fonctions coordonnées

On note $f = (f_1, \dots, f_n)$. Alors f est de classe $\mathscr{C}^1$ si et seulement si chacune des f_i l'est.

Propriété 2 : Structure d'algèbre

Pour tout $n \in \mathbb{N}$, $\mathscr{C}^1(\mathscr{U}, \mathbb{R}^n)$ est un $\mathbb{R}$-espace vectoriel et $\mathscr{C}^1(\mathscr{U}, \mathbb{R})$ est une $\mathbb{R}$-algèbre.

Théorème 1 : DL_1

Soit $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^1$ et $a \in \mathscr{U}$. Pour tout $h \in \mathbb{R}^p$ tel que $a + h \in \mathscr{U}$,

$$\begin{aligned} f(a+h) &= f(a) + h_1 \frac{\partial f}{\partial x_1}(a) + \cdots + h_p \frac{\partial f}{\partial x_p}(a) + \underset{h\to 0}{\mathrm{o}}(\|h\|) \\ &= f(a) + \big(\nabla f(a) \big| h\big) + \underset{h\to 0}{\mathrm{o}}(\|h\|). \end{aligned}$$

en utilisant le produit scalaire canonique sur $\mathbb{R}^p$.

CC BY NC SA

Corollaire 1 : $\mathscr{C}^1 \Longrightarrow$ continue

Une fonction de classe $\mathscr{C}^1$ est continue.

Propriété 3 : Règle de la chaîne

Soient $\mathscr{U},\mathscr{V}$ deux ouverts respectifs de $\mathbb{R}^p$ et $\mathbb{R}^n$ et $a \in \mathscr{U}$. Soient

$$f : \left| \begin{array}{ccc} \mathscr{U} & \longrightarrow & \mathscr{V} \\ (x_1,\dots,x_p) & \longmapsto & (f_1(x_1,\dots,x_p),\dots,f_n(x_1,\dots,x_p)) \end{array} \right.$$

et

$$g : \left| \begin{array}{ccc} \mathscr{V} & \longrightarrow & \mathbb{R} \\ (y_1,\dots,y_n) & \longmapsto & g(y_1,\dots,y_n) \end{array} \right.$$

deux fonctions de classe $\mathscr{C}^1$. Alors $g\circ f = g\circ(f_1,\dots,f_n)$ est de classe $\mathscr{C}^1$ et pour $j \in [\![1,p]\!]$,

$$\frac{\partial(g\circ f)}{\partial x_j}(a) = \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(f(a))\frac{\partial f_k}{\partial x_j}(a)$$

Cas particulier 1 : important – dérivée le long d'un arc

Si $\gamma : \mathbb{R} \to \mathbb{R}^3$ et $f : \mathbb{R}^3 \to \mathbb{R}$ sont de classe $\mathscr{C}^1$, en notant, pour $t \in \mathbb{R}$, $\gamma(t) = (x(t), y(t), z(t))$,

$$(f\circ\gamma)'(t) = x'(t)\frac{\partial f}{\partial x}(\gamma(t)) + y'(t)\frac{\partial f}{\partial y}(\gamma(t)) + z'(t)\frac{\partial f}{\partial z}(\gamma(t)).$$

Corollaire 2 : Règle de la chaîne vectorielle

Soient $\mathscr{U},\mathscr{V}$ deux ouverts respectifs de $\mathbb{R}^p$ et $\mathbb{R}^n$ et $a \in \mathscr{U}$. Soient $f : \mathscr{U} \to \mathscr{V}$ et $g : \mathscr{V} \to \mathbb{R}^q$ deux fonctions de classe $\mathscr{C}^1$. Alors $g\circ f$ est de classe $\mathscr{C}^1$ et pour $j \in [\![1,p]\!]$ et $\ell \in [\![1,q]\!]$,

$$\frac{\partial(g_\ell\circ f)}{\partial x_j}(a) = \frac{\partial(g\circ f)_\ell}{\partial x_j}(a) = \sum_{k=1}^{n} \frac{\partial g_\ell}{\partial y_k}(f(a))\frac{\partial f_k}{\partial x_j}(a)$$

En particulier,

$$J_{g\circ f}(a) = J_g(f(a))J_f(a).$$

4 Fonctions de classe $\mathscr{C}^k$

Définition 7 : Dérivées partielles d'ordre supérieur

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^p$, $f : \mathscr{U} \to \mathbb{R}^n$. On appelle **dérivée partielle d'ordre** $k \in \mathbb{N}^*$, une dérivée $\dfrac{\partial \varphi}{\partial x_j}$ où φ est une dérivée partielle d'ordre $k-1$ de f. Les dérivées partielles d'ordre k sont de la forme

$$\frac{\partial}{\partial x_{i_k}}\left(\frac{\partial}{\partial x_{i_{k-1}}}\left(\cdots\frac{\partial}{\partial x_{i_2}}\left(\frac{\partial f}{\partial x_{i_1}}\right)\cdots\right)\right)$$

notées $\dfrac{\partial^k f}{\partial x_{i_k}\cdots\partial x_{i_1}} = \partial_{i_k,\dots,i_1} f$ où $i_1,\dots,i_k \in [\![1,p]\!]$.

Définition 8 : Classe $\mathscr{C}^k$

Soit $k \in \mathbb{N}^*$, $\mathscr{U}$ ouvert de $\mathbb{R}^p$, $f : \mathscr{U} \to \mathbb{R}^n$. On dit que f **est de classe** $\mathscr{C}^k$ si toutes ses dérivées partielles d'ordre k existent et sont continues.

On dit que f **est de classe** $\mathscr{C}^\infty$ si elle est de classe $\mathscr{C}^k$ pour tout $k \in \mathbb{N}$.

Propriété 4 : Caractérisation par les applications coordonnées

On note $f = (f_1,\dots,f_n)$, $k \in \mathbb{N}\cup\{+\infty\}$. Alors f est de classe $\mathscr{C}^k$ si et seulement si chacune des f_i l'est.

Théorème 2 : de Schwarz

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^p$, $f : \mathscr{U} \to \mathbb{R}^p$ de classe $\mathscr{C}^2$. Alors, pour tout $i,j \in [\![1,p]\!]$ tel que $i \neq j$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Propriété 5 : Opérations sur les fonctions de classe $\mathscr{C}^k$

Soit $k \in \mathbb{N}\cup\{+\infty\}$.

(i) Toute combinaison linéaire, toute composée, tout produit, tout quotient de fonctions de classe $\mathscr{C}^k$ l'est encore.

(ii) Si $M : \mathbb{R}^{m_1}\times\cdots\times\mathbb{R}^{m_q} \to \mathbb{R}^n$ est q-linéaire, $f_1 : \mathscr{U} \to \mathbb{R}^{m_1},\dots,f_q : \mathscr{U} \to \mathbb{R}^{m_q}$ de classe $\mathscr{C}^k$ alors $M(f_1,\dots,f_q)$ est de classe $\mathscr{C}^k$.

(iii) Toute fonction polynomiale à n variables est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^n$.

(iv) Toute fonction rationnelle (quotient de fonctions polynomiales) est de classe $\mathscr{C}^\infty$ sur son domaine de définition.

5 Matrice hessienne et DL_2

Définition 9 : Matrice hessienne

Soit $f : \mathscr{U} \to \mathbb{R}$ où $\mathscr{U}$ est un ouvert de $\mathbb{R}^p$. On appelle, lorsqu'elle existe, **matrice hessienne** de f en $x \in \mathscr{U}$ la matrice

$$H_f(x) = \left(\partial_{i,j}\right)_{\leqslant i,j\leqslant p} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{1\leqslant i,j\leqslant p} \in \mathscr{M}_p(\mathbb{R}).$$

Théorème 3 : DL$_2$: formule de Taylor-Young à l'ordre 2

Soit $\mathcal{U}$ est un ouvert de $\mathbb{R}^p$, $f:\mathcal{U}\to\mathbb{R}$ de classe $\mathscr{C}^2$, $a\in\mathcal{U}$ et $h\in\mathbb{R}^p$ tel que $a+h$ reste dans $\mathcal{U}$ lorsque $h\to 0$.

En confondant $\mathbb{R}^p$ et $\mathcal{M}_{p,1}(\mathbb{R})$, et en utilisant le produit scalaire canonique sur $\mathbb{R}^p$, on a

$$f(a+h)=f(a)+\left(\nabla f(a)\middle|h\right)+\frac{1}{2}\left(H_f(a)h\middle|h\right)+\underset{h\to 0}{\mathrm{o}}\left(\|h\|^2\right)$$

$$f(a+h)=f(a)+\nabla f(a)^\mathsf{T}h+\frac{1}{2}h^\mathsf{T}H_f(a)h+\underset{h\to 0}{\mathrm{o}}\left(\|h\|^2\right)$$

II Applications

1 Équations aux dérivées partielles

Méthode 1

- Savoir résoudre les ÉDP fondamentales auxquelles on se ramène systématiquement :

$$\frac{\partial f}{\partial x}(x,y)=0 \qquad \frac{\partial f}{\partial x}(x,y)=g(x) \qquad \frac{\partial f}{\partial x}(x,y)=g(y)$$
$$\frac{\partial^2 f}{\partial x^2}(x,y)=0 \qquad \frac{\partial^2 f}{\partial x\partial y}(x,y)=0$$

- Dans la pratique, on s'y ramène via un changement de variables $(u,v)=\varphi(x,y)$, en écrivant $f(x,y)=g(u,v)$ et en remplaçant soit (x,y) en fonction de (u,v), soit (u,v) en fonction de (x,y).
- La changement de variable doit être bijectif, entre deux ouverts et suffisamment régulier (classe $\mathscr{C}^1$ ou $\mathscr{C}^2$ suivant l'ordre de l'équation.)
- S'il n'est pas donné, il doit être affine ou polaire.
- Appliquer la règle de la chaîne (ou utiliser des matrices jacobiennes) pour exprimer les dérivées de f en fonction de celle de g ou l'inverse, et simplifier l'ÉDP.

2 Optimisation : recherche d'extremums

Ici, toutes les fonctions sont à valeurs dans $\mathbb{R}$.

On notera plutôt n la dimension au départ : $\mathcal{U}$ désigne un ouvert de $\mathbb{R}^n$.

a Extremums libres

Définition 10 : Extremum

Soit A une partie de $\mathbb{R}^n$, $a\in A$, $f:A\to\mathbb{R}$.

(i) On dit que f présente en a un **maximum** (respectivement **minimum**) local s'il existe $\mathcal{V}$ voisinage de a dans A tel que pour tout $x\in\mathcal{V}$, $f(x)\leqslant f(a)$ (respectivement $f(x)\geqslant f(a)$).

Il est **strict** lorsque, $\forall x\in\mathcal{V}\setminus\{a\}$, , $f(x)<f(a)$ (respectivement $f(x)>f(a)$).

(ii) On dit que f présente un **maximum** (respectivement **minimum**) global si cette inégalité est en fait valable pour tout $x\in A$.

Définition 11 : Point critique

Soit $f:\mathcal{U}\to\mathbb{R}$ admettant des dérivées partielles en $a\in\mathcal{U}$.

Lorsque $\nabla f(a)=0$, c'est-à-dire $\forall i\in[\![1,n]\!]$, $\dfrac{\partial f}{\partial x_i}(a)=0$, on dit que a est un **point critique** de f.

Propriété 6 : Condition nécessaire d'extremum local à l'ordre 1

On suppose que

H1 *$\mathcal{U}$ ouvert (très important !) de $\mathbb{R}^n$.*

H2 *$a\in\mathcal{U}$ tel que $f:\mathcal{U}\to\mathbb{R}$ admettant des dérivées partielles en a.*

H3 *f présente un extremum local en a.*

Alors

C1 *a est un point critique de f.*

*La réciproque est fausse, et un contre-exemple est appelé **point selle** ou **point col**.*

Propriété 7 : Condition nécessaire de minimum local à l'ordre 2

On suppose que

H1 *$\mathcal{U}$ ouvert (très important !) de $\mathbb{R}^n$*

H2 *$a\in\mathcal{U}$ et $f\in\mathscr{C}^2(\mathcal{U},\mathbb{R})$.*

H3 *f présente un minimum local en a*

alors

C1 *a est un point critique de f.*

C2 *$H_f(a)\in\mathscr{S}_n^+(\mathbb{R})$.*

Si c'est un maximum local en a, alors $-H_f(a)\in\mathscr{S}_n^+(\mathbb{R})$ (ie $H_f(a)$ est symétrique « négative », ses valeurs propres sont toutes dans $\mathbb{R}^-$).

Propriété 8 : Condition suffisante d'extremum local à l'ordre 2

On suppose que

H1 *$\mathcal{U}$ ouvert (très important !) de $\mathbb{R}^n$.*

H2 *$a\in\mathcal{U}$ et $f\in\mathscr{C}^2(\mathcal{U},\mathbb{R})$.*

H3 *a est un point critique de f.*

H4 *$H_f(a)\in\mathscr{S}_n^{++}(\mathbb{R})$*

(respectivement $-H_f(a)\in\mathscr{S}_n^{++}(\mathbb{R})$)

alors

C1 *f atteint un minimum (respectivement maximum) local **strict** en a.*

CC BY NC SA

Corollaire 3 : Discussion sur le spectre

*Si $\mathscr{U}$ **ouvert**, $f \in \mathscr{C}^2(\mathscr{U}, \mathbb{R})$ et $a \in \mathscr{U}$ point critique de f.*

- *Si $\mathrm{Sp}\left(H_f(a)\right) \subset \mathbb{R}_+^*$, f atteint en a un minimum local.*
- *Si $\mathrm{Sp}\left(H_f(a)\right) \subset \mathbb{R}_-^*$, f atteint en a un maximum local.*
- *Si $H_f(a)$ possède des valeurs propres non nulles de signes opposés, a est un point selle.*

Si $n = 2$ (fonction de 2 variables), on note

$$H_f(a) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

où $r = \dfrac{\partial^2 f}{\partial x^2}(a)$, $s = \dfrac{\partial^2 f}{\partial x \partial y}(a) = \dfrac{\partial^2 f}{\partial y \partial x}(a)$ (théorème de Schwarz), $t = \dfrac{\partial^2 f}{\partial y^2}(a)$.

On suppose que a est un point critique. On a $\det\left(H_f(a)\right) = rt - s^2$ (égal produit de ses deux valeurs propres réelles) et $\mathrm{tr}\left(H_f(a)\right) = r + t$ (égal à leur somme).

1. Si $rt - s^2 > 0$, les deux valeurs propres de $H_f(a)$ ont même signe et sont non nulles : f **présente en** a **un extremum local strict.**
   (a) Si $r + t > 0$ alors $H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$ et f présente un **minimum local strict** en a.
   (b) Si $r + t < 0$ alors $-H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$ et f présente un **maximum local strict** en a.
2. Si $rt - s^2 < 0$, les deux valeurs propres de $H_f(a)$ ont des signes opposés et sont non nulles.
   Dans ce cas, f **présente en** a **un point col** : dans les directions propres, on a respectivement un minimum et un maximum local.
3. Si $rt - s^2 = 0$, on ne peut rien conclure en général.

Méthode 2 : Recherche d'extremum

(i) Soit $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles sur $\mathscr{U}$. Pour déterminer des extremums locaux de f on cherche ses points critiques.
Si f est de classe $\mathscr{C}^2$, on peut s'intéresser à la Hessienne aux points critiques qui permet de conclure lorsqu'elle est inversible (voir la remarque précédente pour $n = 2$.)
Sinon, pour un point critique a, on étudie $f(a+h) - f(a)$ pour h proche de 0. Comme a est un point critique, les termes obtenus dans la différence sont au moins d'ordre 2.

(ii) Si $\mathscr{U}$ n'est pas un ouvert, on le décompose en son intérieur (ouvert) et son bord. Sur le bord, on étudie « à la main », en général à l'aide d'un paramétrage du bord.

(iii) Si $f : K \to \mathbb{R}$ où K est un compact, on est assuré de l'existence d'un minimum et d'un maximum globaux. On les cherche comme dans la méthode précédente.

b Extremums liés

Théorème 4 : d'optimisation sous contrainte

Soient $f, g \in \mathscr{C}^1(\mathscr{U}, \mathbb{R})$ où $\mathscr{U}$ est un ouvert de $\mathbb{R}^n$, et $X = \{x \in \mathscr{U},\ g(x) = 0\}$.

Si $f_{|X}$ admet un extremum local en $a \in X$ et si $\nabla g(a) \neq 0_{\mathbb{R}^n}$, alors $\nabla f(a)$ est colinéaire à $\nabla g(a)$.

III Différentielle

1 Différentielle en un point

Définition 12 : Application différentiable en un point

Soit f définie sur un ouvert $\mathscr{U}$ de E, à valeurs dans F. Soit a un point de $\mathscr{U}$. On dit que f est **différentiable** en a lorsqu'il existe une application linéaire ℓ_a de E dans F telle que, au voisinage de 0_E,

$$f(a+h) = f(a) + \ell_a(h) + \underset{h \to 0_E}{o}(h)$$

ou encore, au voisinage de a,

$$f(x) = f(a) + \ell_a(x-a) + \underset{x \to a}{o}(x-a).$$

Lorsqu'elle existe, l'application ℓ_a est unique et appelée **différentielle** de f au point a ou encore **application linéaire tangente** à f en a, notée $\mathrm{d}f(a) \in \mathscr{L}(E, F)$. On a donc

$$f(a+h) = f(a) + \mathrm{d}f(a)(h) + \underset{h \to 0_E}{o}(h)$$

Propriété 9 : différentiable ⇒ continue

Si f est différentiable en a, alors f est continue en a.

2 Cas particuliers

Propriété 10 : Cas d'une fonction d'une variable réelle

Dans le cas d'une fonction $f : I \to F$, f est dérivable en a si et seulement si elle est différentiable en a.

Dans ce cas, $\mathrm{d}f(a) : h \mapsto hf'(a)$ et en particulier $f'(a) = \mathrm{d}f(a)(1)$.

Propriété 11 : Cas d'une fonction constante

Si $f : \mathscr{U} \subset E \to F$ est constante, elle est différentiable en tout point de $\mathscr{U}$ et pour tout $a \in \mathscr{U}$, $\mathrm{d}f(a) = 0_{\mathscr{L}(E,F)}$.

Propriété 12 : Cas d'une fonction linéaire

Si $f : \mathcal{U} \subset E \to F$ est la restriction à $\mathcal{U}$ d'une application linéaire $\varphi \in \mathcal{L}(E,F)$, elle est différentiable en tout point de $\mathcal{U}$ et pour tout $a \in \mathcal{U}$, $\mathrm{d}f(a) = \varphi$ (et donc $\mathrm{d}f$ est constante).

3 Dérivée selon un vecteur, dérivées partielles

Définition 13 : Dérivée selon un vecteur

On dit que $f : \mathcal{U} \subset E \to F$ est **dérivable selon le vecteur** $v \in E$ **au point** $a \in \mathcal{U}$, lorsque $\phi : t \mapsto f(a+tv)$ est dérivable en 0.

On note alors $\mathrm{D}_v f(a) = \phi'(0) = \lim_{t\to 0} \dfrac{f(a+tv)-f(a)}{t} \in F$.

Définition 14 : Dérivées partielles

Soient $\mathcal{B} = (e_1,\dots,e_p)$ une base de E, $f : \mathcal{U} \subset E \to F$, $a \in \mathcal{U}$, $j \in [\![1,p]\!]$.

On appelle j^e **dérivée partielle de** f **en** a, lorsqu'elle existe, la dérivée de f selon le vecteur e_j de base en a :

$$\partial_j f(a) = \frac{\partial f}{\partial x_j}(a) = \mathrm{D}_{e_j} f(a) \in F.$$

4 Lien entre différentielle et dérivées partielles

Propriété 13 : Lien entre différentielle et dérivées partielles

Soit $f : \mathcal{U} \to F$, $a \in \mathcal{U}$. Si f est différentiable en a, alors f est dérivable en a selon tout vecteur $v \in E$ et $D_v f(a) = \mathrm{d}f(a)(v)$.

Cas particulier 2 : Dérivée selon un vecteur de base

Soit $f : \mathcal{U} \to F$, $a \in \mathcal{U}$, $\mathcal{B} = (e_1,\dots,e_p)$ une base de E. On note $\dfrac{\partial f}{\partial x_j}$ les dérivées partielles de f dans la base $\mathcal{B}$. Si f est différentiable en a, alors pour tout $j \in [\![1,p]\!]$,

$$\mathrm{d}f(a)(e_j) = D_{e_j} f(a) = \frac{\partial f}{\partial x_j}(a).$$

Propriété 14 : Expression de la différentielle avec les dérivées partielles

Soit $\mathcal{B} = (e_1,\dots,e_p)$ une base de E, $f : \mathcal{U} \to F$, $a \in \mathcal{U}$ tel que f est différentiable en a.

Alors pour tout vecteur $h = \sum_{j=1}^{p} h_j e_j \in E$,

$$\mathrm{d}f(a)(h) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a)$$

On note $\mathrm{d}x_j : \left| \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ h & \longmapsto & h_j \end{array} \right.$ la forme linéaire j^e coordonnée dans $\mathcal{B}$. Alors on a

$$\mathrm{d}f(a) = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a)\,\mathrm{d}x_j.$$

5 Matrice jacobienne

Définition 15 : Matrice jacobienne

Soit $p = \dim E$, $n = \dim F$, $\mathcal{U}$ ouvert de E et $f : E \to F$ différentiable, $\mathcal{B} = (e_1,\dots,e_p)$ une base de E et $\mathcal{C} = (\varepsilon_1,\dots,\varepsilon_n)$ une base de F. On note $f = \sum_{i=1}^{n} f_i \varepsilon_i$.

On appelle **matrice jacobienne** de f en a dans les bases $\mathcal{B}$ et $\mathcal{C}$ la matrice $J_f(a) = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$.

Propriété 15 : Matrice jacobienne et différentielle

$$J_f(a) = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}\left(\mathrm{d}f(a)\right)$$

6 Gradient

Définition 16 : Gradient

Soit E un espace **euclidien**, $\mathcal{U}$ un ouvert de E, $f : \mathcal{U} \to \mathbb{R}$ une fonction différentiable. Alors pour tout $a \in \mathcal{U}$, il existe un unique vecteur noté $\nabla f(a)$ ou $\overrightarrow{\mathrm{grad}}\, f(a)$ et appelé **gradient de** f **en** a tel que pour tout $h \in E$,

$$\mathrm{d}f(a)(h) = (\nabla f(a)|h).$$

Propriété 16 : Coordonnées du gradient

*Soit E un espace **euclidien**, $\mathcal{U}$ un ouvert de E, $f : \mathcal{U} \to \mathbb{R}$ une fonction différentiable. Si l'on fixe une base **orthonormée** de E, le coordonnées de $\nabla f(a)$ dans cette base sont*

$$\left(\frac{\partial f}{\partial x_1}(a),\dots,\frac{\partial f}{\partial x_p}(a)\right).$$

CC BY NC SA

IV Opérations sur les différentielles

1 Combinaisons linéaires

Propriété 17 : Linéarité

Soit $\mathscr{U}$ ouvert de E, $f, g : \mathscr{U} \to F$ différentiables en $a \in \mathscr{U}$, $\lambda, \mu \in \mathbb{R}$.

Alors $\lambda f + \mu g$ est différentiable et

$$\mathrm{d}(\lambda f + \mu g)(a) = \lambda \,\mathrm{d} f(a) + \mu \,\mathrm{d} g(a).$$

2 Image par une application multilinéaire

Propriété 18 : Image par une application multilinéaire

Soit $\mathscr{U}$ ouvert d'un $\mathbb{R}$-espace vectoriel normé de dimension finie, $E_1, \dots, E_q, F$ des $\mathbb{R}$-espaces vectoriels normés de dimension finie, $f_1 : \mathscr{U} \to E_1, \dots, f_q : \mathscr{U} \to E_q$ des applications différentiables en $a \in \mathscr{U}$ et $M : E_1 \times \cdots \times E_q \to F$ une fonction q-linéaire.

Alors $\phi : \mathscr{U} \to F$ définie par $\phi : x \mapsto M(f_1(x), \dots, f_q(x))$ est différentiable en a et

$$\mathrm{d}\phi(a) : h \mapsto \sum_{k=1}^{q} M(f_1(a), \dots, \mathrm{d} f_k(a)(h), \dots, f_q(a)).$$

3 Composition

Propriété 19 : Différentielle d'une composée

Soit $\mathscr{U}$ ouvert de E, $\mathscr{V}$ ouvert de F, $f : \mathscr{U} \to \mathscr{V}$ différentiable en $a \in \mathscr{U}$ et $g : \mathscr{V} \to G$ différentiable en $b = f(a)$.

Alors $g \circ f$ est différentiable en a et

$$\mathrm{d}(g \circ f)(a) = \mathrm{d} g(f(a)) \circ \mathrm{d} f(a).$$

Propriété 20 : Matrice jacobienne d'une composée

En munissant E, F et G de bases, on obtient

$$J_{g \circ f}(a) = J_g(f(a)) \times J_f(a).$$

Corollaire 4 : Dérivée le long d'un arc

Soit $\mathscr{U}$ ouvert de E, $f : \mathscr{U} \to F$, I intervalle de $\mathbb{R}$ et $\gamma : I \to \mathscr{U}$ une application dérivable.

On suppose que f est différentiable en $\gamma(t)$ où $t \in I$. Alors $f \circ \gamma$ est dérivable en t et

$$(f \circ \gamma)'(t) = \mathrm{d} f(\gamma(t)) \left(\gamma'(t)\right).$$

En munissant E d'une base, on retrouve l'expression vue avec la règle de la chaîne, si les coordonnées de $\gamma(t)$ sont $(\gamma_1(t), \dots, \gamma_p(t))$,

$$(f \circ \gamma)'(t) = \sum_{k=1}^{p} \gamma_k'(t) \frac{\partial f}{\partial x_k}(\gamma(t)).$$

En particulier, si $\gamma : t \mapsto x + tv$ où $x, v \in \mathscr{U}$ sont fixés,

$$(f \circ \gamma)'(t) = \mathrm{d} f(\gamma(t))(v) = D_v f(\gamma(t)).$$

V Classe $\mathscr{C}^1$

Définition 17 : Fonctions de classe $\mathscr{C}^1$

$f : \mathscr{U} \to F$ est dite **de classe** $\mathscr{C}^1$ lorsque f est différentiable sur $\mathscr{U}$ et $\mathrm{d} f : \mathscr{U} \to \mathscr{L}(E, F)$ est continue.

Propriété 21 : Opérations

Toute combinaison linéaire, toute composée d'applications de classe $\mathscr{C}^1$ l'est encore.

Si M est q-linéaire et $f_1, \dots, f_q$ sont $\mathscr{C}^1$, $M(f_1, \dots, f_q)$ l'est.

Propriété 22 : Caractérisation par les applications coordonnées

Soit $f : \mathscr{U} \to F$, d'applications coordonnées $(f_1, \dots, f_n)$ dans une base de F. f est de classe $\mathscr{C}^1$ si et seulement si toutes les f_i le sont.

Théorème 5 : IMPORTANT - Caractérisation avec les dérivées partielles

f est de classe $\mathscr{C}^1$ si et seulement si, dans une base quelconque de E, toutes les dérivées partielles de f existent et sont continues.

Propriété 23 : Expression intégrale le long d'un arc

Soit $f \in \mathscr{C}^1(\mathscr{U}, F)$, $\gamma \in \mathscr{C}^1([0,1], \mathscr{U})$, $a = \gamma(0)$ et $b = \gamma(1)$. Alors

$$f(b) - f(a) = \int_0^1 \mathrm{d} f(\gamma(t)) \left(\gamma'(t)\right) \mathrm{d}t.$$

Corollaire 5 : Cas particulier $\gamma(t) = a + tv$

En particulier, avec $\gamma(t) = a + tv$ sur $[0,1]$,

$$f(a+v) - f(a) = \int_0^1 \mathrm{d}f(a+tv)(v)\,\mathrm{d}t = \int_0^1 D_v f(a+tv)\,\mathrm{d}t.$$

Corollaire 6 : Caractérisation des fonctions constantes sur un ouvert connexe par arcs

*Si $f \in \mathscr{C}^1(\mathscr{U}, F)$ et $\mathscr{U}$ est un ouvert **connexe par arcs**, alors*

f est constante si et seulement si $\mathrm{d}f = 0$.

VI Vecteurs tangents

Définition 18 : Vecteur tangent

Si X est une partie de E, $a \in X$. Un vecteur $v \in E$ est dit **tangent** à X en a lorsqu'il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to X$, dérivable en 0 tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

On note $T_a X$ l'ensemble des vecteurs tangents à X en a.

Propriété 24 : Plan tangent pour une surface explicite

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^2$, $f : \mathscr{U} \to \mathbb{R}$ différentiable.

Soit $S \subset \mathbb{R}^3$ la surface représentative de f, c'est-à-dire

$$S = \left\{(x,y,z) \in \mathbb{R}^3,\ (x,y) \in \mathscr{U} \text{ et } z = f(x,y)\right\}.$$

Soit $g : \left| \begin{array}{ccc} \mathscr{U} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x,y,z) & \longmapsto & f(x,y) - z \end{array} \right.$

L'ensemble $T_a S$ des vecteurs tangents à S en $a = (x_0, y_0, z_0) \in S$ est un plan vectoriel $P = \nabla g(a)^{\perp}$ de vecteur normal le vecteur (non nul) $\nabla g(a)$, donc d'équation

$$\left(\nabla g(a) \middle| (x,y,z)\right) = 0.$$

*On appelle **plan tangent à** S **en** a le plan affine $a + P$ passant par a et de direction l'ensemble P des vecteurs tangents à X en a, d'équation*

$$\left(\nabla g(a) \middle| (x - x_0, y - y_0, z - z_0)\right) = 0$$

Propriété 25 : Hyperplan tangent à une surface implicite

Soit g une fonction numérique de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$, $X = \{x \in \mathscr{U},\ g(x) = 0_{\mathbb{R}}\}$ et $a \in X$.

Si $\mathrm{d}g(a) \neq 0_{\mathscr{L}(E,\mathbb{R})}$, alors

$$T_a X = \operatorname{Ker}\left(\mathrm{d}g(a)\right) = \nabla g(a)^{\perp}$$

(L'espace étant supposé euclidien pour utiliser le gradient.)

Autrement dit, X admet un hyperplan affine tangent en a qui est, $a + T_a X = a + \nabla g(a)^{\perp}$ d'équation

$$\left(\nabla g(a) \middle| x - a\right) = 0.$$

Méthode 3 : Trouver une équation d'hyperplan tangent

On retiendra que dans tous les cas, mieux vaut repasser par une équation **implicite** $g(x) = 0$ pour trouver une équation d'hyperplan tangent.

Le théorème précédent donne alors une équation dudit hyperplan tangent en un point a n'annulant pas la $\mathrm{d}g$.

Par exemple, le plan tangent à la surface S d'équation $g(x,y,z) = 0$ en $a = (x_0, y_0, z_0)$ est le plan affine $a + \nabla g(a)^{\perp}$ passant par a et de direction le plan vectoriel $T_a S = \nabla g(a)^{\perp}$ de vecteur normal $\nabla g(a) \neq (0,0,0)$, d'équation

$$\frac{\partial g}{\partial x}(x_0,y_0,z_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0,y_0,z_0)(y - y_0) + \frac{\partial g}{\partial z}(x_0,y_0,z_0)(z - z_0) = 0$$

où $g(a) = g(x_0, y_0, z_0) = 0$.

VII Optimisation (recherche d'extremums)

Définition 19 : Point critique

Soit $f : \mathscr{U} \to \mathbb{R}$ différentiable. On dit que a est un **point critique** de f lorsque $\mathrm{d}f(a) = 0$.

Si E est euclidien, cela équivaut à $\nabla f(a) = 0$.

Propriété 26 : Condition nécessaire d'extremum local

*Soit $\mathscr{U}$ **ouvert** (très important) et $f : \mathscr{U} \to \mathbb{R}$ différentiable en $a \in \mathscr{U}$.*

Si f présente un extremum local en a, alors a est un point critique de f, c'est-à-dire $\mathrm{d}f(a) = 0$.

CC BY NC SA

Propriété 27 : Vecteurs tangents en un extremum local

Si f est une fonction numérique définie sur l'ouvert $\mathscr{U}$, X une partie de $\mathscr{U}$, $f_{|X}$ admet un extremum local en $a \in X$ et f est différentiable en a, alors

$$\forall\, v \in T_a X,\quad \mathrm{d}\, f(a)(v) = 0.$$

Autrement dit, dans le cas où E est euclidien,

$$T_a X \subset \mathrm{Ker}(\mathrm{d}\, f(a)) = \left(\nabla f(a)\right)^{\perp}$$

Théorème 6 : d'optimisation sous contrainte

Soient $f, g \in \mathscr{C}^1(\mathscr{U}, \mathbb{R})$ où $\mathscr{U}$ est un ouvert de E, et $X = \left\{x \in \mathscr{U},\ g(x) = 0\right\}$.

Si $f_{|X}$ admet un extremum local en $a \in X$ et si $\mathrm{d}\, g(a) \neq 0_E$, alors $\mathrm{d}\, f(a)$ est colinéaire à $\mathrm{d}\, g(a)$.

Théorème 7 : HP : multiplicateurs de Lagrange

*Soit $f, g_1, \ldots, g_p$ des fonctions numériques de classe $\mathscr{C}^1$ sur l'**ouvert** $\mathscr{U}$ de E et*

$$X = \left\{x \in \mathscr{U},\ g_1(x) = \cdots = g_p(x) = 0\right\}.$$

Si $f_{|X}$ admet un extremum local en $a \in X$ et si les formes linéaires $\mathrm{d}\, g_1(a), \ldots, \mathrm{d}\, g_p(a)$ sont linéaires indépendantes, alors il existe $\lambda_1, \ldots, \lambda_p \in \mathbb{R}$ tels que

$$\mathrm{d}\, f(a) = \lambda_1\, \mathrm{d}\, g_1(a) + \cdots + \lambda_p\, \mathrm{d}\, g_p(a).$$

*Les réels $\lambda_1, \ldots, \lambda_p$ sont appelés **multiplicateurs de Lagrange**.*