

Calcul différentiel et optimisation

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

DÉRIVÉES PARTIELLES

Dans cette partie, \mathcal{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 $f : \mathcal{U} \rightarrow \mathbb{R}$ une fonction, $a = (a_1, \dots, a_p) \in \mathcal{U}$, $j \in \llbracket 1, p \rrbracket$. On appelle j^{e} **application partielle de f en a** l'application $t \mapsto f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_p)$.

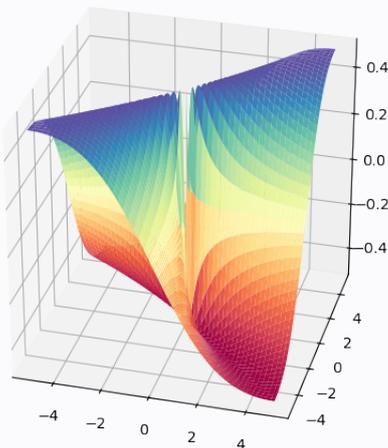
Remarque

R1 – Si $f : (x_1, \dots, x_p) \in \mathbb{R}^p \mapsto f(x_1, \dots, x_p) \in F$ est une fonction de n variables, la continuité des applications partielles $x_i \mapsto f(x_1, \dots, x_p)$ (les x_j pour $j \neq i$ étant fixés) ne garantit pas celle de f .

Exemple

E1 –

$$f : (x, y) \mapsto \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$



Lemme 1 : « de partition »

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

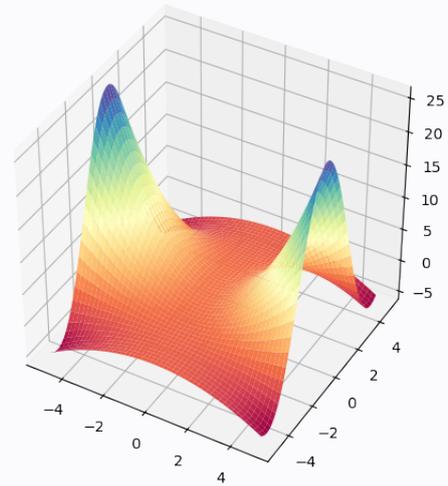
Si $f|_{B_1}(x) \xrightarrow{x \rightarrow a} \ell$ et $f|_{B_2}(x) \xrightarrow{x \rightarrow a} \ell$, alors $f(x) \xrightarrow{x \rightarrow a} \ell$.

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

Exemple

E2 –

$$f : (x, y) \mapsto \begin{cases} \frac{(1+x^2)\sin y}{1+x^2} & \text{si } y \neq 0, \\ y & \text{sinon.} \end{cases}$$



2 Dérivées partielles

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

Remarque

R2 – Le domaine de définition d'une application partielle peut être compliqué, mais c'est toujours un ouvert de \mathbb{R} .

Définition 2 : Dérivées partielles

Soit $f : \mathcal{U} \rightarrow \mathbb{R}$, $a = (a_1, \dots, a_p) \in \mathcal{U}$, $j \in \llbracket 1, p \rrbracket$. 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 $\partial_j f(a) = \frac{\partial f}{\partial x_j}(a)$ le nombre dérivé en ce point.

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

Remarque

R3 – Comme on en a déjà l'habitude, calculer la j^{e} dérivée partielle revient à fixer les coordonnées selon tous les autres vecteurs de bases et dériver par rapport à la seule variable x_j .

$$\frac{f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) - f(a_1, \dots, a_n)}{t - a_j} \xrightarrow{t \rightarrow a_j} \frac{\partial f}{\partial x_j}(a).$$

R4 – En général, $p = 2$ ou 3 . Si $p = 2$, par exemple, on note plutôt $f(x, y)$ que $f(x_1, x_2)$. On notera alors volontiers $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ les deux dérivées partielles selon la PREMIÈRE ou la DEUXIÈME variable. En mathématiques, on dérive selon une position et non « par rapport à une va-



riable », dont le nom n'importe pas.

Exemple

$$\text{E3} - f : (y, x) \mapsto xy^2$$

Que désigne dans ce cas $\frac{\partial f}{\partial x}$??

Définition 3 : Vecteur gradient

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

$$\nabla f(a) = \overrightarrow{\text{grad}} f(a) = \quad .$$

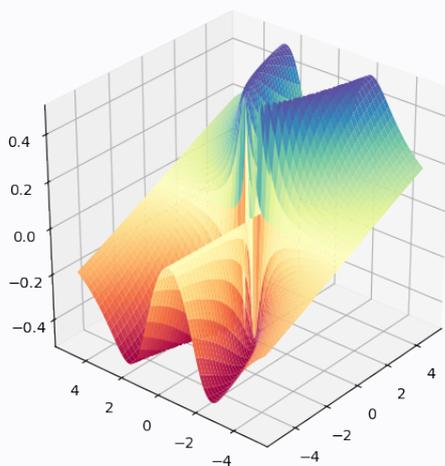
Exemple

E4 – Dérivées partielles en tout point de

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Les application partielles $x \mapsto f(x, 0)$ et $y \mapsto f(0, y)$ sont-elles continues en 0 ?

f est-elle continue en $(0, 0)$?

**Remarque**

R5 – Et donc \triangleleft l'existence de dérivées partielles n'implique pas la continuité !

Définition 4 : Extension aux fonctions à valeurs vectorielles

Soit \mathcal{U} ouvert de \mathbb{R}^p et $f : \mathcal{U} \rightarrow \mathbb{R}^n$. 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 $a \in \mathcal{U}$ si chacune des f_i admet une dérivée partielle en a .

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

Exemple

E5 – Dérivées partielles en tout point de

$$f : (r, \theta, \phi) \mapsto (r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta).$$

Définition 5 : Matrice jacobienne

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

Remarque

R6 – Coefficient (i, j) : dérivée de f_i par rapport à x_j , dans l'ordre.

R7 – $J_f(a) \in \mathcal{M}_{n,p}(\mathbb{R})$

Donc si $f : \mathbb{R}^p \rightarrow \mathbb{R}$, c'est une matrice ligne : la transposée du gradient $(\nabla f(a))^T$ en confondant n -uplet et vecteur colonne.

Et si $f : \mathbb{R} \rightarrow \mathbb{R}^n$, c'est une matrice colonne. En fait, $J_f(a) = f'(a)$ en confondant toujours n -uplet et vecteur colonne.

Exemple : Matrice jacobienne de changement de variables

E6 – Matrice jacobienne du changement de variables en coordonnées polaires :

$$f : (r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

E7 – Matrice jacobienne du changement de variables en coordonnées cylindriques :

$$f : (r, \theta, z) \mapsto (r \cos \theta, r \sin \theta, z).$$

E8 – Matrice jacobienne du changement de variables en coordonnées sphériques :

$$f : (r, \theta, \phi) \mapsto (r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta).$$

3 Fonctions de classe \mathcal{C}^1 **Définition 6 : Classe \mathcal{C}^1**

Soit $f : \mathcal{U} \rightarrow \mathbb{R}^n$. On dit que f est de classe \mathcal{C}^1 sur \mathcal{U} lorsqu'en tout point de \mathcal{U} , les dérivées partielles de f existent, et que ces dérivées partielles sont continues. On note $\mathcal{C}^1(\mathcal{U}, \mathbb{R}^n)$ 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 \mathcal{C}^1 si et seulement si chacune des f_i l'est.

Propriété 2 : Structure d'algèbre

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

Remarque

R8 – Il suffit donc de travailler sur les fonctions $f : \mathcal{U} \rightarrow \mathbb{R}$.

R9 – On rappelle que l'on note, pour $f : E \rightarrow F$, $f(h) = \underset{h \rightarrow 0_E}{o} (h)$ ou $f(h) = \underset{h \rightarrow 0_E}{o} (\|h\|_E)$ lorsque

$$\|f(h)\|_F = \underset{h \rightarrow 0_E}{o} (\|h\|_E),$$

autrement dit lorsque

$$\frac{\|f(h)\|_F}{\|h\|_E} \xrightarrow{h \rightarrow 0_E} 0_{\mathbb{R}}.$$

Cela revient aussi à écrire que

$$f(h) = \|h\|_E \varepsilon(h)$$

où

$$\varepsilon(h) \xrightarrow{h \rightarrow 0_E} 0_F.$$

Comme on travaille en dimension finie, n'importe quel quelles normes conviennent.

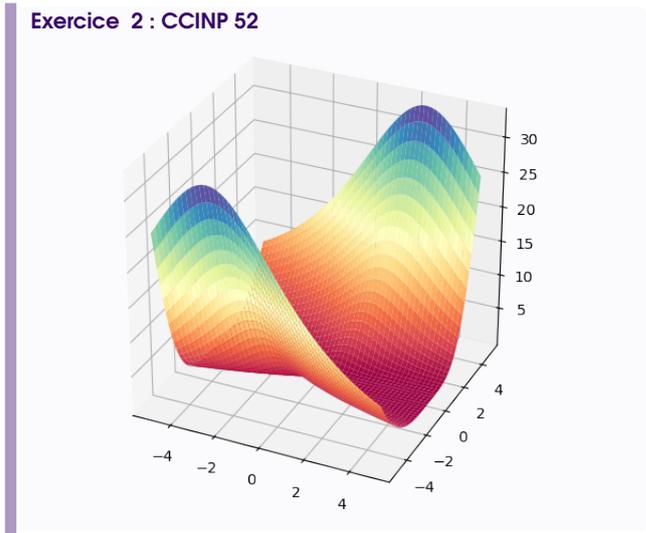
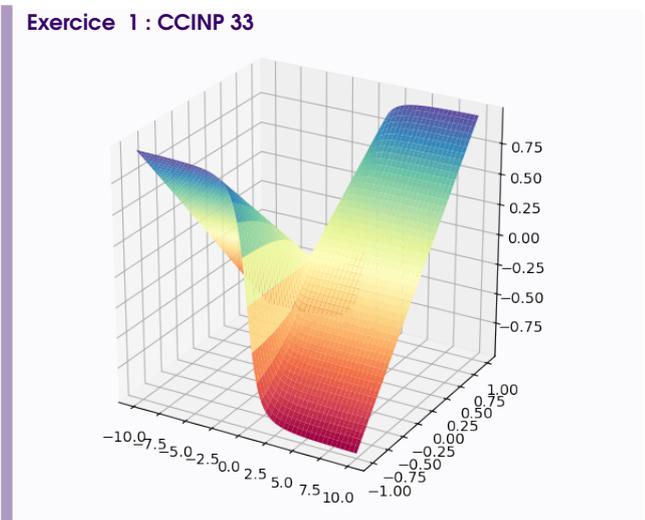
Théorème 1 : DL₁

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

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

Corollaire 1 : $\mathcal{C}^1 \implies$ continue

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



Propriété 3 : Règle de la chaîne

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

$$f : \begin{cases} \mathcal{U} & \longrightarrow \mathcal{V} \\ (x_1, \dots, x_p) & \longmapsto (f_1(x_1, \dots, x_p), \dots, f_n(x_1, \dots, x_p)) \end{cases}$$

et

$$g : \begin{cases} \mathcal{V} & \longrightarrow \mathbb{R} \\ (y_1, \dots, y_n) & \longmapsto g(y_1, \dots, y_n) \end{cases}$$

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

$$\frac{\partial (g \circ f)}{\partial x_j}(a) =$$

Exercice 3 : Changement de variable et gradient en polaire

Soit $\mathcal{U} = \mathbb{R}^2 \setminus \{(0,0)\}$, $V =]0, +\infty[\times \mathbb{R}$, $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R})$,

$$g : \begin{cases} V & \longrightarrow \mathbb{R} \\ (r, \theta) & \longmapsto f(r \cos \theta, r \sin \theta) \end{cases}$$

$\vec{u}(\theta) = (\cos \theta, \sin \theta)$ et $\vec{v}(\theta) = \vec{u}'(\theta)$.

1. Montrer que g est de classe \mathcal{C}^1 sur V .
2. Si $(r, \theta) \in V$, on note $x = r \cos \theta$ et $y = r \sin \theta$. Exprimer les dérivées partielles de g en (r, θ) en fonction des dérivées partielles de f en (x, y) .
3. Exprimer le gradient $\left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right)$ dans la base $(\vec{u}(\theta), \vec{v}(\theta))$ en fonction des dérivées partielles de g en (r, θ) .

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

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

$$(f \circ \gamma)'(t) =$$

**Corollaire 2 : Règle de la chaîne vectorielle**

Soient \mathcal{U}, \mathcal{V} deux ouverts respectifs de \mathbb{R}^p et \mathbb{R}^n et $a \in \mathcal{U}$. Soient $f : \mathcal{U} \rightarrow \mathcal{V}$ et $g : \mathcal{V} \rightarrow \mathbb{R}^q$ deux fonctions de classe \mathcal{C}^1 . Alors $g \circ f$ est de classe \mathcal{C}^1 et pour $j \in \llbracket 1, p \rrbracket$ et $\ell \in \llbracket 1, q \rrbracket$,

$$\frac{\partial (g \circ f)}{\partial x_j}(a) = \frac{\partial (g \circ f)_\ell}{\partial x_j}(a) =$$

En particulier,

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

Exercice 4 : Calculer les dérivées partielles de $h : (x, y) \mapsto g(x + y, xy)$ par la règle de la chaîne puis par les matrices jacobiennes.

Exercice 5 : Calculer la dérivée de $g : t \mapsto f(ta_1, \dots, ta_n)$ où $f \in \mathcal{C}^1(\mathbb{R}^p, \mathbb{R})$ et $a = (a_1, \dots, a_n) \in \mathbb{R}^p$.

4 Fonctions de classe \mathcal{C}^k **Définition 7 : Dérivées partielles d'ordre supérieur**

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

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

notées $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} = \partial_{i_k, \dots, i_1} f$ où $i_1, \dots, i_k \in \llbracket 1, p \rrbracket$.

Définition 8 : Classe \mathcal{C}^k

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

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

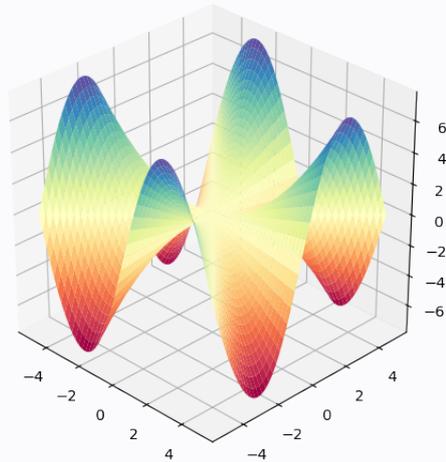
Propriété 4 : Caractérisation par les applications coordonnées

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

Théorème 2 : de Schwarz

Soit \mathcal{U} ouvert de \mathbb{R}^p , $f : \mathcal{U} \rightarrow \mathbb{R}^p$ de classe \mathcal{C}^2 . Alors, pour tout $i, j \in \llbracket 1, p \rrbracket$ tel que $i \neq j$,

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

Exercice 6 : CCINP 57

Exercice 7 : Contre-exemple au théorème de Schwarz : avec la fonction de l'exercice précédent (due à Péano) : calculer $\frac{\partial^2 f}{\partial x \partial y}(0, 0)$

et $\frac{\partial^2 f}{\partial y \partial x}(0, 0)$. Qu'en conclut-on ?

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

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

- (i) Toute combinaison linéaire, toute composée, tout produit, tout quotient de fonctions de classe \mathcal{C}^k l'est encore.
- (ii) Si $M : \mathbb{R}^{m_1} \times \dots \times \mathbb{R}^{m_q} \rightarrow \mathbb{R}^n$ est q -linéaire, $f_1 : \mathcal{U} \rightarrow \mathbb{R}^{m_1}, \dots, f_q : \mathcal{U} \rightarrow \mathbb{R}^{m_q}$ de classe \mathcal{C}^k alors $M(f_1, \dots, f_q)$ est de classe \mathcal{C}^k .
- (iii) Toute fonction polynomiale à n variables est de classe \mathcal{C}^∞ sur \mathbb{R}^n .
- (iv) Toute fonction rationnelle (quotient de fonctions polynomiales) est de classe \mathcal{C}^∞ sur son domaine de définition.

5 Matrice hessienne et DL_2 **Définition 9 : Matrice hessienne**

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

Remarque

R 10 – Le théorème de Schwarz assure que si f est de classe \mathcal{C}^2 sur \mathcal{U} , pour tout $x \in \mathcal{U}$,

Théorème 3 : DL₂ : formule de Taylor-Young à l'ordre 2

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

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

Remarque

R 11 – Se réécrit, en posant $h = (h_1, \dots, h_p)$,

$$f(a+h) = f(a) + \sum_{i=1}^p h_i \partial_i f(a) + \frac{1}{2} \sum_{1 \leq i, j \leq p} h_i h_j \partial_{i,j}^2 f(a) + o_{h \rightarrow 0}(\|h\|^2).$$

Exercice 8 : Résoudre

$$2 \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial y}(x, y) = 0$$

sur $\mathcal{C}^1(\mathbb{R}^2)$ à l'aide du changement de variable $(u, v) = \varphi(x, y) = (x + y, x + 2y)$, en vérifiant que φ est une bijection de \mathbb{R}^2 sur \mathbb{R}^2 .

Exercice 9 : Soient $a, b \in \mathbb{R}$. Résoudre

$$a \frac{\partial f}{\partial x}(x, y) + b \frac{\partial f}{\partial y}(x, y) = 0$$

sur $\mathcal{C}^1(\mathbb{R}^2)$ à l'aide d'un changement de variable affine.

Exercice 10 : Résoudre l'équation des cordes vibrantes

$$\frac{\partial^2 f}{\partial x^2}(x, t) - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2}(x, t) = 0$$

sur $\mathcal{C}^2(\mathbb{R}^2)$.

Exercice 11 : Coordonnées polaires

Soit $\mathcal{V} =]0, +\infty[\times]-\frac{\pi}{2}, \frac{\pi}{2}[$ et $\varphi : \mathcal{V} \rightarrow \mathbb{R}^2$ défini par $\varphi(r, \theta) = (r \cos \theta, r \sin \theta)$.

Déterminer un ouvert \mathcal{U} de \mathbb{R}^2 tel que φ soit une bijection de \mathcal{V} sur \mathcal{U} .

Résoudre

$$y \frac{\partial f}{\partial x}(x, y) - x \frac{\partial f}{\partial y}(x, y) = 0$$

sur $\mathcal{C}^1(\mathcal{U}, \mathbb{R})$.

Exercice 12 : À l'aide du changement de variable

$(u, v) = (x, \frac{y}{x})$, résoudre sur $\mathcal{C}^2(\mathbb{R}_*^+ \times \mathbb{R})$,

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

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 , $a \in A$, $f : A \rightarrow \mathbb{R}$.

(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 $x \in \mathcal{V}$, $f(x) \leq f(a)$ (respectivement $f(x) \geq f(a)$).

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 $x \in A$.

II APPLICATIONS

1 Équations aux dérivées partielles

Exemple : Quelques exemples fondamentaux

E 9 – Résoudre $\frac{\partial f}{\partial x}(x, y) = 0$ dans $\mathcal{C}^1(\mathcal{U}, \mathbb{R})$ où $\mathcal{U} = \mathbb{R}^2$ puis $\mathbb{R}^2 \setminus \{(0, 0)\}$ puis $\mathbb{R}^2 \setminus B'(0, 1)$.

E 10 – Résoudre $\frac{\partial f}{\partial x}(x, y) = g(x)$ dans $\mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ où $g \in \mathcal{C}(\mathbb{R})$.

E 11 – Résoudre $\frac{\partial f}{\partial x}(x, y) = g(y)$ dans $\mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ où $g \in \mathcal{C}(\mathbb{R})$.

E 12 – Résoudre $\frac{\partial^2 f}{\partial x^2}(x, y) = 0$ dans $\mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$.

E 13 – Résoudre $\frac{\partial^2 f}{\partial x \partial y}(x, y) = 0$ dans $\mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$.



Méthode 1

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

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) = 0 & \quad \frac{\partial f}{\partial x}(x, y) = g(x) & \quad \frac{\partial f}{\partial x}(x, y) = g(y) \\ \frac{\partial^2 f}{\partial x^2}(x, y) = 0 & \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 0 \end{aligned}$$

- 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 \mathcal{C}^1 ou \mathcal{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 jacobiniennes) pour exprimer les dérivées de f en fonction de celle de g ou l'inverse, et simplifier l'ÉDP.

**Définition 11 : Point critique**

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

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

Comme en dimension 1, une condition nécessaire d'extremum local est liée à l'annulation du terme d'ordre 1 dans le développement limité, donc du gradient.

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

On suppose que

H1

H2

H3

Alors

C1 a est un point critique de f .

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

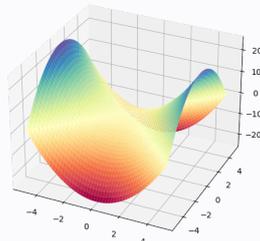
Exemple

E 14 –

$$f : (x, y) \mapsto x^2 - y^2$$

présente un point selle en $(0, 0)$.

La surface porte le doux nom de **paraboloïde hyperbolique**.



Le développement limité à l'ordre 2 permet de conditionner le fait d'avoir un minimum ou un maximum local (donc un signe constant pour $f(a+h) - f(a)$) au signe du terme d'ordre 2.

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

On suppose que

H1

H2

H3

alors

C1 a est un point critique de f .

C2

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

Si jamais, de plus, les termes d'ordre 2 du développement limité ne sont pas nuls, on obtient un équivalent pour $f(a+h) - f(a)$ ayant le même signe, ce qui permet de formuler une condition suffisante.

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

On suppose que

H1

H2

H3

H4

alors

C1

Corollaire 3 : Discussion sur le spectre

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

- Si $\text{Sp}(H_f(a)) \subset \mathbb{R}_+^*$, f atteint en a un minimum local.
- Si $\text{Sp}(H_f(a)) \subset \mathbb{R}_-^*$, 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 $r = \frac{\partial^2 f}{\partial x^2}(a)$, $s = \frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a)$ (théorème de Schwarz), $t = \frac{\partial^2 f}{\partial y^2}(a)$.

On suppose que a est un point critique. On a $\det(H_f(a)) = rt - s^2$ (égal produit de ses deux valeurs propres réelles) et $\text{tr}(H_f(a)) = r + t$ (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 $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$ et f présente un **minimum local strict** en a .
 - (b) Si $r + t < 0$ alors $-H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$ 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.

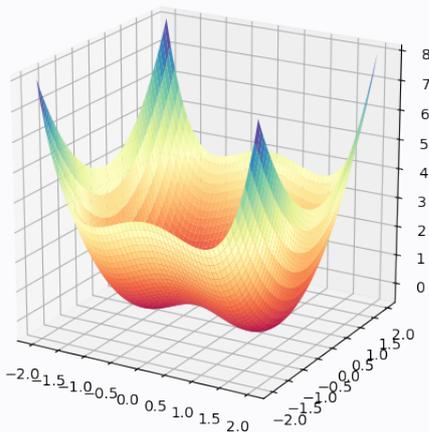


Méthode 2 : Recherche d'extremum

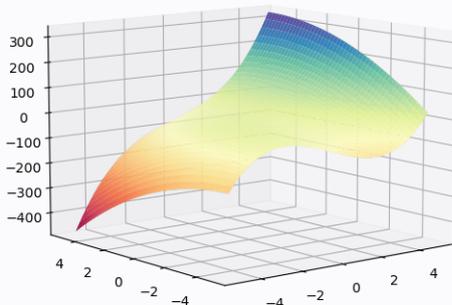
- (i) Soit $f : \mathcal{U} \rightarrow \mathbb{R}$ admettant des dérivées partielles sur \mathcal{U} . Pour déterminer des extremums locaux de f on cherche ses points critiques.
Si f est de classe \mathcal{C}^2 , on peut s'intéresser à la Hessienne aux points critiques qui permet de conclure lorsqu'elle 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 \mathcal{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 $f : K \rightarrow \mathbb{R}$ 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.

Exercice 13 : Déterminer les extremums de

$$f : (x, y) \in \mathbb{R}^2 \mapsto y^2 - x^2 + \frac{x^4}{2}.$$



Exercice 14 : CCINP 56



b Extremums liés

On cherche désormais les extremums de $f : \mathcal{U} \rightarrow \mathbb{R}$ non pas sur \mathcal{U} entier, mais sur l'ensemble des points annulant une certaine fonction numérique $g \in \mathcal{C}^1(\mathcal{U}, \mathbb{R})$.
Une première approche pour calculer $\min_{g(x)=0} f(x)$ consiste à paramétrer l'ensemble $X = \{x \in \mathcal{U}, g(x) = 0\}$, c'est-à-dire trouver

une fonction $t \in I \mapsto (x_1(t), \dots, x_n(t))$ dont X est l'image. On est alors ramené à un problème à une seule variable que l'on sait résoudre : calculer

$$\min_{t \in I} f(x_1(t), \dots, x_n(t)).$$

Une seconde approche consiste à transformer la contrainte $g(x_1, \dots, x_n) = 0$ en une ou plusieurs expressions de la forme $x_n = \varphi(x_1, \dots, x_{n-1})$ (le théorème hors-programme des fonctions implicites nous assure la possibilité de le faire au voisinage d'un point tel que $\nabla g(x) \neq 0_{\mathbb{R}^n}$) puis à calculer

$$\min_{(x_1, \dots, x_{n-1}) \in \mathcal{U}'} f(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1}))$$

où \mathcal{U}' est à préciser avec les techniques précédentes.

Exercice 15 : Déterminer de deux manières différentes

$$\min_{x^2+y^2=1} xy$$

En bleu :

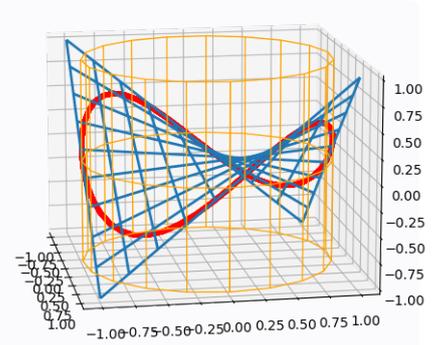
$$z = f(x, y)$$

En orange :

$$g(x, y) = 0$$

En rouge :

$$z = f|_X(x, y)$$



Malheureusement, en général, il n'est pas aisé de trouver un paramétrage de X ou de remplacer la contrainte implicite $g(x_1, \dots, x_n) = 0$ en contrainte explicite $x_n = \varphi(x_1, \dots, x_{n-1})$...

On dispose au programme d'un outil permettant d'étudier ce problème.

Théorème 4 : d'optimisation sous contrainte

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

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 $\nabla g(a)$.

Un scalaire $\lambda \in \mathbb{R}$ tel que $\nabla f(a) = \lambda \nabla g(a)$ est appelé **multiplicateur de Lagrange**.

Cette condition étant seulement nécessaire, la résolution d'un problème d'optimisation sous contrainte s'accompagne en général d'une recherche de point critique (sur un ouvert, bien sûr) suivi d'une étude locale, ou d'un argument de compacité.

Exercice 16 : Déterminer de nouveau $\min_{x^2+y^2=1} xy$

Exercice 17 : (Oral CCINP) Montrer que $f : (x, y) \mapsto 4x^2 + 12xy - y^2$ **admet un minimum et un maximum sur** $\mathcal{C} = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 13\}$ **et les déterminer.**

Exercice 18 : En étudiant l'application $f : (x_1, \dots, x_n) \mapsto x_1 \cdot \dots \cdot x_n$ **sur** $\mathcal{C}_s = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n, x_1 + \dots + x_n = s\}$, **retrouver l'inégalité arithmético-géométrique.**

**DIFFÉRENTIELLE****1****Différentielle en un point**

Rappel : si I est intervalle de \mathbb{R} , $f : I \rightarrow F$ est dérivable en $a \in I$ si et seulement si f admet un développement limité à l'ordre 1 en a si et seulement si on a un vecteur $b \in F$ tel que

$$f(a+h) = f(a) + hb + \underset{h \rightarrow 0_{\mathbb{R}}}{o}(h)$$

(et, dans ce cas, $b = f'(a)$).

Nous allons généraliser cette idée aux fonctions définies sur une espace vectoriel normé de dimension finie E .

Définition 12 : Application différentiable en un point

Soit f définie sur un ouvert \mathcal{U} de E , à valeurs dans F . Soit a un point de \mathcal{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 ,

$$f(a+h) =$$

ou encore, au voisinage de a ,

$$f(x) =$$

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 $df(a) \in \mathcal{L}(E, F)$. On a donc

$$f(a+h) =$$

Remarque

R 12 – Lorsque cela a du sens, on définit donc une application $df : \begin{cases} E & \rightarrow \mathcal{L}(E, F) \\ a & \mapsto df(a) \end{cases}$ appelée **différentielle** de f .

R 13 – On note parfois $df(a) \cdot h$ au lieu de $df(a)(h)$.

R 14 – L'unicité vient du

Lemme 2

Si $\varphi \in \mathcal{L}(E, F)$ est telle que $\varphi(h) = \underset{h \rightarrow 0}{o}(h)$, alors φ est l'application nulle.

Démonstration

$$\varphi(h) = \|h\| \varepsilon(h) \text{ avec } \varepsilon(h) \xrightarrow{h \rightarrow 0_E} 0_F.$$

Alors pour tout $x \in E$, $t \in \mathbb{R}_*^+$, $t\varphi(x) = \varphi(tx) = t\|x\|\varepsilon(tx)$ donc $\varphi(x) = \|x\| \varepsilon(tx) \xrightarrow{t \rightarrow 0} 0_F$.

$$\text{Donc pour tout } x \in E, \varphi(x) = 0_F. \quad \blacksquare$$

Propriété 9 : différentiable \Rightarrow continue

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

Démonstration

Par continuité de l'application linéaire en dimension finie (celle de E suffit) $df(a)$, en passant à la limite dans le développement limité, $f(a+h) - f(a) \xrightarrow{h \rightarrow 0_E} 0_F$. \blacksquare

2**Cas particuliers****Propriété 10 : Cas d'une fonction d'une variable réelle**

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

Dans ce cas,

Propriété 11 : Cas d'une fonction constante

Si $f : \mathcal{U} \subset E \rightarrow F$ est constante, elle est différentiable en tout point de \mathcal{U} et pour tout $a \in \mathcal{U}$,

Propriété 12 : Cas d'une fonction linéaire

Si $f : \mathcal{U} \subset E \rightarrow F$ 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 $a \in \mathcal{U}$,

Exercice 19 : Montrer que $f : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \rightarrow \mathcal{M}_n(\mathbb{R}) \\ M & \mapsto M^2 \end{cases}$ est différentiable en toute $A \in \mathcal{M}_n(\mathbb{R})$ et calculer $df(A)$.

Exercice 20 : Montrer que, si E est un espace euclidien $f : \begin{cases} E & \rightarrow \mathbb{R} \\ x & \mapsto \|x\|^2 \end{cases}$ est différentiable en toute $a \in E$ et calculer $df(a)$.

Exercice 21 : Montrer que, si E est un espace euclidien, $u \in \mathcal{L}(E)$, $f : \begin{cases} E & \rightarrow \mathbb{R} \\ x & \mapsto (x|u(x)) \end{cases}$ est différentiable en toute $a \in E$ et calculer $df(a)$. Que se passe-t-il si, de plus, u est symétrique ?

3 Dérivée selon un vecteur, dérivées partielles

Le fait de travailler sur un ouvert \mathcal{U} assure, pour a point de \mathcal{U} et v vecteur de E , l'existence d'un $\delta > 0$ (« distance de sécurité ») tel que pour tout $t \in]-\delta, \delta[$, $a + tv \in \mathcal{U}$ (on s'éloigne de a dans la direction de v), cela permet de définir l'application $\phi : t \mapsto f(a + tv)$ d'une variable réelle au voisinage de 0.

Définition 13 : Dérivée selon un vecteur

On dit que $f : \mathcal{U} \subset E \rightarrow F$ est **dérivable selon le vecteur** $v \in E$ au point $a \in \mathcal{U}$, lorsque

On note alors

Exemple

E 15 – La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto \frac{xy^2}{x^2 + y^4}$ si $(x, y) \neq (0, 0)$ et $(0, 0)$ sinon admet des dérivées selon tout vecteur en $(0, 0)$ et si $v = (\alpha, \beta) : D_v f(0, 0) =$

Définition 14 : Dérivées partielles

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

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 :

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

Remarque

R 15 – Les dérivées partielles vues pour les fonctions définies sur un ouvert de \mathbb{R}^p sont les dérivées partielles dans la base canonique de \mathbb{R}^p .

Bien sûr la notion générale de dérivée partielle dépend de la base choisie. En ce sens, la notation $D_{e_j} f(a)$ est la plus précise.

R 16 – Comme on en a déjà l'habitude, calculer la j^{e} dérivée partielle revient à fixer les coordonnées selon tous les autres vecteurs de bases et dériver par rapport à la seule variable x_j . : en effet, en dérive

$$\phi : t \mapsto f(a + te_j) = f(a_1 e_1 + \dots + (a_j + t)e_j + \dots + a_n e_n)$$

en 0, ce qui revient aussi à dériver

$$t \mapsto f(a_1 e_1 + \dots + te_j + \dots + a_n e_n)$$

en a_j .

Exemple

E 16 – Avec $f : (x, y) \in \mathbb{R}^2 \mapsto \frac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et $(0, 0)$ sinon, on voit que Δ on peut avoir des dérivées partielles en $(0, 0)$ sans avoir de dérivée selon certains vecteurs ($v = (\alpha, \beta) \in \mathbb{R}_*^2$, ici).

4 Lien entre différentielle et dérivées partielles

Propriété 13 : Lien entre différentielle et dérivées partielles

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

Cas particulier 2 : Dérivée selon un vecteur de base

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

$$df(a)(e_j) =$$

Propriété 14 : Expression de la différentielle avec les dérivées partielles

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

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

$$df(a)(h) =$$

On note $dx_j : \begin{matrix} E & \rightarrow & \mathbb{R} \\ h & \mapsto & h_j \end{matrix}$ la forme linéaire j^{e} coordonnée dans \mathcal{B} . Alors on a

$$df(a) =$$

Démonstration

C'est la linéarité de $df(a)$:

$$df(a) \left(\sum_{j=1}^p h_j e_j \right) = \sum_{j=1}^p h_j df(a)(e_j) = \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a).$$

Remarque

R 17 – Le DL₁ de f différentiable en a s'écrit alors

$$f(a + h) = f(a) + \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a) + o(h).$$



5 Matrice jacobienne

Définition 15 : Matrice jacobienne

Soit $p = \dim E$, $n = \dim F$, \mathcal{U} ouvert de E et $f : E \rightarrow F$ 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 $J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket}$.

Propriété 15 : Matrice jacobienne et différentielle

$$J_f(a) = \text{Mat}_{\mathcal{B}, \mathcal{C}}(df(a))$$

Démonstration

$$df(a)(e_j) = \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_j}(a) \varepsilon_i$$

permet bien de remplir la j^{e} colonne de $J_f(a)$. ■

6 Gradient

Définition 16 : Gradient

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

$$df(a)(h) = (\nabla f(a)|h).$$

Démonstration

C'est le théorème de représentation de Riesz : la forme linéaire $df(a)$ s'écrit $x \mapsto (b|x)$ pour un certain vecteur $b...$ ■

Remarque

R 18 – On comprend mieux la notation alternative $df(a) \cdot h$ faisant penser à un produit scalaire. Le DL₁ de f différentiable en a dans E euclidien se réécrit

$$f(a+h) = f(a) + \underbrace{(\nabla f(a)|h)}_{\text{produit scalaire}} + o(h)$$

R 19 – Dans \mathbb{R}^2 muni du produit scalaire canonique,

$$df(a)(h) = h_1 \frac{\partial f}{\partial x}(a) + h_2 \frac{\partial f}{\partial y}(a) = \begin{pmatrix} \frac{\partial f}{\partial x}(a) \\ \frac{\partial f}{\partial y}(a) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

On retrouve notre gradient habituel.

Exactement le même raisonnement (expression du produit scalaire en base orthonormée) donne la propriété suivante.

Propriété 16 : Coordonnées du gradient

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

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

Remarque : Interprétation géométrique

R 20 – On suppose f différentiable en a et $\nabla f(a) \neq 0_E$.

Notons S la sphère unité de E pour la norme euclidienne $\|\cdot\|$. D'après l'inégalité de Cauchy-Schwarz

$$\forall h \in S, df(a)(h) = (\nabla f(a)|h) \leq |(\nabla f(a)|h)| \leq \|\nabla f(a)\| \|h\| = \|\nabla f(a)\|.$$

Or ce majorant est atteint pour $h_0 = \frac{\nabla f(a)}{\|\nabla f(a)\|}$.

Ainsi le vecteur $\nabla f(a)$ donne la direction et le sens de plus forte variation de la fonction f .

C'est à la base d'algorithmes d'optimisation (descente de gradient), utilisé par exemple en Machine Learning (apprentissage automatique).



Ci-dessus, on représente, au cœur de Mafate, des « lignes de niveau » de la fonction $(x, y) \mapsto f(x, y)$ qui au point de coordonnées (x, y) associe son altitude (on parle d'isoplethes d'altitude).

Ces lignes de niveau sont des courbes d'équation $f(x, y) = \text{constante}$.

Tracer la direction et le sens de ∇f en quelques points.

On peut montrer que le gradient est normal aux lignes de niveaux en tous points.

IV

OPÉRATIONS SUR LES DIFFÉRENTIELLES

1 Combinaisons linéaires

Propriété 17 : Linéarité

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

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

Remarque

R 21 – $f \mapsto df$ est donc linéaire.

2 Image par une application multilinéaire

Propriété 18 : Image par une application multilinéaire

Soit \mathcal{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, $f_1 : \mathcal{U} \rightarrow E_1, \dots, f_q : \mathcal{U} \rightarrow E_q$ des applications différentiables en $a \in \mathcal{U}$ et $M : E_1 \times \dots \times E_q \rightarrow F$ une fonction q -linéaire.

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

$$d\phi(a) : h \mapsto$$

Démonstration

On traite le cas bilinéaire. Le cas général est analogue.

1^{re} tentative On peut voir le faire avec les dérivées selon les vecteurs. On sait, sous réserve d'existence, que $d\phi(a)(h) = D_h\phi(a)$. Or

$$\psi : t \mapsto \phi(a + th) = B(f(a + th), g(a + th))$$

est dérivable en 0 (car f et g sont différentiables en a) de dérivée

$$\begin{aligned} \psi'(0) &= D_h\phi(a) = B(D_h f(a), g(a)) + B(f(a), D_h g(a)) \\ &= B(df(a)(h), g(a)) + B(f(a), dg(a)(h)) \end{aligned}$$

Le problème est que l'existence de dérivée selon tout vecteur ne garantit pas la différentiabilité...

On peut par exemple montrer que la fonction

$$f : (x, y) \mapsto \begin{cases} \frac{y^2}{x} & \text{si } x \neq 0 \\ y & \text{sinon} \end{cases}$$

admet des dérivées selon tout vecteur en $(0, 0)$ mais n'est pas continue, et encore moins différentiable.

Bonne méthode On utilise les DL₁ : avec $\varepsilon_1(h) \xrightarrow{h \rightarrow 0} 0$ et $\varepsilon_2(h) \xrightarrow{h \rightarrow 0} 0$,

$$\begin{aligned} \phi(a + h) &= B(f(a + h), g(a + h)) \\ &= B(f(a) + df(a)(h) + \|h\| \varepsilon_1(h), g(a) + dg(a)(h) + \|h\| \varepsilon_2(h)) \\ &= B(f(a), g(a)) + B(df(a)(h), g(a)) + B(f(a), dg(a)(h)) + \psi(h) \end{aligned}$$

avec $h \mapsto B(df(a)(h), g(a)) + B(f(a), dg(a)(h))$ linéaire et

$$\begin{aligned} \psi(h) &= \|h\| B(df(a)(h), \varepsilon_2(h)) + \|h\| B(\varepsilon_1(h), g(a) + dg(a)(h)) \\ &\quad + \|h\|^2 B(\varepsilon_1(h), \varepsilon_2(h)) + B(df(a)(h), dg(a)(h)) \end{aligned}$$

donc

$$\begin{aligned} \frac{\psi(h)}{\|h\|} &= B(df(a)(h), \varepsilon_2(h)) + B(\varepsilon_1(h), g(a) + dg(a)(h)) \\ &\quad + \|h\| B(\varepsilon_1(h), \varepsilon_2(h)) + \frac{B(df(a)(h), dg(a)(h))}{\|h\|}. \end{aligned}$$

Les trois premiers termes tendent vers 0 lorsque $h \rightarrow 0$ par continuité de l'application bilinéaire en dimension finie.

Pour le dernier, on remarque que

$$(x, y) \mapsto B(df(a)(x), dg(a)(y))$$

est bilinéaire, donc on a, par continuité, un réel C tel que

$$\begin{aligned} \left\| \frac{B(df(a)(h), dg(a)(h))}{\|h\|} \right\| &= \frac{\|B(df(a)(h), dg(a)(h))\|}{\|h\|} \\ &\leq \frac{C \|h\|^2}{\|h\|} = C \|h\| \xrightarrow{h \rightarrow 0} 0. \end{aligned}$$

Donc $\psi(h) = o_{h \rightarrow 0}(h)$, ce qui permet d'obtenir toute la conclusion. ■

3 Composition

Propriété 19 : Différentielle d'une composée

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

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

$$d(g \circ f)(a) =$$

Propriété 20 : Matrice jacobienne d'une composée

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

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

Remarque

R22 – D'où la règle de la chaîne.

Démonstration

Il suffit de composer les DL₁ : on écrit, avec $\varepsilon_1(h) \xrightarrow{h \rightarrow 0} 0$ et $\varepsilon_2(k) \xrightarrow{k \rightarrow 0} 0$, et pour simplifier la lecture, les applications linéaires $\varphi = df(a)$ et $\psi = dg(f(a))$,

$$f(a + h) = f(a) + \varphi(h) + \|h\| \varepsilon_1(h)$$

et

$$g(f(a) + k) = g(f(a)) + \psi(k) + \|k\| \varepsilon_2(k)$$

puis

$$\begin{aligned} g(f(a + h)) &= g\left[f(a) + \underbrace{(\varphi(h) + \|h\| \varepsilon_1(h))}_k\right] \\ &= g(f(a)) + \psi(\varphi(h) + \|h\| \varepsilon_1(h)) \\ &\quad + \|\varphi(h) + \|h\| \varepsilon_1(h)\| \varepsilon_2(\varphi(h) + \|h\| \varepsilon_1(h)) \\ &= g \circ f(a) + [\psi \circ \varphi](h) + \zeta(h) \end{aligned}$$

avec

$$\frac{\zeta(h)}{\|h\|} = \psi(\varepsilon_1(h)) + \frac{\|\varphi(h) + \|h\| \varepsilon_1(h)\|}{\|h\|} \varepsilon_2(\varphi(h) + \|h\| \varepsilon_1(h)).$$

φ et ψ étant linéaires en dimension finie, elles sont continues, donc $\psi(\varepsilon_1(h)) \xrightarrow{h \rightarrow 0} 0$, $\varphi(h) \xrightarrow{h \rightarrow 0} 0$ puis $\varepsilon_2(\varphi(h) + \|h\| \varepsilon_1(h)) \xrightarrow{h \rightarrow 0} 0$.

Enfin,

$$\frac{\|\varphi(h) + \|h\| \varepsilon_1(h)\|}{\|h\|} \leq \frac{\|\varphi(h)\|}{\|h\|} + \varepsilon_1(h)$$

et par continuité de l'application linéaire φ , on a $C \in \mathbb{R}$ tel que pour tout h , $\|\varphi(h)\| \leq C \|h\|$ donc $\frac{\|\varphi(h) + \|h\| \varepsilon_1(h)\|}{\|h\|}$ est

borné et finalement $\frac{\zeta(h)}{\|h\|} \xrightarrow{h \rightarrow 0} 0$, c'est-à-dire $\zeta(h) = o(h)$.

Par unicité, on en déduit que $g \circ f$ est différentiable en a et

$$d(g \circ f)(a) = \psi \circ \varphi = dg(f(a)) \circ df(a).$$

**Corollaire 4 : Dérivée le long d'un arc**

Soit \mathcal{U} ouvert de E , $f: \mathcal{U} \rightarrow F$, I intervalle de \mathbb{R} et $\gamma: I \rightarrow \mathcal{U}$ 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

$$(f \circ \gamma)'(t) =$$

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

$$(f \circ \gamma)'(t) =$$

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

$$(f \circ \gamma)'(t) =$$

Démonstration

On a vu que $df(\gamma(t))(\gamma'(t)) = (f \circ \gamma)'(t)$, c'est donc une simple application du théorème fondamental de l'analyse. ■

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

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

$$f(a+v) - f(a) =$$

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

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

f est constante si et seulement si $df = 0$.

Démonstration

Le sens direct a déjà été vu.

Pour le sens réciproque, seul le cas convexe est au programme. On suppose donc \mathcal{U} convexe et $df = 0$.

Soient $a, b \in \mathcal{U}$, on veut montrer que $f(a) = f(b)$.

Mais on peut relier a et b par un segment inclus dans \mathcal{U} par convexité : $\gamma: t \in [0, 1] \mapsto (1-t)a + tb$ de classe \mathcal{C}^1 .

Alors, la propriété précédente donne $f(b) - f(a) = 0$. ■

V CLASSE \mathcal{C}^1 **Définition 17 : Fonctions de classe \mathcal{C}^1**

$f: \mathcal{U} \rightarrow F$ est dite **de classe \mathcal{C}^1** lorsque f est différentiable sur \mathcal{U} et $df: \mathcal{U} \rightarrow \mathcal{L}(E, F)$ est continue.

Propriété 21 : Opérations

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

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

Propriété 22 : Caractérisation par les applications coordonnées

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

Théorème 5 : IMPORTANT - Caractérisation avec les dérivées partielles

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

Démonstration

Non exigible. ■

Propriété 23 : Expression intégrale le long d'un arc

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

$$f(b) - f(a) =$$

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 $\gamma:]-\varepsilon, \varepsilon[\rightarrow X$, 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 .

Exemple

E 17 – Si \mathcal{G} est un sous-espace affine de E de direction G (sous-espace vectoriel de E), pour tout $a \in \mathcal{G}$,

$$T_a \mathcal{G} = G$$

En effet, on peut écrire $\mathcal{G} = a + G$.

Si $v \in G$, soit $\gamma: \begin{cases}]-1, 1[& \rightarrow & \mathcal{G} \\ t & \mapsto & a + tv \end{cases}$ est dérivable en 0, $\gamma(0) = a$ et $\gamma'(0) = v$.

Réciproquement, si $v \in T_a \mathcal{G}$, on a $\varepsilon > 0$ et $\gamma:]-\varepsilon, \varepsilon[\rightarrow \mathcal{G}$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

Alors v est la limite lorsque $t \rightarrow 0$ de $\frac{\gamma(t) - a}{t} \in G$ et comme G est un sous-espace de E qui est de dimension finie, il l'est aussi, donc il est fermé et donc $v \in G$.

É 18 – Sphère d'un espace euclidien : si $a \in S(0_E, 1)$,

$$T_a S(0_E, 1) = a^\perp$$

En effet, si $v \in a^\perp$, soit $\gamma : \begin{cases}]-1, 1[\rightarrow S(0_E, 1) \\ t \rightarrow \frac{a+tv}{\|a+tv\|} \end{cases}$ est dérivable en 0, $\gamma(0) = a$ et $\gamma'(0) = v$ car, avec $\|a\| = 1$,

$$\gamma : t \mapsto \frac{a+tv}{\sqrt{\|a+tv\|^2}} = \frac{a+tv}{\sqrt{1+2t(a|v) + t^2\|v\|^2}}$$

donc

$$\gamma'(0) = \frac{\|a+0 \cdot v\|v - \frac{2(a|v)+2 \cdot 0 \cdot \|v\|^2}{2\|a+0 \cdot v\|}(a+0 \cdot v)}{\|a+0 \cdot v\|^2}$$

Réciproquement, si $v \in T_a S(0_E, 1)$, on a $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\rightarrow S(0_E, 1)$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

Alors $\frac{\gamma(t)-a}{t} \xrightarrow{t \rightarrow 0} v$ donc

$$\begin{aligned} \left(\frac{\gamma(t)-a}{t} \middle| a \right) &= \frac{(\gamma(t)|a) - 1}{t} = \frac{(\gamma(t)|a - \gamma(t))}{t} \\ &= - \left(\gamma(t) \middle| \frac{\gamma(t)-a}{t} \right) \xrightarrow{t \rightarrow 0} (v|a) = -(a|v) \end{aligned}$$

Donc $(v|a) = 0 : v \in a^\perp$.

Comme dernier exemple, on va obtenir que le plan tangent à une surface d'équation $z = f(x, y)$ admet comme vecteur normal le gradient (non nul) de $g : (x, y, z) \mapsto f(x, y) - z$.

Cela revient à transformer l'équation explicite $z = f(x, y)$ en équation implicite $g(x, y, z) = 0$.

Propriété 24 : Plan tangent pour une surface explicite

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

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

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

Soit $g : \begin{cases} \mathcal{U} \times \mathbb{R} \rightarrow \mathbb{R} \\ (x, y, z) \mapsto f(x, y) - z \end{cases}$

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

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

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 S en a , d'équation

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

Remarque

R 23 – On retrouve, pour le plan tangent, une équation type « équation de la tangente » :

Démonstration

On raisonne toujours par double inclusion.

■ Si $v = (x, y, z) \in T_a S$, on a $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\rightarrow S$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$. Alors, pour tout $t \in]-\varepsilon, \varepsilon[$, $\gamma(t) = (x(t), y(t), z(t))$ avec $z(t) = f(x(t), y(t))$ donc

$$\begin{aligned} z = z'(0) &= x'(0) \frac{\partial f}{\partial x}(x(0), y(0)) + y'(0) \frac{\partial f}{\partial y}(x(0), y(0)) \\ &= x \frac{\partial f}{\partial x}(x_0, y_0) + y \frac{\partial f}{\partial y}(x_0, y_0) \end{aligned}$$

donc

$$0 = x \frac{\partial f}{\partial x}(x_0, y_0) + y \frac{\partial f}{\partial y}(x_0, y_0) - z = (\nabla g(a)|(x, y, z))$$

car $\nabla g(a) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right)$.

■ Si $v = (x, y, z)$ est orthogonal à $\nabla g(a)$, On définit

$$\gamma : \begin{cases}]-1, 1[\rightarrow S \\ t \mapsto (x_0 + tx, y_0 + ty, f(x_0 + tx, y_0 + ty)) \end{cases}$$

Alors γ est dérivable en 0, $\gamma(0) = a$ et $\gamma'(0) = \left(x, y, x \frac{\partial f}{\partial x}(x_0, y_0) + y \frac{\partial f}{\partial y}(x_0, y_0) \right) = (x, y, z)$ car $(\nabla g(a)|v) = 0$.

Mais toutes les surfaces n'ont pas une équation explicite $z = f(x, y)$.

Le programme propose un dernier énoncé, à la démonstration hors programme, permettant de traiter le cas général des surfaces décrites par une équation implicite $g(x, y, z) = 0$ et généralisant l'énoncé précédent.

Propriété 25 : Hyperplan tangent à une surface implicite

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

Si $dg(a) \neq 0_{\mathcal{L}(E, \mathbb{R})}$, alors

$$T_a X = \text{Ker}(dg(a)) = \nabla g(a)^\perp$$

(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

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

Remarque

R 24 – Bien sûr le **vecteur** x de cet énoncé ne doit pas être confondu avec le **scalaire** x d'un triplet (x, y, z) du cas particulier \mathbb{R}^3 .

**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 du dit hyperplan tangent en un point a n'annulant pas la dg .

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 $g(a) = g(x_0, y_0, z_0) = 0$.

Exemple

E 19 – Équation du plan tangent en un point d'une sphère de \mathbb{R}^3

Démonstration

Soit $v \in T_a X$. On a $\gamma :]-\varepsilon, \varepsilon[\rightarrow X$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$. Alors, $f \circ \gamma$ admet un extremum local en $0 \in]-\varepsilon, \varepsilon[$ qui est ouvert, donc

$$0 = (f \circ \gamma)'(0) = df(\gamma(0))(\gamma'(0)) = df(a)(v).$$

Remarque

R 25 – Autrement dit, les vecteurs tangents à X en a extremum local de $f|_X$ sont dans l'hyperplan vectoriel de vecteur normal $\nabla f(a)$.

Théorème 6 : d'optimisation sous contrainte

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

Si $f|_X$ admet un extremum local en $a \in X$ et si $dg(a) \neq 0_E$, alors $df(a)$ est colinéaire à $dg(a)$.

Démonstration

- Soit $df(a)$ est nulle et $df(a) = 0 \cdot dg(a)$.
- Soit $df(a)$ n'est pas nulle et on prouve que $\text{Ker}(df(a)) = \text{Ker}(dg(a))$. En effet, un résultat du programme dit que si deux formes linéaires non nulles ont même noyau, elles sont colinéaires.

Comme $dg(a) \neq 0$, et avec la propriété précédente, $\text{Ker}(dg(a)) = T_a X \subset \text{Ker}(df(a))$.

Par égalité des dimensions, $\text{Ker} df(a) = \text{Ker} dg(a)$, ce qui conclut. ■

Terminons avec une extension hors programme, pour la culture.

Théorème 7 : HP : multiplicateurs de Lagrange

Soit f, g_1, \dots, g_p des fonctions numériques de classe \mathcal{C}^1 sur l'ouvert \mathcal{U} de E et

$$X = \{x \in \mathcal{U}, g_1(x) = \dots = g_p(x) = 0\}.$$

Si $f|_X$ admet un extremum local en $a \in X$ et si les formes linéaires $dg_1(a), \dots, dg_p(a)$ sont linéaires indépendantes, alors il existe $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ tels que

$$df(a) = \lambda_1 dg_1(a) + \dots + \lambda_p dg_p(a).$$

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

VII**OPTIMISATION (RECHERCHE D'EXTREMUMS)****Définition 19 : Point critique**

Soit $f : \mathcal{U} \rightarrow \mathbb{R}$ différentiable. On dit que a est un **point critique** de f lorsque $df(a) = 0$.

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

Propriété 26 : Condition nécessaire d'extremum local

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

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

Démonstration

Il suffit de revenir aux dérivées partielles. ■

Propriété 27 : Vecteurs tangents en un extremum local

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

$$\forall v \in T_a X, df(a)(v) = 0.$$

Autrement dit, dans le cas où E est euclidien,

$$T_a X \subset \text{Ker}(df(a)) = (\nabla f(a))^\perp$$