Calcul différentiel

I Vocabulaire

I.1 Définitions

Soient E et F deux espaces vectoriels (réels) de dimension finie . Soit Ω un ouvert de E.

Dans ce chapitre on se propose d'étudier des fonctions du type :

$$f : x \in \Omega \longmapsto f(x) \in F$$

Communément, une telle fonction s'appelle une fonction d'une variable vectorielle.

Modulo le choix de $(e_1, ..., e_n)$ une base de E, toute fonction $f : x \in \Omega \longmapsto f(x)$ d'une variable vectorielle s'identifie à $\widetilde{f} : (x_1, ..., x_n) \longmapsto f(x_1 e_1 + \cdots + x_n e_n)$.

On définit donc également :

Définition (fonction de n variables réelles)

Soit Ω un ouvert de $\mathbb{R}^n$, soit F un sous-espace vectoriel de dimension finie. Une application

$$f : (x_1, ..., x_n) \in \Omega \longmapsto f(x_1, ..., x_n) \in F$$

s'appelle une fonction de n variables réelles (à valeurs vectorielles à priori).

I.2 Coordonnées d'une fonction vectorielle

Modulo le choix d'une base $B_F = (u_1, ..., u_p)$ de F, on peut identifier f à :

$$\widetilde{f} : \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R}^p \\ x & \longmapsto & \begin{pmatrix} f_1(x) \\ \vdots \\ f_p(x) \end{pmatrix} \end{array} \right.$$

Les applications $f_1, ..., f_p$ sont appelées **coordonnées** de f dans B_F. Ce sont évidemment des applications à valeurs réelles. Nous verrons dans la suite l'importance de ce choix.
Exemples :
l'application exponentielle complexe $z \mapsto e^z$ peut être vue comme la fonction de deux variables réelle :

$$f(x, y) =$$

dont les coordonnées sont les deux fonctions de deux variables réelles

I.3 Restriction à une direction

Soit Ω un ouvert de E, soit $a \in \Omega$, soit $u \in E$ un vecteur non nul. On note $\mathcal{D}_{u,a}$ la droite $x_0 + \text{Vect}(u)$. $\mathcal{D}_{u,a}$ admet le paramétrage $\phi_{u,a} : t \longmapsto a + tu$. Si f est définie sur Ω, alors

$$f \circ \phi_{u,a} : t \mapsto f(a + tu)$$

est définie au voisinage de 0 et est une fonction d'une seule variable réelle . On l'appellera restriction de f selon le vecteur u en a.

Représentation géométrique :

I.4 Applications partielles en un point a

Si $B = (e_1, ..., e_n)$ est une base et si u est le vecteur e_i on définit :

Définition

Soit $f : \Omega \longrightarrow \mathbb{R}$ une fonction de n variables réelles. En un point $a = (a_1, ..., a_n)$ on définit :

$$\left| \begin{array}{lll} I_a \subset \mathbb{R} & \longrightarrow & F \\ x_i & \longmapsto & f(a_1, ..., a_{i-1}, x_i, a_{i+1}, ..., a_n) \end{array} \right.$$

Cette application s'appelle la $i^{\text{ème}}$ application partielle dans la base B au point a.

Cette application définie au voisinage (réel) de a_i n'est qu'un cas particulier de restriction à une direction puisque c'est l'application $t \mapsto f(a + (t - a_i)e_i)$.

Ces applications joueront un rôle important par la suite.

Exemple. Déterminer les applications partielles de f en tout point (x_0, y_0) ou f est définie par :

$$f : \left| \begin{array}{lll} \mathbb{R}^2 \backslash \{(0,0)\} & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & \dfrac{xy}{x^2 + y^2} \end{array} \right.$$

I.5 Graphe

On appelle graphe de l'application f l'ensemble

$$\Gamma_f = \{(x, f(x)), x \in \Omega\}$$

Ce graphe est une partie de l'espace produit $E \times F$. Comme on ne s'intéresse ici qu'au cas des fonctions ayant au moins deux variables réelles, on ne peut le représenter que lorsque E est de dimension 2 et f à valeurs réelles. Dans ce cas, le graphe de f s'interprète comme une surface de $\mathbb{R}^3$.

Cette représentation s'avérera très utile pour comprendre les objets intervenant dans la suite.
L'objectif du chapitre est de dériver les fonctions que nous venons d'introduire. Avant celà, commençons par un rappel sur la continuité.

II Continuité

II.1 Rappel de définitions et propriétés

On considère une application $f : \Omega \subset E \longrightarrow F$. f est continue en x_0 lorsque :

$$\forall \epsilon > 0,\ \exists \alpha > 0,\ \|h\|_E < \alpha \Longrightarrow \|f(x) - f(x_0)\|_F \leqslant \epsilon$$

Souvenons-nous qu'en dimension finie, la continuité d'une fonction ne dépend pas de la norme choisie. De plus, une fonction $f \in F^E$ est continue si et seulement si chacune de ses composantes dans une base de F est continue.

La continuité est conservé par les opérations linéaires, les produits et les compositions. En particulier, pour les fonctions de n variables réelles :

- Les applications polynômiales des coordonnées sont continues
- Les applications linéaires ou bilinéaires sont continues
- Les composées de fonctions continues sont continues

II.2 Continuité selon une direction. Continuité partielle

Définition

Soit $a \in \Omega$, soit $u \neq 0$. On dit que f est continue en a selon u lorsque $t \longmapsto f(a + tu)$ est continue en 0.

De même, on dit que $f : (x_1, ..., x_n) \longmapsto f(x_1, ..., x_n)$ est continue en $(a_1, ..., a_n)$ selon x_i lorsque $x_i \longmapsto f(a_1, ..., x_i, ...a_n)$ est continue en a_i.

Proposition

Si f est continue en x_0, alors f est continue en x_0 selon tout vecteur. En particulier, les applications partielles sont continues en x_0.

Démonstration. Comme f est continue en x_0, et que $t \longmapsto x_0 + tu$ est continue en 0, l'application composée $t \longmapsto f(x_0 + tu)$ est donc continue en 0.

Attention :La réciproque est fausse !
L'exemple suivant est canonique.
l'application f suivante est continue en 0 selon x et selon y mais n'est pas continue en 0 :

$$f : \left| \begin{array}{lcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases} \end{array} \right.$$

Remarque. On peut exhiber des exemples (voir TD) de fonctions qui sont continues en un point selon toute direction sans être continues en ce point.

II.3 Technique d'étude de la continuité

i. On repère les points problématiques (c'est à dire les points où les théorèmes généraux ne marchent pas. En général il n'y en a pas, ou très peu)
ii. En un point problématique x_0 :
   - Si on soupçonne une discontinuité, on essaie d'exhiber une direction de non continuité.
   - Si on soupçonne la continuité, on essaie de majorer $\|f(x+h) - f(x)\|$ à l'aide de $\|h\|$.

Exemple. Etudier la continuité de la fonction f suivante :

$$f : \left| \begin{array}{lcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \begin{cases} \dfrac{x^2y}{(x^2+y^2)^\alpha} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases} \end{array} \right.$$

Exemple. Montrer que la fonction suivante est continue sur $\mathbb{R}^2$:

$$f : \left| \begin{array}{lcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \begin{cases} \dfrac{e^y - e^x}{y-x} & \text{si } x \neq y \\ 0 & \text{sinon} \end{cases} \end{array} \right.$$

III Dérivée selon un vecteur, dérivée partielle

III.1 Dérivée selon un vecteur

Définition

Soit $f : \Omega \subset E \longrightarrow F$, soit $x_0 \in \Omega$ et $u \neq 0$. On dit que f possède une dérivée en a selon le vecteur u lorsque $t \longmapsto f(a + tu)$ est dérivable en 0, autrement dit lorsque :

$$\lim_{t \to 0} \frac{f(a + tu) - f(a)}{t} \quad \text{existe}$$

En cas d'existence, on la notera $\partial_u f(a)$ ou $D_u f(a)$ ou parfois $f'_u(a)$.

Remarquons qu'en cas d'existence de la limite, celle-ci appartient à F.

III.2 Dérivées partielles.

Définition

Soit $f : \Omega \subset \mathbb{R}^n \longrightarrow F$, soit $a \in \Omega$. On dit que f possède une $i^{\text{ème}}$ dérivée partielle en $a = (a_1, ..., a_n)$ lorsque $t \longmapsto f(a_1, ..., t, ..., a_n)$ est dérivable en a_i.

Cette dérivée se note $\partial_i f(a)$ ou $D_{e_i} f(a)$ ou parfois $f'_{e_i}(a)$, mais le plus souvent $\frac{\partial f}{\partial x_i}(a)$

Modulo le choix d'une base B de E, on peut définir pour $f : \Omega \subset E \longrightarrow F$ les dérivées partielles de f dans B.

III.3 Propriétés

Les dérivées selon un vecteur on les mêmes propriétés que les dérivées usuelles :
linéarité $\partial_u(f + \lambda g)(a) = \partial_u(f)(a) + \lambda \partial_u(g)(a)$.

produit $\partial_u(fg)(a) = \partial_u(f)(a)g(a) + f(a)\partial_u(g)(a)$.

Méthode de calcul : pour calculer la $i^{\text{ème}}$ dérivée partielles, on fixe toutes les variables sauf x_i, puis on dérive comme si l'on avait affaire à une fonction d'une variable.

Exemple. Déterminer les dérivées partielles de la fonction f de l'exemple précédents.

Remarque Pour une fonction à valeurs vectorielles les dérivées partielles sont des vecteurs. Par exemple les dérivées partielles de l'application $(t, x) \mapsto x^2 exp(tA)$ sont

Pour la dérivée selon un vecteur, on dérive par rapport à t la fonction $f(x_0 + tu)$ en $t = 0$.

Exemples :

Soient x_0 et u deux vecteurs d'un espace euclidien. Calculer la dérivée en x_0 selon u de la fonction $x \mapsto \|x\|^2$.

Calculer la dérivée en M_0 selon la matrice I de la fonction ϕ suivante :

$$\phi : \left| \begin{array}{lcl} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ M & \longmapsto & M^2 \end{array} \right.$$

III.4 Lien avec la continuité et mise en garde

Attention : l'existence d'une dérivée partielle n'implique pas en général celle des autres, ni l'existence des dérivées selon d'autres vecteurs.

Notons le fait plus inattendu suivant :
L'existence des dérivées partielles n'implique pas la continuité de la fonction.

III.5 Interprétation géométrique dans le cas des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$.

IV Différentiabilité

Après avoir exploré le point de vue directionnel, dont l'intérêt est de se ramener à des fonctions d'une variable, nous examinons ici un point le point de vue global puis nous mettons en évidence les liens entre les deux points de vue.

IV.1 Définitions

Définition (petit o vectoriel)

Soit V un voisinage de 0 dans E, et $u : h \in V \longmapsto u(h) \in F$. On dit que $u(h) = o_{h \to 0}(h)$ si $\frac{\|u(h)\|}{\|h\|} \to 0$ quand $\|h\| \to 0$.

Une autre façon de le dire est que chaque coordonnée du vecteur $u(h)$ est un $o(h)$.
Pour alléger les notations, on écrira souvent $o(h)$ à la place de $o_{h\to 0}(h)$.
Avec cette notation, on donne la définition suivante :

Définition

Soit $f : \Omega \in E \longrightarrow F$. Soit $x_0 \in \Omega$. On dit que f est **différentiable** en x_0 lorsqu'il existe $L \in \mathcal{L}(E, F)$ telle que :

$$f(x_0 + h) = f(x_0) + L(h) + o(h)$$

On peut montrer que , en cas d'existence, l'application L est unique, ce qui autorise la définition suivante :

Définition (différentielle en un point)

Si f est différentiable en x_0, l'application linéaire L associée est notée $L = \mathrm{d}f(x_0)$, on l'appelle la différentielle de f en x_0.

Retenir : Si f est différentiable en x_0, on a donc l'égalité

$$f(x_0 + h) = f(x_0) + \mathrm{d}f(x_0)(h) + o(h)$$

Cette égalité s'appelle également le **développement limité à l'ordre** 1 de f en x_0.
La différentielle de f en a s'appelle aussi parfois l'application linéaire tangente à f en a.

Remarque. On note parfois $\mathrm{d}f(x_0)(h) = \mathrm{d}f(x_0) \cdot h$.

Une fois définie la différentielle en chaque point, on peut la définir globalement. C'est la différentielle de f. C'est un objet assez compliqué. Nous le manipulerons peu.

Définition (application différentielle)

Soit $f : \Omega \subset E \longrightarrow F$. On suppose que pour tout $x \in \Omega$, f est différentiable en x. Alors on dit que f est différentiable, et on appelle différentielle de f l'application :

$$\mathrm{d}f : \left| \begin{array}{lcl} \Omega \subset E & \longrightarrow & \mathcal{L}(E, F) \\ x & \longmapsto & \mathrm{d}f(x) \end{array} \right.$$

IV.2 Exemples

Exemple calculatoire

On considère l'application $f : (x, y) \longmapsto \exp\left(x + y^2\right)$. Déterminer la différentielle de f au point $(1, 3)$

Cet exemple fournit une méthode générale en 2 (ou 3) variables : en $x_0 = (a, b)$ on pose $h = (h_1, h_2)$. On fait le DL en h_1, h_2 à l'ordre 1 (tout ce qui est négligeable devant h_1 ou h_2 rentre dans le $o(h)$). Le DL obtenu détermine la différentielle en x_0.

Fonctions d'une variable

Soit $f : I \subset \mathbb{R} \longrightarrow F$. Montrer que f est différentiable en $x \in \mathbb{R}$ si et seulement si f est dérivable en x.

Fonctions constantes et linéaires

La différentielle d'une fonction constante est définie en tout point et c'est l'application linéaire nulle. la différentielle d'une fonction linéaire est constante égale à elle même.

Produit scalaire. Norme euclidienne

IV.3 Opérations algébriques

Proposition

Si f et g sont différentiables en x_0, alors $f + \lambda g$ l'est aussi, avec $\lambda \in \mathbb{R}$.

Proposition

Si B est une forme bilinéaire, et que f et g sont différentiables en x_0, alors $h = B(f, g)$ est différentiable en x_0 et :

$$\mathrm{d}h(x_0)(x) = B(f(x_0), \mathrm{d}g(x_0)(x)) + B(\mathrm{d}f(x_0)(x), g(x_0))$$

La démonstration se fait comme pour la formule de dérivation d'un produit.

IV.4 Lien avec les notions précédentes

Proposition

Si f est différentiable en x_0, alors f est continue en x_0.

Démonstration. En écrivant le développement limité à l'ordre 1 en x_0 :

$$f(x_0 + h) - f(x) = \mathrm{d}f(x_0)(h) + o(h) \xrightarrow[h \to 0]{} 0$$

Proposition

Si f est différentiable en x_0, alors f possède une dérivée selon tout vecteur u en x_0 (réciproque fausse). De plus, cette dérivée est :

$$\partial_u f(x_0) = \mathrm{d}f(x_0)(u)$$

En particulier, si f est différentiable, alors f possède des dérivée partielles et :

$$\boxed{\frac{\partial f}{\partial x_i}(x_0) = \mathrm{d}f(x_0)(e_i)}$$

Démonstration.

Corollaire

Soit f une fonction différentiable en x_0 à valeurs dans $\mathbb{R}$. Alors :

$$\mathrm{d}f(x_0)(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_0)$$

avec $h = (h_1, ..., h_n)$.
En particulier, la dérivée selon tout vecteur u se calcule dans ce cas à l'aide des dérivées partielles.

Remarque. C'est cette formule qui s'écrit en physique

$$\boxed{\mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i}$$

Regardons la traduction de ce résultat dans nos exemples précédents.

IV.5 Matrice Jacobienne

La différentielle en x_0 est une application linéaire. Il est donc naturel d'introduire :

Définition

Soit $f : E \longrightarrow F$ une application différentiable, soit $B = (e_1, ..., e_n)$ une base de E et soit $C = (f_1, ..., f_p)$ une base de F. On appelle matrice Jacobienne de f en x_0 la matrice :

$$Jf(x_0) = \mathrm{Mat}_{B,C}(\mathrm{d}f(x_0))$$

Proposition

Si l'on note $x_1, ..., x_n$ les coordonnées de x dans la base B et $f_1, ..., f_p$ les coordonnées de f dans la base C, alors :

$$Jf(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$$

Ainsi, d$f(x_0)(h)$est le vecteurs de coordonnées dans C la colonne $Jf(x_0) \cdot H$ (où H est le vecteur colonne représentant h dans la base B). On a en particulier $Jf(x_0) \in \mathrm{Mat}_{p,n}(\mathbb{R})$.

Exemple (coordonnées polaires)**.**
On considère f l'application suivante :

$$f : \left| \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (r,\theta) & \longmapsto & (x = r\cos\theta, y = r\sin\theta) \end{array} \right.$$

Alors la matrice Jacobienne de f est :

$$Jf(r,\theta) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Exemple (lien avec la dérivée)**.**
Lorsque f est une fonction réelle de la variable réelle, sa jacobienne $Jf(x)$ est la matrice $(1,1)$ dont l'unique élément est $f'(x)$.

IV.6 Le vecteur gradient pour les fonctions à valeurs réelles

Lorsque $f : E \to \mathbb{R}$, sa matrice jacobienne est le vecteur ligne égal à

$$Jf(x) = \left(\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}\right)$$

et la différentielle est une forme linéaire. Lorsque l'espace E est euclidien (en général on aura $E = \mathbb{R}^n$), on peut utiliser le **produit scalaire** :

Définition

Soit $f : \Omega \subset E \longrightarrow \mathbb{R}$ une fonction différentiable. Alors pour tout $x \in \Omega$, il existe un unique vecteur noté $\nabla f(x)$ tel que :

$$\forall y \in E,\ \mathrm{d}f(x)(y) = \langle \nabla f(x) | y \rangle$$

$\nabla f(x)$ s'appelle le gradient de f en x.

On a implicitement muni le $\mathbb{R}$ espace vectoriel E d'un produit scalaire (en général sur $\mathbb{R}^n$ le produit scalaire sera le produit scalaire canonique). L'existence d'un tel vecteur est une conséquence directe du théorème de Riesz, puisque E est de dimension fini et que $\mathrm{d}f(x)$ est une forme linéaire sur E.

Proposition

Soit $B = (e_1, ..., e_n)$ une base orthonormée de E. Notons $x_1, ..., x_n$ les coordonnées dans B. Alors :

$$\nabla f(x) = \sum_{i=1}^{n} \partial_i f(x) e_i$$

Matriciellement on a donc (dans une BON) $\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$

Démonstration.

Le gradient possède les propriétés algébriques usuelles des dérivées :

Proposition

Si f et g sont deux fonctions différentiables à valeurs dans $\mathbb{R}$ et si $\lambda \in \mathbb{R}$ alors :

$$\nabla(f + \lambda g) = \nabla f + \lambda \nabla g$$

$$\nabla(f \times g) = f \nabla g + g \nabla f$$

Géométriquement, le gradient donne la direction dans laquelle la fonction f varie le plus au point x.

Illustration :

Exemple.
Calculons par deux méthodes le gradient de la norme euclidienne : $f(x) = \|x\|^2$

IV.7 Fonctions de classe $\mathcal{C}^1$

Définition

On dit que $f : \Omega \subset E \longrightarrow F$ est de classe $\mathcal{C}^1$ lorsque :

i. f est différentiable en tout point de Ω

ii. l'application $df : \Omega \to L(E, F)$ est continue sur Ω.

L'important théorème suivant permet d'établir le lien définitif entre les objets que nous venons d'étudier.

Il signale également que le cadre agréable pour étudier les fonctions de n variables est celui des fonctions au moins de classe C^1.

Théorème

Soit $f : \Omega \subset E \longrightarrow F$. Il est équivalent de dire :

i. f est de classe $\mathcal{C}^1$

ii. f possède des dérivées partielles continues sur Ω

Nous avons donc le tableau d'implication suivant :

f continue		f possède des dérivées partielles
$\Uparrow$		$\Uparrow$
f différentiable	$\Longrightarrow$	f possède des dérivées partielles selon tout vecteur
$\Uparrow$		$\Uparrow$
f $\mathcal{C}^1$	$\Longleftrightarrow$	f possède des dérivées partielles continues

Démonstration.

IV.8 Fonction de classe C^k

On définit les fonctions de classe $\mathcal{C}^k$ par récurrence, comme pour les fonctions réelles.

Définition (fonctions de classe $\mathcal{C}^k$)

Soit $f : \Omega \longrightarrow F$. On dit que f est de classe $\mathcal{C}^k$ lorsque :

i. f est de classe $\mathcal{C}^1$ sur Ω

ii. les dérivées partielles de f sont de classe $\mathcal{C}^{k-1}$ sur Ω.

Notations

Soient i, j deux indices. La dérivée partielle $\partial_i(\partial_j f)$ est notée au choix :

$$\partial_{i,j} f = \partial_i \partial_j f = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

(Si $i = j$ on notera plutôt $\frac{\partial^2 f}{\partial x_i^2}$)

Plus généralement, pour k indices $j_1, \ldots, j_k$ on notera $\partial_{j_k} \ldots \partial_{j_1} f$ ou $\frac{\partial^k f}{\partial x_{j_k} \ldots \partial x_{j_1}}$.

Remarque : toute fonction de classe $\mathcal{C}^k$ est aussi de classe $\mathcal{C}^{k-1}$.

Exemple :

Calculons les dérivées partielles secondes de la fonction définie sur $\mathbb{R}^2 \setminus \{(0,0\}$ par $f(x, y) = \dfrac{xy}{x^2 + y^2}$

On constate qu'en fait 2 de ces dérivées sont égales : c'est un fait général

Théorème (Schwarz)

Si f est de classe C^2 alors pour tout i, j et pour tout x on a $\partial_i \partial_j f(x) = \partial_j \partial_i f(x)$

Ce théorème est admis.

Il s'étend bien sûr à la classe C^k. A titre d'exemple, déterminons les dérivées partielles d'ordre k possibles pour une fonction de 2 variables.

IV.9 Complément : Extension des théorèmes d'interversion aux fonctions de plusieurs variables.

On réécrit ici les théorèmes d'analyse pour les fonctions de plusieurs variables réelles. Ils se démontrent exactement comme dans le cas d'une seule variable.
Pour les séries : On se donne pour tout $n \in \mathbb{N}$, $u_n : \Omega \subset E \longrightarrow F$.
On suppose que la série $\sum u_n$ converge simplement. On pose $f(x) = \sum_{n=0}^{+\infty} u_n(x)$.

Théorème (continuité)

On suppose que :

i. pour tout $n \in \mathbb{N}$, u_n est continue sur Ω
ii. $\sum u_n$ converge uniformément sur tout compact de Ω

Alors f est continue sur Ω.

Théorème (dérivation)

On suppose que :

i. Pour tout $n \in \mathbb{N}$, pour tout $i \in [\![1, n]\!]$, u_n possède une dérivée partielle selon x_i
ii. $x \longmapsto \partial_i u_n$ est continue sur Ω
iii. $\sum \partial_i u_n(x)$ converge uniformément sur tout compact de Ω

Alors f est de classe $\mathcal{C}^1$ sur Ω, et $\partial_i f = \sum_{n=0}^{+\infty} \partial_i u_n$.

Pour les intégrales : On se donne une fonction f définie ainsi : $f : \left| \begin{array}{ccc} I \times \Omega & \longrightarrow & F \\ (t, x) & \longmapsto & f(t, x) \end{array} \right.$ On pose, sous-réserve d'existence, $g(x) = \int_I f(t, x) \, \mathrm{d}t$

Théorème (continuité)

On suppose que :

i. $t \longmapsto f(t, x)$ est continue par morceaux
ii. $x \longmapsto f(t, x)$ est continue sur Ω
iii. pour tout K compact de Ω, il existe $\phi_K \in \mathcal{L}^1(I, \mathbb{R})$ tel que pour tout $t \in I$, pour tout $x \in K$, $\|f(t, x)\| \leqslant \phi_K(t)$.

Alors g est continue sur Ω.

Théorème (Leibniz)

On suppose que :

i. $x \longmapsto f(t; x)$ possède des dérivées partielles notées $\partial_i f$ continues sur Ω
ii. pour tout i, $t \longmapsto \partial_i f(t, x)$ est continue par morceaux
iii. pour tout i, pour tout K compact de Ω, il existe $\phi_{i,K} \in \mathcal{L}^1(I, \mathbb{R})$ tel que pour tout $x \in K$, pour tout $t \in I$, $\|\partial_i f(t, x)\| \leqslant \phi_{i,K}(t)$

Alors g est de classe $\mathcal{C}^1$ sur Ω.

V Compositions des dérivées partielles et des différentielles

Nous énonçons d'abord les différents résultats que nous démontrerons à la fin

V.1 Régle de la chaine : cas des arcs.

Situation : On compose une fonction f de n variable par un arc paramétré γ (c'est à à dire une fonction d'une seule variable).

Proposition

On considère $f : \Omega \subset \mathbb{R}^n \longrightarrow F$ et $\gamma : I \subset \mathbb{R} \longrightarrow \Omega$ deux applications $\mathcal{C}^1$. On note $g = f \circ \gamma$, et $g(t) = f(x_1(t), ..., x_n(t))$. Alors :

$$g'(t) = \sum_{i=1}^{n} x_i'(t) \frac{\partial f}{\partial x_i}(x_1(t), ..., x_n(t))$$

Exemples

La formule suivante exprime exactement la même propriété mais en termes de différentielle plutôt que de dérivées partielles.

V.2 Dérivation le long d'un arc

Proposition

Soit γ un arc $\mathcal{C}^1$ sur I à valeurs dans Ω et f est de classe $\mathcal{C}^1$.
La fonction $f \circ \gamma$ est aussi de classe $\mathcal{C}^1$ et :

$$\forall\, t \in I,\ (f \circ \gamma)'(t) = \mathrm{d}f(\gamma(t))(\gamma'(t))$$

Prenons le temps de comprendre les objets et le lien avec la règle de la chaîne.

V.3 Règle de la chaîne : cas général

Proposition (Règle de la chaîne)

On considère f et g deux fonctions de classe $\mathcal{C}^1$, avec :

$$f : \left| \begin{array}{lcl} \Omega_1 \subset \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ (x_1, ..., x_n) & \longmapsto & (f_1(x_1, ..., x_n), ..., f_p(x_1, ..., x_n)) \end{array} \right.$$

$$g : \left| \begin{array}{lcl} \Omega_2 \subset \mathbb{R}^p & \longrightarrow & \mathbb{R}^q \\ (y_1, ..., y_p) & \longmapsto & (g_1(y_1, ..., y_p), ..., g_q(y_1, ..., y_p)) \end{array} \right.$$

Soit $x_0 \in \Omega$. Alors

$$Jg \circ f(x_0) = Jg(f(x_0)) \times Jf(x_0)$$

On a pour tout i, j l'égalité :

$$\frac{\partial f_i}{\partial x_j}(x_0) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial y_k}(f(x_0)) \times \frac{\partial f_k}{\partial x_j}(x_0)$$

Un premier exemple :

Tous les résultats précédents sont des conséquences du théorème suivant :

V.4 Théorème de composition des applications différentiables

Théorème

Soient $f : \Omega \subset E \longrightarrow F$ et $g : \Omega' \subset F \longrightarrow G$ deux applications différentiables. On suppose que f est différentiable en $x_0 \in \Omega$, avec $f(x_0) \in \Omega'$ et g différentiable en $f(x_0)$. Alors $g \circ f$ est différentiable en x_0, et $\mathrm{d}(g \circ f)(x_0) = \mathrm{d}g(f(x_0)) \circ \mathrm{d}f(x_0)$.

Démonstration.

V.5 Exemples d'utilisation

Manipulations élémentaires

Composition avec les fonctions d'une variable :
Soit f une fonction de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ dans $\mathbb{R}$. Dériver $x \longmapsto f(x, x^2)$, et $(x, y) \longmapsto f(x, xy)$.
Soit g une fonction de classe $\mathcal{C}^1$ de $\mathbb{R}$ dans $\mathbb{R}$. Dériver $(x, y) \longmapsto g(y + x^2)$.

Fonctions homogènes

On suppose que f vérifie pour tout $t > 0, x_1, \ldots x_n$ l'égalité $f(tx_1, \ldots, tx_n) = t^\alpha f(x_1, \ldots, x_n)$.

Montrer que pour tout x,

$$\sum_1^n x_i \partial_i f(x) = \alpha f(x)$$

Coordonnées polaires

Soit $f : \Omega \subset \mathbb{R}^2 \longrightarrow F$ une application $\mathcal{C}^1$. On pose $g : (r, \theta) \longmapsto f(r\cos\theta, r\sin\theta)$. Déterminer les dérivées partielles de g en fonction de celles de f.

V.6 Complément : Variation le long d'un arc. Caractérisation des constantes

La dérivation selon un arc est un outil pour démontrer des propriétés des fonctions de plusieurs variables : pour comparer $f(a)$ et $f(b)$ on considère un arc reliant a à b. (si celui-ci est inclus dans Ω).

le résultat ci dessous figure au programme mais n'est pas central : on peut le passer dans une première lecture.

Proposition

Soit $\gamma : [0,1] \to \Omega$ un arc C^1 tel que $\gamma(0) = a$ et $\gamma(1) = b$
Alors

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

Pour trouver un arc joignant a et b le plus simple est de paramétrer le segment $[a,b]$.Dans ce cas, on considère l'arc γ :

$$\gamma : \left| \begin{array}{lll} [0,1] & \longrightarrow & \Omega \\ t & \longmapsto & a + t(b-a) \end{array} \right.$$

Si f est $\mathcal{C}^1$ et que $g(t) = f \circ \gamma$, alors $g'(t) = \mathrm{d}f(a + t(b-a))(b-a)$.
Dans ce cas on a la formule suivante :

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(a + t(b-a))(b-a)dt$$

Voici un exemple d'application de cette technique :

Proposition (caractérisation des constantes)

Soit Ω un ouvert connexe par arcs. Soit $f \in \mathcal{C}^1(\Omega, F)$. Alors f est constante si et seulement si sa différentielle est nulle.

V.7 Composition des dérivées d'ordre 2

On peut également composer les dérivées partielles secondes. Aucunes connaissance théorique n'est au programme mais il est important de manipuler.

L'exemple des coordonnées polaires

V.8 Application à la résolution d'équations aux dérivées partielles

Cette section n'est pas explicitement au programme. Elle est consacrée à des exemples issus de la physique.

L'idée principale de cette partie est que l'on sait résoudre trois types d'équations aux dérivées partielles :

$$(E_1)\ \frac{\partial f}{\partial x} = 0 \qquad (E_2)\ \frac{\partial^2 f}{\partial x^2} = 0 \qquad (E_3)\ \frac{\partial^2 f}{\partial x \partial y} = 0$$

Le but est donc de s'y ramener par composition.

Proposition

Soit $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$. Alors $\frac{\partial f}{\partial x} = 0$ si et seulement s'il existe $\phi \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ telle que pour tout (x, y), $f(x, y) = \phi(y)$.

De cette propriété découlent les deux propositions suivantes.

Proposition

Soit $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$. Alors $\frac{\partial^2 f}{\partial x \partial y} = 0$ si et seulement s'il existe $\phi_1, \phi_2 \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ telle que pour tout (x, y), $f(x, y) = \phi_1(x) + \phi_2(y)$.

Proposition

Soit $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$. Alors $\frac{\partial^2 f}{\partial x^2} = 0$ si et seulement s'il existe $\phi_1, \phi_2 \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ telle que pour tout (x, y), $f(x, y) = x\phi_1(y) + \phi_2(y)$.

Exemple ('Equation de d'Alembert)**.**
À l'aide d'un changement de variable de la forme $u = x + \lambda y$ et $v = x + \mu y$, résoudre l'équation de d'Alembert :

$$\frac{\partial^2 f}{\partial x^2} - \alpha^2 \frac{\partial^2 f}{\partial y^2} = 0$$

Exemple. À l'aide d'un passage en coordonnées polaires, résoudre l'équation suivante :

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = \lambda f$$

Exemple. Chercher les solutions stationnaire (i.e. de la forme $f(x, y) = g(x)h(y)$) de l'équation de la chaleur en dimension 1 :

$$\frac{\partial^2 f}{\partial x^2} + \lambda \frac{\partial f}{\partial y} = 0$$

VI Fonctions à valeurs réelles. Problèmes d'extrema

Dans cette section, toutes les fonctions sont à valeurs réelles.

Leur différentielle est donc déterminée par leur **gradient** ∇f.

Le développement limité de f s'écrit :

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0) | h \rangle + o(h)$$

Le but de cette partie est de généraliser les théorèmes sur les extrema des fonctions d'une variable.

VI.1 Condition nécessaire du premier ordre pour un extremum local

Rappelons d'abord les différentes définitions

Définition (extremum global, extremum local)

On considère $f : \mathcal{D} \subset E \longrightarrow \mathbb{R}$ avec $\mathcal{D}$ non nécessairement ouverte. On dit que f présente un maximum global en $x_0 \in \mathcal{D}$ lorsque $f(x) \leqslant f(x_0)$ pour tout $x \in \mathcal{D}$.

On dit que f présente un maximum local en $x_0 \in \mathcal{D}$ lorsque $x_0 \in \overset{\circ}{\mathcal{D}}$ et qu'il existe $\epsilon > 0$ tel que $f(x) \leqslant f(x_0)$ pour tout $x \in B(x_0, \epsilon)$.

On définit de la même façon un minimum local et global.

Définition (point critique)

Soit $x_0 \in \overset{\circ}{\mathcal{D}}$. On dit que x_0 est un point critique de f lorsque $\nabla f(x_0) = 0$.

On peut désormais énoncer :

Théorème

Si f présente un extremum local en x_0, alors x_0 est un point critique de f.

Démonstration.

Un exemple

Déterminer les extrema de $f : (x, y) \in \mathbb{R}^2 \longmapsto (x - y)^2 + 6(x^2 + y^2)$

La condition nécessaire d'extremum n'est pas suffisante !
Par exemple prenons $f(x, y) = xy$.

Il nous faut donc des arguments supplémentaires pour trouver les extrema. Il sont de deux natures :

VI.2 Utilisation d'arguments topologiques

On peut pour trouver les extrema utiliser le théorème des bornes atteintes : Pour cela il faut étudier f sur un compact.

Premier exemple

On pose $f(x, y) = xy(1 - x - y)$ Trouver le maximum de f sur le domaine $\mathcal{D} = \{(x, y), x \geq 0, y/geq0, x + y \leq 1\}$

Second exemple : fonctions coercives

Soit $f : E \longrightarrow \mathbb{R}$ une application continue telle que $\lim\limits_{\|x\| \to +\infty} f(x) = +\infty$ (on dit que f est coercive). Alors f admet un minimum global.

Corollaire

Si f est $\mathcal{C}^1$ et coercive, alors elle a un point critique.

La seconde condition suffisante d'extremum repose sur un développement à l'ordre 2

VI.3 Développement limité à l'ordre 2 (HP)

Dans cette partie, on s'intéresse à une fonction $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$.

Définition

On appelle matrice hessienne de f en x la matrice :

$$Hf(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j}$$

Proposition

Par le théorème de Schwarz, $Hf(x)$ est une matrice symétrique.

Théorème

Soit $x_0 \in \Omega$. On dispose de la formule de Taylor-Young à l'ordre 2 :

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0) | h \rangle + \frac{1}{2} \langle Hf(x_0) h | h \rangle + o\left(\|h\|^2\right)$$

La preuve de ce théorème, présentée ci-dessous, n'est pas exigible.

Démonstration. Sans perdre de généralité, on se ramène au cas $x_0 = 0$.

$$f(h) - f(0) = \int_0^1 \left[\frac{\mathrm{d}}{\mathrm{d}t} f(th)\right] \mathrm{d}t = \int_0^1 \sum_{i=1}^n h_i \partial_i f(th) \,\mathrm{d}t$$

Pour tout i, $\partial_i f$ est de classe $\mathcal{C}^1$, elle admet donc un développement limité à l'ordre 1 :

$$\partial_i f(h) = \partial_i f(0) + \sum_{i=1}^n h_j \partial_{ij} f(0) + \|h\| \epsilon_i(h)$$

avec $\epsilon_i(h)$ qui tend vers 0 lorsque h tend vers 0. On fixe $\epsilon > 0$, puis α tel que pour tout i, pour tout $h \in B(0, \alpha)$, $\|\epsilon(h)\| \leqslant \epsilon$. On prend $h \in B(0, \alpha)$. On obtient alors :

$$f(h) - f(0) = \left(\sum_{i=1}^n h_i \partial_i f(0)\right) + \frac{1}{2}\left(\sum_{i,j} h_i h_j \partial_i \partial_j f(0)\right) + \|h\| \int_0^1 \sum_{i=1}^n t h_i \epsilon_i(th) \,\mathrm{d}t$$

On conclut en remarquant :

$$\left| \int_0^1 \sum_{i=1}^n t h_i \epsilon_i(th) \,\mathrm{d}t \right| \leqslant \frac{\epsilon}{2} \sum_{i=1}^n \|h_i\|$$

VI.4 Condition suffisante du second ordre pour un extremum local

L'intérêt majeur de la formule de Tayor à l'ordre 2 est de donner une condition suffisante d'extremum local.

Théorème (condition suffisante d'extremum local)

On note $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ le spectre ordonné de $Hf(x_0)$, où x_0 est un point critique de f. Alors :
- Si $Hf(x_0)$ est définie positive (ie $\lambda_1 > 0$) alors x_0 est un minimum local
- Si $Hf(x_0)$ est définie négative (ie $\lambda_n < 0$) alors x_0 est un maximum local
- Si $\lambda_1 < 0$ et $\lambda_n > 0$, x_0 est un point col, ce n'est pas un extremum

(Dans les cas restants, on ne peut rien dire sans une étude plus précise.)
Démonstration :

Exemple

Déterminer à l'aide de la méthode ci-dessus les extrema locaux de la fonction $f(x, y) = x^4 + y^4 - 4xy$

Remarque : Cas particulier de la dimension 2

Proposition

Soit $f : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On suppose que f possède un point critique x_0, et on note $d = \det Hf(x_0)$. Alors :
- Si $d > 0$, f présente un extremum local en x_0
- Si $d < 0$, f présente un point col en x_0

On ne peut conclure dans les autres cas.

Il s'agit d'une conséquence directe de la proposition précédente.

VII Compléments

VII.1 Vecteurs tangents à une partie

Définition

Soit E un espace vectoriel normé de dimension finie, soit $X \subset \Omega$, soit $a \in X$. On appelle vecteur tangent à X en a tout vecteur $u \in E$ qui est vecteur tangent à un arc tracé sur X en a. C'est à dire tel que :

$$\exists\, \gamma \in \mathcal{C}^1(I, X),\ \exists\, t_0 \in I,\ \gamma(t_0) = a \text{ et } \gamma'(t_0) = u$$

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

Exemple (de référence :Vecteurs tangents à la sphère euclidienne)**.**
Soit $S = \{x \in \mathbb{R}^3,\ \|x\|_2 = 1\}$. Pour tout $x_0 \in S$, on a $TS_{x_0} = \langle x_0 \rangle^{\perp}$.

Proposition (vecteurs tangents aux ensembles de niveau)

Soit $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^1$. Soit $C \in \mathbb{R}$. On pose $N_C = \{x \in \Omega,\ f(x) = C\}$. Soit $x_0 \in N_C$. Alors les vecteur tangent à N_C en x_0 sont les vecteurs orthogonaux à $\nabla f(x_0)$.
Autrement dit $TX_{x_0} = \{\nabla f(x_0)\}^{\perp}$

Ce théorème est admis.
Exemple : Tangente à une ellipse.

Plan tangent à une surface

VII.2 Extrema sur un sous ensemble X. Extrema liés

La notion de vecteur tangent permet d'étudier la question des extrema d'une fonction de classe C^1 en restriction à un sous ensemble X. Ceci est utile lorsque cet ensemble X n'est pas ouvert car alors la condition nécessaire d'extremum local n'est plus satisfaite (les variables sont liées).

Lemme

Soit f une fonction de classe C^1 sur un ouvert $\Omega \subset \mathbb{R}^n$ et X une partie de Ω. Soit $x_0 \in X$. Si $f_{|X}$ présente en x_0 un extremum, alors $< \nabla(f)(x_0)|u >= 0$ pour tout vecteur u tangent à X en x_0

Démonstration.

Ce lemme a pour conséquence l'important théorème suivant :

Théorème (des extrema liés)

Soit g une fonction de classe C^1 sur l'espace vectoriel E à valeurs réelles et $X = \{x \in E, g(x) = 0\}$. On suppose que $\nabla g(x)$ **est non nul pour tout** $x \in X$.
Alors pour que la restriction à X d'une fonction numérique f de classe C^1 possède un extremum local en $x_0 \in X$ il est nécessaire que $\nabla f(x_0)$ soit colinéaire à $\nabla g(x_0)$

Démonstration

Ce théorème est extrêmement utile dans les questions d'extrema. En voici plusieurs illustrations.
Exemples