Calcul différentiel

Dans tout le cours, E, F (éventuellement G) sont des $\mathbb{R}$-e.v.n. de dimensions finies p, n respectivement (éventuellement q). Comme dans le chapitre précédent, si $\varphi \in \mathscr{L}(E,F)$ et $v \in E$, on note $\varphi.v$ plutôt que $\varphi(v)$ l'image de v par l'application linéaire φ.

1 Dérivées partielles

1.1 Dérivée selon un vecteur

Définition. Soit U un ouvert de E, $f : U \to F$ et $a \in U$, $v \in E - \{0\}$, t un réel.

On dit que f possède une dérivée en a selon le vecteur v quand la fonction $t \mapsto f(a+tv)$ est dérivable en 0, *i.e.* quand $\dfrac{f(a+tv)-f(a)}{t}$ a une limite finie dans F quand t tend vers 0.

Dans ce cas, on pose $D_v f(a) = \lim\limits_{t\to 0} \dfrac{f(a+tv)-f(a)}{t}$.

1.2 Dérivées partielles dans une base

Définition. Soit $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E. Soit U un ouvert de E, $f : U \to F$ et $a \in U$.

On dit que f possède des dérivées partielles en a dans la base $\mathscr{B}$ quand pour tout $j \in [\![1,p]\!]$, f possède une dérivée en a selon le vecteur e_j.

Dans ce cas, on note $\partial_j f(a) = D_{e_j} f(a) = \lim\limits_{t\to 0} \dfrac{f(a+te_j)-f(a)}{t}$.

Avec les mêmes notations, en notant $x = \sum\limits_{j=1}^{p} x_j e_j$ un vecteur générique de E, on identifie en général les notations $f(x)$ et $f(x_1,\ldots,x_p)$: moyennant le choix d'une base de E, toute fonction de E dans F peut être vue comme une fonction de $\mathbb{R}^p$ dans F, donc comme une fonction de p variables réelles et à valeurs dans F. On note alors aussi $\dfrac{\partial f}{\partial x_j}(a) = \partial_j f(a)$.

Autrement dit, f possède une j-ème dérivée partielle en $a = (a_1,\ldots,a_p)$ quand la fonction $\varphi_j : x_j \mapsto f(a_1,\ldots,a_{j-1},x_j,a_{j+1},\ldots,a_p)$ est dérivable en a_j. Dans ce cas, $\partial_j f(a) = \varphi_j'(a_j)$.

1.3 Absence de lien entre la continuité et l'existence de dérivées selon tout vecteur

Pour une fonction d'une seule variable, l'existence d'une dérivée en un point implique la continuité en ce point.

Dès que $p \geqslant 2$, ce lien est faux : une fonction peut très bien avoir des dérivées selon tout vecteur mais n'être pas continue.

Exercices :

1) Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par : si $(x,y) \neq 0$, $f(x,y) = \dfrac{\sin(x)\sin^2(y)}{x^2+y^2}$ et $f(0,0) = 0$. Vérifiez que f a des dérivées selon tout vecteur en $(0,0)$ et qu'elle est continue en ce point.
2) Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par : si $y \neq 0$, $f(x,y) = \dfrac{x^2}{y}$ et si $y = 0$, $f(x,0) = 0$. La fonction f est-elle continue en $(0,0)$? A-t-elle une dérivée selon un vecteur v en $(0,0)$?
3) Soit $f : (x,y) \mapsto \dfrac{x^3+y^3}{x^2+y^2}$ prolongée en $(0,0)$ en 0. Montrez que f possède des dérivées partielles en tout point de $\mathbb{R}^2$. Ces dérivées partielles sont-elles continues ?
4) Soit $f : \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ définie par : $f(A) = \mathrm{tr}(A^2)$. Montrez que f possède en tout point des dérivées selon tout vecteur.
   En choisissant comme base de $\mathscr{M}_n(\mathbb{R})$ la base canonique et en notant $x_{i,j}$ les coordonnées dans cette base, calculez $\dfrac{\partial f}{\partial x_{i,j}}(A)$.

2 Différentielle

2.1 Application différentiable

Définition. Soit U un ouvert de E, $f : U \to F$ et $a \in U$.
On dit que f est différentiable en a quand il existe une application linéaire $L : E \to F$ continue, un réel $r > 0$ et une application $\varepsilon : B(0,r) \to F$ telles que $\begin{cases} \forall v \in B(0,r) \quad f(a+v) = f(a) + L.v + \|v\|\,\varepsilon(v) \\ \lim\limits_{v \to 0} \varepsilon(v) = 0 \end{cases}$

On note classiquement $o(\|v\|)$ toute expression du type $\|v\|\,\varepsilon(v)$ où $\lim\limits_{v\to 0} \varepsilon(v) = 0$.
La définition précédente affirme donc l'existence d'un d.l. en a à l'ordre 1 : $f(a+v) = f(a) + L.v + o(\|v\|)$.

Remarque. Dans la définition précédente, l'hypothèse de continuité de L est superflue, car E est de dimension finie donc toutes les applications linéaires de E dans F sont continues. Mais au cas où un sujet hors-programme vous placerait dans un espace de dimension infinie E, vous avez la définition complète.

Exercices **:**

5) Soit $f : (x,y) \mapsto y \ln x + e^{xy}$. Montrez que f est différentiable en $(1,1)$.
6) Soit $C : v \mapsto \|v\|^2$ (où la norme est ici la norme euclidienne). Montrez que C est différentiable en tout point. En est-il de même pour l'application norme elle-même ?
7) Soit $f : \mathscr{M}_n(\mathbb{R}) \to \mathscr{M}_n(\mathbb{R})$ définie par : $f(A) = A^2$. Montrez que f est différentiable en tout point de $\mathscr{M}_n(\mathbb{R})$.
8) Soit $g : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ définie par $g(A) = A^{-1}$. Rappelez pourquoi $GL_n(\mathbb{R})$ est un ouvert de $\mathscr{M}_n(\mathbb{R})$.
Montrez qu'il existe $r > 0$ tel que pour tout $V \in B(0,r)$, la série $\sum\limits_{n \geqslant 0} (-V)^n$ converge.
Dans ce cas, que vaut sa somme ?
Montrez que g est différentiable en I_n, puis en tout point de $GL_n(\mathbb{R})$.

Proposition 1. *Avec les mêmes notations, si f est différentiable en a, alors*
- *f est continue en a.*
- *f admet des dérivées en a selon tout vecteur.*

Bien évidemment, la réciproque est fausse.

2.2 Différentielle

Proposition 2. *Avec les mêmes notations, l'application L est unique.*

Dans le cas où f est différentiable en a, l'application L s'appelle la différentielle de f en a ou l'application linéaire tangente en a. Elle est notée $\mathrm{d}f(a)$.
Le d.l. en a à l'ordre 1 est donc

$$f(a+v) = f(a) + \mathrm{d}f(a).v + o(\|v\|)$$

et la dérivée selon le vecteur v en a est

$$D_v f(a) = \mathrm{d}f(a).v$$

2.3 Différentiabilité sur un ouvert

Définition. Soit U un ouvert de E, $f : U \to F$.
On dit que f est différentiable sur U quand f est différentiable en tout point de U. On lui associe donc une unique application $\mathrm{d}f : U \to \mathscr{L}(E,F)$, appelée la différentielle de f sur U.

Remarque. En fait, on devrait noter $\mathrm{d}f : U \to \mathscr{L}_c(E,F)$

Deux cas particuliers :
- si f est constante sur U, alors $\mathrm{d}f = 0$;
- si f est linéaire, alors $\mathrm{d}f = f$.

2.4 Lien avec les dérivées partielles

Proposition 3. *Soit U un ouvert de E, $f : U \to F$ et $a \in U$. Soit $\mathscr{B}$ une base de E.*
Si f est différentiable en a, alors f possède des dérivées partielles en a dans la base $\mathscr{B}$ et pour tout $v \in E$, de coordonnées $(h_1, \dots, h_p)$, $\mathrm{d}f(a).v = \displaystyle\sum_{j=1}^{p} \partial_j f(a)\, h_j = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a)\, h_j$.

Remarque. Dans la notation précédente, on a exceptionnellement fait une entorse à la convention habituelle qui consiste à écrire les produits externes scalaire-vecteur dans le sens λv.

Ici, les dérivées partielles $\partial_j f(a)$ sont des vecteurs de F et les h_j sont des scalaires, on devrait donc noter les produits externes $h_j\, \partial_j f(a)$. Mais l'usage veut que dans le cas des différentielles, on respecte l'ordre des objets plutôt que la convention du produit externe.

Un cas particulier : si $p = 1$, alors f est différentiable en a si et s.si f est dérivable en a et dans ce cas, pour tout $h \in \mathbb{R}$, $\mathrm{d}f(a).h = f'(a)\, h$.

Exercices :

9) Soit $f : (x, y) \mapsto y \ln x + e^{xy}$. En admettant momentanément que f est différentiable sur $\mathbb{R}_+^* \times \mathbb{R}$, calculez sa différentielle en tout point.
10) Soit $g : \mathscr{M}_2(\mathbb{R}) \to \mathscr{M}_2(\mathbb{R})$ définie par : $g(A) = A^\mathsf{T} A$. En notant $A = \begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix}$, calculez la différentielle de g en la matrice $J = (i + j)_{1 \leqslant i, j \leqslant 2}$.

2.5 Caractérisation des fonctions à dérivée partielle nulle

Proposition 4. *Soit f une fonction différentiable sur $\mathbb{R}^2$, ou plus généralement sur un ouvert de la forme $D = I \times J$ où I et J sont deux intervalles ouverts de $\mathbb{R}$.*
Si pour tout $a \in D$, $\dfrac{\partial f}{\partial x}(a) = 0$, alors il existe une fonction à une seule variable v dérivable sur J telle que pour tout $(x, y) \in D$, $f(x, y) = v(y)$.
Si pour tout $a \in D$, $\dfrac{\partial f}{\partial y}(a) = 0$, alors il existe une fonction à une seule variable u dérivable sur I telle que pour tout $(x, y) \in D$, $f(x, y) = u(x)$.

Exprimé de façon plus grossière, si la dérivée partielle par rapport à une variable est constamment nulle, alors la fonction ne dépend pas de cette variable. Bien sûr, ce résultat s'étend à des fonctions à plus de trois variables.

Corollaire 1. *Avec les mêmes hypothèses,*
si pour tout $a \in D$, $\dfrac{\partial f}{\partial x}(a) = \dfrac{\partial f}{\partial y}(a) = 0$, alors f est constante sur D.

Remarque. Ce résultat reste valable sur des ouverts de forme plus générale (voir plus loin) ou en adaptant légèrement l'énoncé, avec plus de trois variables.

Exercices :

11) Déterminez les fonctions f différentiables sur $\mathbb{R}^2$ telle que pour tout $(x, y) \in \mathbb{R}^2$, $\dfrac{\partial f}{\partial x}(x, y) = \lambda$ (λ constante).
12) Faites de même avec la condition $\dfrac{\partial f}{\partial x}(x, y) = \cos x$.
13) Faites de même avec la condition $\dfrac{\partial f}{\partial x}(x, y) = f(x, y)$.

Ce genre de problème s'appelle des équations aux dérivées partielles (EDP).

2.6 Matrice jacobienne

Définition. Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}^n$ et $a \in U$.
Si f est différentiable en a, alors on appelle (matrice) jacobienne de f en a la matrice dans les bases canoniques de $\mathrm{d}f(a)$, souvent notée $Jf(a)$. C'est une matrice de $\mathscr{M}_{n,p}(\mathbb{R})$.

On note $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$ où $f_1, \ldots, f_n$ sont n fonctions de U dans $\mathbb{R}$.

Proposition 5. *Avec les mêmes notations, f est différentiable en a si et s.si pour tout $i \in [\![1, p]\!]$, f_i est différentiable en a.*

Dans ce cas,

$$Jf(a) = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \ldots & \partial_p f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \ldots & \partial_p f_2(a) \\ \vdots & \vdots & & \vdots \\ \partial_1 f_n(a) & \partial_2 f_n(a) & \ldots & \partial_p f_n(a) \end{pmatrix} = \Big(\partial_j f_i(a)\Big)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$$

Exercices :

14) Calculez la jacobienne du changement de variables en polaires $(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$.

2.7 Cas particulier où $F = \mathbb{R}$

Si $n = 1$, alors $f : U \to \mathbb{R}$, donc si f est différentiable en a, alors $\mathrm{d}f(a)$ est une forme linéaire sur E.
Si on a muni E d'un produit scalaire, on appelle alors gradient de f en a, noté $\nabla f(a)$, l'unique vecteur de E tel que

$$\mathrm{d}f(a).v = \nabla f(a) \cdot v$$

(produit scalaire dans E).
Le dév. limité en a à l'ordre 1 devient donc dans ce cas :

$$f(a + v) = f(a) + \nabla f(a) \cdot v + o(\|v\|)$$

et la dérivée selon le vecteur v en a est donc

$$D_v f(a) = \nabla f(a) \cdot v$$

Dans une base orthonormée $\mathscr{B} = (e_1, \ldots, e_p)$, le vecteur $\nabla f(a)$ a pour coordonnées $\begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_p f(a) \end{pmatrix}$,

tandis que la matrice jacobienne est donc une matrice-ligne :

$$Jf(a) = \begin{pmatrix} \partial_1 f(a) & \partial_2 f(a) & \ldots & \partial_p f(a) \end{pmatrix}$$

Remarque. Si $\nabla f(a) \neq 0$, alors $\nabla f(a)$ est positivement colinéaire au vecteur unitaire selon lequel la dérivée de f en a est maximale :
l'application $v \mapsto \nabla f(a) \cdot v = D_v f(a)$, définie sur la sphère-unité, est maximale en le vecteur $v = \dfrac{1}{\|\nabla f(a)\|} \nabla f(a)$.

3 Opérations sur les fonctions différentiables

3.1 Combinaison linéaire

Proposition 6. *Soit U un ouvert de E, $f, g : U \to F$ et $a \in U$.*
Si f et g sont différentiables en a, alors $f + g$ l'est aussi et $\mathrm{d}(f + g)(a) = \mathrm{d}f(a) + \mathrm{d}g(a)$.
Si f est différentiable en a, alors pour tout $\lambda \in \mathbb{R}$, λf l'est aussi et $\mathrm{d}(\lambda f)(a) = \lambda \mathrm{d}f(a)$.

Autrement dit, l'application $f \mapsto \mathrm{d}f(a)$ est linéaire.

3.2 Composition par une application linéaire

Proposition 7. *Soit U un ouvert de E, $f : U \to F$ et $a \in U$. Soit L une application linéaire de F dans G. Si f est différentiable en a, alors $L \circ f$ l'est aussi et $\mathrm{d}(L \circ f)(a) = L \circ \mathrm{d}f(a)$.*

Autrement dit, pour tout vecteur v de E,

$$\mathrm{d}(L \circ f)(a).v = L\Big(\mathrm{d}f(a).v\Big)$$

3.3 Composition par une application k-linéaire

Proposition 8. *Soit U un ouvert de E, $f_1, \dots, f_k$ k fonctions définie sur U à valeurs dans F et $a \in U$. Soit M une application k-linéaire de F^k dans G.*
Si $f_1, \dots, f_k$ sont différentiables en a, alors $M(f_1, \dots, f_k)$ l'est aussi et
$\mathrm{d}\Big[M(f_1, \dots, f_k)\Big](a) = \sum_{i=1}^{k} M(f_1(a), \dots, f_{i-1}(a), \mathrm{d}f_i(a), f_{i+1}(a), \dots, f_k(a)).$

Autrement dit, pour tout vecteur v de E,

$$\mathrm{d}\Big[M(f_1, \dots, f_k)\Big](a).v = \sum_{i=1}^{k} M(f_1(a), \dots, f_{i-1}(a), \mathrm{d}f_i(a).v, f_{i+1}(a), \dots, f_k(a))$$

Un cas particulier important : le produit externe.

Proposition 9. *Si $\lambda : U \to \mathbb{R}$, $g : U \to F$ sont deux fonctions différentiables en a, alors* $\mathrm{d}(\lambda g)(a) = \mathrm{d}\lambda(a)\, g(a) + \lambda(a)\, \mathrm{d}g(a)$.

D'une manière générale, tout produit vérifie le même genre de relation (produit de deux réels, de deux matrices, de deux polynômes, composée de deux endomorphismes, etc).

Exemples.

— en notant $(\ |\ )$ un produit scalaire sur F,
$\mathrm{d}(f|g)(a) = (\mathrm{d}f(a)|g) + (f|\mathrm{d}g(a))$ donc pour tout vecteur v, $\mathrm{d}(f|g)(a).v = (\mathrm{d}f(a).v|g) + (f|\mathrm{d}g(a).v)$
avec le produit scalaire canonique de $\mathbb{R}^2$, $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}$ des fonctions de deux variables (x, y), $v = (h, k)$, on obtient

$$\begin{aligned}
\mathrm{d}(f|g)(a).v &= \mathrm{d}(f_1g_1 + f_2g_2)(a).v \\
&= \mathrm{d}f_1(a).v \times g_1(a) + f_1(a) \times \mathrm{d}g_1(a).v + \mathrm{d}f_2(a).v \times g_2(a) + f_2(a) \times \mathrm{d}g_2(a).v \\
&= \Big(\partial_1 f_1(a)h + \partial_2 f_1(a)k\Big)g_1(a) + f_1(a)\Big(\partial_1 g_1(a)h + \partial_2 g_1(a)k\Big) + \\
&\quad \Big(\partial_1 f_2(a)h + \partial_2 f_2(a)k\Big)g_2(a) + f_2(a)\Big(\partial_1 g_2(a)h + \partial_2 g_2(a)k\Big) \\
&= \Big(\partial_1 f_1(a)g_1(a) + f_1(a)\partial_1 g_1(a) + \partial_1 f_2(a)g_2(a) + f_2(a)\partial_1 g_2(a)\Big)h \\
&\quad \Big(\partial_2 f_1(a)g_1(a) + f_1(a)\partial_2 g_1(a) + \partial_2 f_2(a)g_2(a) + f_2(a)\partial_2 g_2(a)\Big)k
\end{aligned}$$

Quand on a écrit ça une fois dans sa vie, on comprend mieux l'intérêt des notations symboliques !

— en écrivant une matrice carrée A de $\mathscr{M}_k(\mathbb{R})$ dépendant de p variables $x_1, \dots, x_p$ comme une famille de vecteurs-colonnes dépendant eux-mêmes de $x_1, \dots, x_p$: $A = [c_1, \dots, c_k]$,
$\mathrm{d}(\det A)(a).v = \sum_{i=1}^{k} \det(B_i(a, v))$ où $B_i(a, v)$ est la matrice A évaluée en a mais dans laquelle on a remplacé la i-ème colonne $c_i(a)$ par sa différentielle en a évaluée en v $\mathrm{d}c_i(a).v$
avec une matrice (2,2) $A = \begin{pmatrix} b & c \\ e & f \end{pmatrix}$, où b, c, e, f sont des fonctions à deux variables x, y et $a = (x_0, y_0)$, $v = (h, k)$ on obtient

$$\begin{aligned}
\mathrm{d}(\det A)(x_0, y_0).(h, k) &= \begin{vmatrix} \mathrm{d}b(x_0, y_0).v & c(x_0, y_0) \\ \mathrm{d}e(x_0, y_0).v & f(x_0, y_0) \end{vmatrix} + \begin{vmatrix} b(x_0, y_0) & \mathrm{d}c(x_0, y_0).v \\ e(x_0, y_0) & \mathrm{d}f(x_0, y_0).v \end{vmatrix} \\
&= \begin{vmatrix} \partial_1 b(x_0, y_0)h + \partial_2 b(x_0, y_0)k & c(x_0, y_0) \\ \partial_1 e(x_0, y_0)h + \partial_2 e(x_0, y_0)k & f(x_0, y_0) \end{vmatrix} + \begin{vmatrix} b(x_0, y_0) & \partial_1 c(x_0, y_0)h + \partial_2 c(x_0, y_0)k \\ e(x_0, y_0) & \partial_1 f(x_0, y_0)h + \partial_2 f(x_0, y_0)k \end{vmatrix}
\end{aligned}$$

3.4 Composition d'applications différentiables

Ce résultat est souvent appelé règle de la chaîne.

Proposition 10. *Soit U un ouvert de E, V un ouvert de F, $f : U \to V$ et $g : V \to G$, $a \in U$. Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et*

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

Autrement dit, pour tout vecteur v de E, $\mathrm{d}(g \circ f)(a).v = \mathrm{d}g(f(a)).\big(\mathrm{d}f(a).v\big)$ ce qui est noté par convention $\mathrm{d}g(f(a)).\mathrm{d}f(a).v$.

Proposition 11. *Si E est identifié à $\mathbb{R}^p$, F à $\mathbb{R}^n$ et G à $\mathbb{R}^m$ par choix de bases, alors cela se traduit sur les matrices jacobiennes :*

$$J(g \circ f)(a) = Jg(f(a)) \times Jf(a)$$

produit de matrices habituel.

On retrouve la règle de composition des dérivées partielles de MP2I. En particulier, les changements de variables entrent dans cette catégorie des composées de fonctions différentiables.

Exercices :

15) Soit U, V deux ouverts, $f : U \to V$ une bijection différentiable sur U telle que f^{-1} le soit aussi. Que dire de la matrice jacobienne de f en tout point de U ?

Exemple.
Soit f une fonction différentiable de $\mathbb{R}^2$ dans $\mathbb{R}$, définie sur un ouvert D, et x, y 2 fonctions différentiables sur un ouvert U de $\mathbb{R}^2$ tel que pour tout $(u, v) \in U$, $\Phi(u, v) = (x(u, v), y(u, v)) \in D$.
Alors la fonction composée $g : (u, v) \mapsto f(\Phi(u, v))$ est différentiable sur U et on a l'égalité matricielle

$$\begin{pmatrix} \partial_1 g(u,v) & \partial_2 g(u,v) \end{pmatrix} = \begin{pmatrix} \partial_1 f(\Phi(u,v)) & \partial_2 f(\Phi(u,v)) \end{pmatrix} \cdot \begin{pmatrix} \partial_1 x(u,v) & \partial_2 x(u,v) \\ \partial_1 y(u,v) & \partial_2 y(u,v) \end{pmatrix}$$

Ce qui se traduit en clair par

$$\forall (u,v) \in U \quad \left\{ \begin{array}{rcl} \partial_1 g(u,v) & = & \partial_1 f(\Phi(u,v)).\partial_1 x(u,v) + \partial_2 f(\Phi(u,v)).\partial_1 y(u,v) \\ \partial_2 g(u,v) & = & \partial_1 f(\Phi(u,v)).\partial_2 x(u,v) + \partial_2 f(\Phi(u,v)).\partial_2 y(u,v) \end{array} \right.$$

Avec les notations des physiciens, c'est plus clair, à condition de fixer les noms des variables selon leur rang :

$$\forall (u,v) \in U \quad \left\{ \begin{array}{rcl} \dfrac{\partial g}{\partial u}(u,v) & = & \dfrac{\partial f}{\partial x}(\Phi(u,v)).\dfrac{\partial x}{\partial u}(u,v) + \dfrac{\partial f}{\partial y}(\Phi(u,v)).\dfrac{\partial y}{\partial u}(u,v) \\ \dfrac{\partial g}{\partial v}(u,v) & = & \dfrac{\partial f}{\partial x}(\Phi(u,v)).\dfrac{\partial x}{\partial v}(u,v) + \dfrac{\partial f}{\partial y}(\Phi(u,v)).\dfrac{\partial y}{\partial v}(u,v) \end{array} \right.$$

voire même de façon encore plus abrégée

$$\left\{ \begin{array}{rcl} \dfrac{\partial g}{\partial u} & = & \dfrac{\partial f}{\partial x}.\dfrac{\partial x}{\partial u} + \dfrac{\partial f}{\partial y}.\dfrac{\partial y}{\partial u} \\ \dfrac{\partial g}{\partial v} & = & \dfrac{\partial f}{\partial x}.\dfrac{\partial x}{\partial v} + \dfrac{\partial f}{\partial y}.\dfrac{\partial y}{\partial v} \end{array} \right.$$

Remarque. Attention ! Plus on utilise des notations abrégées, plus il y a de sous-entendus ! Donc pour comprendre correctement ces égalités, il faut les replacer dans le contexte, donc ne pas oublier ces sous-entendus.

Un cas particulier important : le passage en coordonnées polaires.
Par exemple, si f est de classe C^1 sur $\mathbb{R}^2$, alors $g : (r, \theta) \mapsto f(r\cos\theta, r\sin\theta)$ est de classe C^1 sur $\mathbb{R}^2$ et

$$\forall (r,\theta) \in \mathbb{R}^2 \quad \left\{ \begin{array}{rcl} \dfrac{\partial g}{\partial r}(r,\theta) & = & \dfrac{\partial f}{\partial x}(\Phi(r,\theta)).\cos\theta + \dfrac{\partial f}{\partial y}(\Phi(r,\theta)).\sin\theta \\ \dfrac{\partial g}{\partial \theta}(r,\theta) & = & -\dfrac{\partial f}{\partial x}(\Phi(r,\theta)).r\sin\theta + \dfrac{\partial f}{\partial y}(\Phi(r,\theta)).r\cos\theta \end{array} \right.$$

Avec ces changements de variable, on peut résoudre quelques EDP simples, la difficulté étant de trouver un bon changement de variables. En pratique, il est presque toujours donné par l'énoncé.

Exercices :

16) Déterminez les fonctions de classe C^1 sur $\mathbb{R}_+^* \times \mathbb{R}$ telles que $x\dfrac{\partial f}{\partial x} + y\dfrac{\partial f}{\partial y} = 0$.

17) Déterminez les fonctions de classe C^1 sur $\mathbb{R}^2$ telles que $2\dfrac{\partial f}{\partial x} - \dfrac{\partial f}{\partial y} = 0$ en posant $x = 2u - v$, $y = v - u$.

18) Déterminez les fonctions de classe C^1 sur $\mathbb{R}_+^* \times \mathbb{R}$ telles que $2xy\dfrac{\partial f}{\partial x} + (1+y^2)\dfrac{\partial f}{\partial y} = 0$ en posant $x = \dfrac{u^2+v^2}{2}$, $y = \dfrac{u}{v}$

3.5 Dérivation le long d'un chemin

On appelle ici chemin une fonction $\gamma : I \to E$ continue sur I, comme dans la définition de connexité par arcs.

Proposition 12. *Soit U un ouvert de E, $f : U \to F$ et γ un chemin de I dans U.*
Si γ est dérivable en $t \in I$ et f est différentiable en $\gamma(t)$, alors $f \circ \gamma$ est dérivable en t et

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

Un cas particulier : si f est à valeurs réelles.

Proposition 13. *Soit U un ouvert de E, $f : U \to \mathbb{R}$ et γ un chemin de I dans U.*
Si γ est dérivable en $t \in I$ et f est différentiable en $\gamma(t)$, alors $f \circ \gamma$ est dérivable en t et

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

Dans le cas où E est identifié à $\mathbb{R}^p$ par choix d'une base, en notant $\gamma(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_p(t) \end{pmatrix}$, on a donc dans ce cas $(f \circ \gamma)'(t) = \displaystyle\sum_{i=1}^{p} \frac{\partial f}{\partial x_j}(\gamma(t))u_i'(t)$.

Si la situation précédente est valable pour tout $t \in I$, alors le support Γ de γ (*i.e.* son image) est une courbe.

- Si pour tout $t \in I$, les vecteurs $\nabla f(\gamma(t))$ et $\gamma'(t)$ sont orthogonaux, alors sur la courbe Γ, la fonction f est constante : on dit que Γ est une ligne de niveau.
- Si pour tout $t \in I$, les vecteurs $\nabla f(\gamma(t))$ et $\gamma'(t)$ sont colinéaires de même sens, alors la courbe Γ est une courbe sur laquelle quand on se déplace, les variations relatives de f sont maximales : on dit que Γ est une ligne de champ de f.
- Par conséquent, si une ligne de niveau et une ligne de champ se croisent en un point, les vecteurs tangents en ce point sont orthogonaux : on dit que les lignes de champ sont orthogonales aux lignes de niveau.

Exemples.

- Les lignes de niveau dans $\mathbb{R}^2$ de la fonction $f : (x, y) \mapsto x^2+y^2$ sont les courbes inscrites dans des cercles ; les lignes de champ sont les courbes inscrites dans les droites passant par l'origine.
- Les lignes de niveau dans $\mathbb{R}^2$ de la fonction $f : (x, y) \mapsto xy$ sont les courbes inscrites dans les hyperboles d'asymptotes (Ox) et (Oy) ; les lignes de champ sont les courbes inscrites dans les hyperboles d'asymptotes les deux bissectrices des axes.

4 Fonctions de classe C^1

4.1 Définition

Si une fonction $f : U \to F$ est différentiable en tout point d'un ouvert U, alors pour tout $a \in U$, $\mathrm{d}f(a)$ est une application linéaire (continue) de E dans F, donc un élément de $\mathscr{L}_c(E, F)$ (qui est égal à $\mathscr{L}(E, F)$ dans ce cours, puisque les dimensions de E et F sont finies). On peut alors définir l'application $\mathrm{d}f : U \to \mathscr{L}(E, F)$ (remarque : $\mathscr{L}(E, F)$ est aussi de dimension finie selon les hypothèses de ce cours).

Définition. On dit que f est de classe C^1 sur U quand $\mathrm{d}f$ est une application continue de U dans $\mathscr{L}(E, F)$.

Un exemple fondamental : les applications linéaires ou k-linéaires.

4.2 Caractérisation

Théorème 1. *Soit U un ouvert de E, $f : U \to F$.*
Alors en identifiant E et $\mathbb{R}^p$ par choix d'une base quelconque, f est de classe C^1 sur U si et s.si f possède des dérivées partielles en tout point de U et si toutes ses dérivées partielles sont continues sur U.

4.3 Opérations sur les fonctions de classe C^1

Grâce aux th. d'opérations et de composition des fonctions continues, il devient évident que
- toute combinaison linéaire d'applications C^1 est C^1 ;
- toute composée de fonctions C^1 est C^1.

Exemples.
- Les applications coordonnées $(x_1, \dots, x_p) \mapsto x_i$ sont de classe C^1, donc toute application de U dans $\mathbb{R}$ qui est polynomiale en ses p variables est elle-même de classe C^1 : les produits scalaires, le déterminant sont des applications de classe C^1.
- On retrouve ainsi que l'application $A \mapsto A^{-1}$ est de classe C^1 sur $GL_n(\mathbb{R})$.

Exercices :

19) La fonction $(x, y) \mapsto \dfrac{y^4}{x^2 + y^2}$ prolongée en $(0, 0)$ par 0 est-elle de classe C^1 sur $\mathbb{R}^2$?

20) Soit $f : \mathbb{R} \to \mathbb{R}$ de classe C^1. On pose $\varphi(x, y) = \dfrac{1}{y - x} \displaystyle\int_x^y f$ si $x \neq y$ et $\varphi(x, x) = f(x)$. Montrez que φ est de classe C^1 sur $\mathbb{R}^2$.

4.4 Caractérisation des fonctions constantes parmi les C^1

Proposition 14. *Soit U un ouvert de E, $f : U \to F$ de classe C^1, a, b deux points de U.*
Alors pour tout chemin γ de $[0, 1]$ dans U de classe C^1 tel que $\gamma(0) = a$ et $\gamma(1) = b$,

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

Théorème 2. *Si U est un ouvert* ***connexe par arcs*** *et f est de classe C^1 sur U, alors f est constante sur U si et s.si $\mathrm{d}f = 0$ sur U.*

Remarque. Attention ! Ce résultat n'est valable sur un ouvert connexe par arcs ! Dans le cas contraire, on est dans une situation analogue à celle rencontrée sur $\mathbb{R}$:
la fonction $x \mapsto \arctan x + \arctan \dfrac{1}{x}$ est de dérivée nulle sur $\mathbb{R}^*$, mais n'est pas constante sur $\mathbb{R}^*$.

5 Vecteurs tangents à une partie

Définition. Soit A une partie de E, $a \in A$, $v \in E$.
On dit que v est un vecteur tangent à A en a quand il existe un chemin γ défini au voisinage de 0, dérivable en 0 et à valeurs dans A tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

L'ensemble des vecteurs tangents en a à A est noté $T_a(A)$. Il contient toujours le vecteur nul. Quand il s'agit d'un sous-espace vectoriel de E, on l'appelle le sous-espace tangent à A en a.

Proposition 15. *Soit U un ouvert de E, $g : U \to \mathbb{R}$ de classe C^1. On pose $A = g^{-1}(\{0\})$.*
Pour tout $a \in$ tel que $\mathrm{d}g(a) \neq 0$, l'ensemble des vecteurs tangents à A en a est le sous-espace $\mathrm{Ker}\, \mathrm{d}g(a)$.

Dans le cas où $dg(a) = 0$, il n'y a pas de résultat général.

Remarque. On note $A = g^{-1}(\{0\})$.

- Dans le cas où $p = 2$, A est appelé une ligne de niveau d'équation $g(x,y) = 0$. Si $M = (a,b) \in A$ et $\mathrm{d}g(M) \neq 0$, alors $T_M(A)$ est la droite vectorielle d'équation $\dfrac{\partial g}{\partial x}(M)\,x + \dfrac{\partial g}{\partial y}(M)\,y = 0$, et $\nabla g(M)$ est un vecteur normal à cette droite.
  La droite affine $M + T_M(A)$ d'équation $\dfrac{\partial g}{\partial x}(M)\,(x-a) + \dfrac{\partial g}{\partial y}(M)\,(y-b) = 0$ est la tangente en M.
- Dans le cas où $p = 3$, A est appelé une surface de niveau d'équation $g(x,y,z) = 0$. Si $M = (a,b,c) \in A$ et $\mathrm{d}g(M) \neq 0$, alors $T_M(A)$ est le plan d'équation $\dfrac{\partial g}{\partial x}(M)(x-a) + \dfrac{\partial g}{\partial y}(M)(y-b) + \dfrac{\partial g}{\partial z}(M)(z-c) = 0$, et $\nabla g(M)$ est un vecteur normal à ce plan.
  Le plan affine $M + T_M(A)$ d'équation $\dfrac{\partial g}{\partial x}(M)(x-a) + \dfrac{\partial g}{\partial y}(M)(y-b) + \dfrac{\partial g}{\partial z}(M)(z-c) = 0$ est le plan tangent en M.

Les physiciens utilisent aussi le terme d'équipotentielle. Par extension, en dimension quelconque, le vecteur $\nabla g(M)$ est appelé vecteur normal à la ligne de niveau d'équation $g(M) = 0$

Exercices :

21) Soit A l'ellipse d'équation $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$. Soit M_0 et M_1 deux points de A tels que les tangentes en M_0 et M_1 se coupent en un point N. Montrez que la droite (ON) passe par le milieu de $[M_0M_1]$.

6 Optimisation au premier ordre

6.1 Vocabulaire

Définition. Soit f une fonction de E dans $\mathbb{R}$, définie sur une partie A et $a_0 \in A$.

On dit que f possède un maximum local sur A en a_0 quand

il existe $r > 0$ tel que pour tout $a \in B(a_0, r) \cap A$, $f(a) \leqslant f(a_0)$.

On dit que f possède un minimum local sur A en a_0 quand

il existe $r > 0$ tel que pour tout $a \in B(a_0, r) \cap A$, $f(a) \geqslant f(a_0)$.

On dit que f possède un extrémum local sur A en a_0 quand f possède un maximum ou un minimum local en a_0.

Définition. Soit f une fonction de E dans $\mathbb{R}$, définie sur une partie A, et $a_0 \in A$.

On dit que f possède un maximum global sur A en a_0 quand pour tout $a \in A$, $f(a) \leqslant f(a_0)$.

On dit que f possède un minimum global sur A en a_0 quand pour tout $a \in A$, $f(a) \geqslant f(a_0)$.

On dit que f possède un extrémum global sur A en a_0 quand f possède un maximum ou un minimum global sur A en a_0.

La recherche des points en lesquels une fonction possède des extrémums (locaux ou globaux) dépend à la fois des propriétés de la fonction et de l'ensemble sur lequel la fonction est définie.

6.2 Points critiques, extréma locaux d'une fonction sur un ouvert

Définition. Soit U un ouvert de E, $f : U \to F$ différentiable sur U.

Un point critique de f est un point $a \in U$ tel que $\mathrm{d}f(a) = 0$.

Comme dans le cours de MP2I, on retrouve la condition nécessaire d'existence d'un extrémum local pour les fonctions à valeurs réelles.

Proposition 16. *Soit U un ouvert de E, $f : U \to \mathbb{R}$ différentiable sur U, $a \in U$.*

Si f possède un extrémum local en a, alors a est un point critique de f.

Remarque.

- La réciproque est fausse (contre-exemple : la selle de cheval $(x,y) \mapsto x^2 - y^2$).
- Ce résultat n'est valable que sur un ouvert, donc en particulier en tout point intérieur à une partie, mais pas sur les bords. En général, on distingue donc dans l'étude des extréma les points sur la frontière et ceux à l'intérieur.

***Exercices* :**

22) Déterminez les extréma de la fonction $(x, y) \mapsto x^3 + y^3 - 3xy$.
23) Même question avec $(x, y) \mapsto x^3 + x^2 + y^2$.
24) Même question avec $(x, y) \mapsto x^2 + x^2 y + y^3$.

6.3 Extréma locaux d'une fonction sur une partie A

La recherche des extréma locaux sur une partie quelconque est souvent un problème difficile. Néanmoins, on dispose de quelques résultats.

D'abord, on généralise le th. précédent (admis).

Proposition 17. *Soit U un ouvert de E, $f : U \to \mathbb{R}$ différentiable sur U, A une partie de U et $a \in A$.*
Si f possède un extrémum local en a sur A, alors $\mathrm{d}f(a)$ s'annule sur $T_a(A)$.

En conséquence, on a un résultat dans certains cas particuliers d'équipotentielles, appelé th. d'optimisation sous une contrainte.

Proposition 18. *Soit U un ouvert de E, $f, g : U \to \mathbb{R}$ de classe C^1.*
Si a est un élément de $A = g^{-1}(\{0\})$, si f possède un extrémum local en a sur A et si $\mathrm{d}g(a) \neq 0$, alors $\mathrm{d}f(a)$ est colinéaire à $\mathrm{d}g(a)$, ce qui revient à dire que les vecteurs gradients de f et g en a sont colinéaires.

Remarque. Là encore, il s'agit de conditions nécessaires, mais pas suffisantes en général. Une fois les points candidats trouvés, il faut toujours une étude locale pour les accepter ou non, ou bien utiliser un théorème d'existence comme le théorème des bornes atteintes.

***Exercices* :**

25) Déterminez les extréma de la fonction $(x, y) \mapsto xy$ sur la courbe d'équation $x^4 + y^4 = 1$.
26) Même question avec la fonction $(x, y) \mapsto x^3 + 2y^3$ sur la courbe d'équation $x^2 - y^2 = 1$.
27) Même question avec la fonction $(x, y, z) \mapsto x + y + z$ sur la surface d'équation $x^2 + y^2 - 2z^2 = 2$.

7 Fonctions de classe C^k

7.1 Dérivées partielles d'ordre supérieur

Pour une fonction quelconque définie sur un ouvert de $\mathbb{R}^p$, on peut éventuellement définir récursivement ses dérivées partielles d'ordre k comme étant les dérivées partielles des dérivées partielles d'ordre $k-1$ quand cela a un sens.

Évidemment, l'ordre dans lequel on effectue les dérivations a une importance.

Si $i = (i_1, \ldots, i_k) \in [\![1, p]\!]^k$ est un multi-indice, on note $\partial_i f$, $\partial_{i_1,\ldots,i_k} f$ ou $\dfrac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}$ la dérivée partielle $\partial_{i_1}(\partial_{i_2} \ldots \partial_{i_k} f)$.

L'ordre des dérivées partielles se lit donc de la droite vers la gauche.

A priori si une fonction a p variables, elle peut posséder jusqu'à p^k dérivées partielles d'ordre k.

7.2 Fonctions de classe C^k

La classe C^0 étant celle des fonctions continues, on peut donner une définition récursive de la classe C^k pour $k \in \mathbb{N}$.

Définition. Soit U un ouvert de $\mathbb{R}^p$, $f : U \to F$. Soit $k \in \mathbb{N}^*$.
On dit que f est de classe C^k sur U quand elle a des dérivées partielles en tout point de U et que celles-ci sont de classe C^{k-1} sur U.
On dit que f est de classe C^∞ sur U quand pour tout $k \in \mathbb{N}$, elle est de classe C^k.

Par application récursive des th. d'opérations précédents, on obtient les th. d'opérations sur les fonctions de classe C^k sans difficultés.

7.3 Th. de Schwarz

Dans le cas des fonctions de classe C^k, il y a finalement bien moins de dérivées partielles que prévu.

Théorème 3. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to F$.*
Si F est de classe C^2 sur U, alors pour tout $(i,j) \in [\![1,p]\!]^2$, $\partial_i(\partial_j f) = \partial_j(\partial_i f)$

On en déduit par récurrence que pour toute fonction de classe C^k, l'ordre des dérivations (jusqu'à k dérivations successives) n'importe pas.

8 Optimisation au second ordre

8.1 Hessienne

Définition. Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.
On appelle hessienne de f en a la matrice $H_f(a) = (h_{i,j}) \in \mathscr{M}_p(\mathbb{R})$ telle que pour tout $(i,j) \in [\![1,p]\!]^2$,

$$h_{i,j} = \partial_i \partial_j f(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

D'après le th. de Schwarz, la hessienne de f est alors une matrice symétrique.
On lui associe l'application $h_f(a)$ de $\mathbb{R}^p$ dans $\mathbb{R}$ définie de la façon suivante :
si $v = \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix}$ est un vecteur de $\mathbb{R}^p$, $h_f(a).(v,v) = v^\top H_f(a) v$.
C'est un produit scalaire (canonique) entre le vecteur-colonne v et $H_f(a)v$. Cette application est appelée la forme hessienne de f en a (à titre culturel, on appelle ce genre d'application des formes quadratiques).
En clair,

$$h_f(a).(v,v) = \sum_{1 \leqslant i,j \leqslant p} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j$$

8.2 Dév. limité à l'ordre 2

On admet le résultat suivant, appelé formule de Taylor-Young à l'ordre 2.

Proposition 19. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.*
Alors pour tout vecteur v au voisinage de 0,

$$f(a+v) = f(a) + \mathrm{d}f(a).v + h_f(a).(v,v) + o(\|v\|^2)$$

Ce d.l. est souvent écrit à l'aide du gradient et de la hessienne :

$$f(a+v) = f(a) + \nabla f(a) \cdot v + (H_f(a).v) \cdot v + o(\|v\|^2)$$

8.3 Application à l'étude des points critiques

Proposition 20. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.*
Si f possède un minimum local en a, alors a est un point critique et $H_f(a)$ est une matrice symétrique positive.

Remarque. Attention ! La réciproque est fausse. Néanmoins, elle est « presque vraie » en ajoutant une précision.

Proposition 21. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.*
Si a est un point critique et $H_f(a)$ est une matrice symétrique définie-positive, alors f possède un minimum local strict en a.

Cas particulier très courant : les applications à deux variables.
Soit f une fonction de classe C^2 sur un ouvert U de $\mathbb{R}^2$, $a \in U$.
Le gradient de f en a est le vecteur $\begin{pmatrix} \dfrac{\partial f}{\partial x}(a) \\ \dfrac{\partial f}{\partial y}(a) \end{pmatrix}$, la hessienne en a est la matrice $\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a) & \dfrac{\partial^2 f}{\partial x \partial y}(a) \\ \dfrac{\partial^2 f}{\partial x \partial y}(a) & \dfrac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}$
Cette matrice est définie-positive si et s.si sa trace et son déterminant sont strictement positifs.

En remplaçant f par $-f$, on en déduit des résultats similaires à propos des maxima locaux.

Proposition 22. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.*
Si f possède un maximum local en a, alors a est un point critique et $H_f(a)$ est une matrice symétrique négative (c'est-à-dire à valeurs propres toutes négatives).

Proposition 23. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.*
Si a est un point critique et $H_f(a)$ est une matrice symétrique définie-négative, alors f possède un maximum local strict en a.

Pour une application à deux variables, sa hessienne est définie-négative si et s.si sa trace est strictement négative et son déterminant strictement positif.

Proposition 24. *Soit U un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}$ de classe C^2, $a \in U$.*
Si a est un point critique et $H_f(a)$ est une matrice symétrique ayant deux valeurs propres non nulles de signes opposés, alors f ne possède pas d'extrémum en a (le point a est appelé point-col ou point-selle).

Exercices :

28) Déterminez les extréma de la fonction $(x, y) \mapsto 2y^2 + 2x^2 - x^4$.
29) Même question avec $(x, y) \mapsto y(x^2 + (\ln y)^2)$.
30) Soit $n \in \mathbb{N}^*$, $a_1, \dots, a_n$ n réels et $f(x, y) = \dfrac{1}{2x} \displaystyle\sum_{i=1}^{n} (y - a_i)^2 + \dfrac{n}{2} \ln x$. Déterminez (s'ils existent) les extréma de f.
31) Déterminez les extréma de la fonction $(x, y) \mapsto x^3 + y^2$.