Chapitre 15

Fonctions de plusieurs variables : calcul différentiel

Dans ce chapitre les fonctions , souvent notées f, sont définies sur un ouvert U d'un espace vectoriel normé réel de dimension finie E, à valeurs dans un espace vectoriel normé réel ou complexe de dimension finie F. Les résultats obtenus s'étendraient sans difficulté au cas où F serait un espace de Banach. Dans le cas où E serait un espace de Banach, il faudrait légèrement modifier la définition de la différentiabilité.

15.1 Application continûment différentiables

15.1.1 Fonctions différentiable, application linéaire tangente

Définition 15.1 *L'application f est différentiable en un point a de U si et seulement si il existe une application linéaire L telle que*

$$f(a+h) = f(a) + L(h) + \circ(h).$$

Si f est différentiable en a l'application L est unique et s'appelle l'application linéaire tangente à f en a, ou différentielle de f en a. On la notera $df(a)$.

Si f est différentiable en tout point de U on dit que f est différentiable sur U, et on définit sa différentielle df par $df : x \mapsto df(x)$.

Exemple : Une fonction de la variable réelle est différentiable si et seulement si elle est dérivable. Sa différentielle est alors l'application $h \mapsto df(a)(h) = hf'(a)$.

15.1.2 Cas des fonctions à valeur dans un espace produit

Si $f = (g, h)$ est à valeurs dans l'espace produit $F = G \times H$, alors f est différentiable en a si et seulement si g et h sont différentiables en a et $df(a) = (dg(a), dh(a))$. En particulier si $(v_1, \dots, v_n)$ est une base de F et si $f = \sum_{i=1}^{n} f_i v_i$, alors f est diférentiable en a si et seulement si chaque f_i est différentiable en a et $df(a)(h) = \sum_{i=1}^{n} df_i(a)(h)v_i$.

15.1.3 Dérivée selon un vecteur

Si a est un point de U et h un vecteur de E, il existe un nombre réel $\delta > 0$ tel que pour tout $|t| \leq \delta$ le point $a + th$ soit élément de U. On peut donc définir $\phi_h : [-\delta, \delta] \to F$ par $\phi_h(t) = f(a + th)$. Si ϕ_h possède une dérivée en 0, celle-ci s'appelle la dérivée de f en a suivant h. On la notera $D_h f(a)$.

Proposition 15.1 *Si f est différentiable en a, alors f admet une dérivée en a selon tout vecteur h, égale à $df(a)(h)$.*

Mais la réciproque est fausse. On peut par exemple envisager le contre-exemple de l'application f définie par $f(x, y) = y$ si $x \neq 0$, 0 sinon, étudiée en 0.

15.1.4 Dérivées partielles

Si $(e_1, \dots, e_n)$ est une base de E les dérivées de f suivant les vecteurs de la base s'appellent les dérivées partielles dans la base E. On notera :

$$D_j f(a) \text{ ou } \frac{\partial f}{\partial x_j}(a) \text{ la valeur de } df(a)(e_j) = D_{e_j} f(a).$$

Dans la plupart des cas E sera $\mathbb{R}^n$, et la base sera la base canonique, sans que cela soit précisé.

Dans la deuxième écriture l'usage de la lettre x_j sert à nous indiquer l'indice de dérivation. Par exemple usuellement, si $E = \mathbb{R}^3$ la dérivée partielle $\frac{\partial f}{\partial y}(a)$ est $D_2 f(a)$. Noter aussi la position du (a), sur la même ligne que la barre de fraction, ce qui permettra ultérieurement de différencier les notations

$$\frac{\partial f}{\partial y}(xy, x+y)$$

et

$$\frac{\partial f(xy, x+y)}{\partial y} = x\frac{\partial f}{\partial x}(xy, x+y) + \frac{\partial f}{\partial y}(xy, x+y).$$

Proposition 15.2 *Si f est différentiable en a elle admet des dérivées partielles dans toute base, et de plus :*

$$df(a)(h) = \sum_{i=1}^{n} h_i D_i f(a).$$

Si pour tout a de U, $D_j f(a)$ existe, on définit la dérivée partielle de f d'indice j par $D_j : x \mapsto D_j(x)$, notée aussi $\dfrac{\partial f}{\partial x_j}$.

Définition 15.2 *Une application f est continuement différentiable si et seulement si pour une base les dérivées partielles dans cette base sont définies et continues sur U*

Cette définition, qui est celle du programme, n'est pas bonne. On rappelle que l'existence pour tout h et tout x de $D_h f(x)$ n'implique pas la différentiablité de f même en un unique point. Il eut été plus naturel (et c'est ce qui se fait en général) prendre pour définition d'une application continûment différentiable la propriété d'être différentiable et de posséder une différentielle continue.

Fort heureusement les deux définitions sont équivalentes, et équivalentes à la troisième caractérisation suivante, qui est celle que vous devez retenir :

Théorème 15.1 *f est continûment différentiable sur E si et seulement pour une base de E toute les dérivées partielles $D_j f$ existent et sont continues sur U. Cette propriété est alors conservée pour toute base.*

Il est facile de prouver que si f est différentiable alors elle est continûment différentiable si et seulement si ses dérivées partielles sont continues. Le théorème précédent est bien plus fort, il affirme que si les dérivées partielles existent et sont continues, alors f est différentiable, puis de facto de classe C^1.

Corollaire 15.1 *L'ensemble des applications de classe C^1 sur un ouvert U de E, à valeurs dans F est un sous-espace vectoriel de $\mathcal{A}(U, F)$, noté $\mathcal{C}^1(U, F)$.*

On vérifierait facilement que l'application qui à une fonction associe sa différentielle, ou une dérivée partielle est linéaire.

15.1.5 Cas d'une application linéaire, d'une application bilinéaire

Toute application linéaire $f$est différentiable. De plus sa différentielle est constante, pour tout a $df(a) = f$. En particulier f est de classe C^1. Plus généralement toute application affine est de classe C^1 et sa différentielle est constante, sa valeur étant l'application linéaire associée.

Toute application bilinéaire B est de classe C^1, et sa différentielle est définie par

$$dB(a, b)(h, k) = B(a, k) + B(h, b).$$

15.1.6 Composée de deux applications de classe C^1

Proposition 15.3 *La composée de deux fonctions de classe C^1, si elle est définie, est de classe C^1. Plus précisément $d(g \circ f)(a) = dg(f(a)) \circ df(a)$ pour tout a.*

On en déduit la formule fondamentale suivante, clé de voute du calcul sur les dérivée partielle. On se donne f une application de U ouvert de E vers F, g une application de l'ouvert V de F vers G. On suppose f et g différentiables. On fixe des bases dans E, F et G, ce qui permet de parler des dérivées partielles. On peut aussi introduire les fonctions coordonnées $(f_1, \ldots, f_n)$ de f. Avec ces hypothèses, on a, pour tout i :

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

Certain emploient la notation $\dfrac{\partial g}{\partial f_k}$ au lieu de $\dfrac{\partial g}{\partial y_k}$.

Exemple : Dérivées partielles de $g \circ f$ si $f : \mathbb{R} \to \mathbb{R}^n$ et $g : \mathbb{R}^n \to \mathbb{R}$, ou bien si $f : \mathbb{R}^n \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$

Application : Différentielle de $B(f, g)$ si f et g sont de classe C^1

Corollaire 15.2 *L'ensemble des application de classe C^1 sur U à valeurs dans $\mathbb{K}$ est une sous–algèbre de $\mathcal{A}(U, \mathbb{K})$, notée $\mathcal{C}^1(U)$. De plus si f ne s'annule pas sur U, alors $\frac{1}{f}$ est différentiable sur U.*

Application : Toute les fonctions dont une expression dans des bases fixée s'exprime à l'aide des fractions rationnelles est de classe C^1 sur son domaine de définition. Par exemple l'application

$$M \mapsto M^{-1}$$

est de classe C^1 sur $GL_n(\mathbb{K})$. On peut montrer que sa différentielle en M est l'application

$$H \mapsto -M^{-1}HM^{-1}.$$

15.1.7 Matrice jacobienne, jacobien

On fixe deux bases de E et F, la matrice jacobienne de f en a est la matrice de $df(a)$ dans ces bases. On parle aussi de matrice jacobienne de f, tout simplement, pour l'application qui à a associe la matrice jacobienne de f en a.

Si $E = F$, le déterminant de la matrice jacobienne de f s'appelle le jacobien de f.

Proposition 15.4 *La matrice jacobienne d'une composée est le produit des matrices jacobiennes*

15.2 Fonctions numériques continûment différentiables

15.2.1 Gradient

On suppose que l'espace E est euclidien.

15.2.2 Inégalité des accroissements finis

On se place sur un ouvert convexe.

Proposition 15.5 *Soit f une fonction numérique définie sur un ouvert connexe U d'un ouvert d'un espace vectoriel normé E. On suppose que f est de classe C^1 sur U et qu'il existe M dans $\mathbb{R}$ telle que $\forall a \in U$: $\|df(a)\| \leq M$. Alors pour tout couple (x, y) de points de U on a*

$$|f(x) - f(y)| \leq M\|x - y\|.$$

Dans le cas où E est un espace euclidien on peut remplacer $\|df(a)\|$ par $\nabla f(a)$. On peut aussi adapter la démonstration pour obtenir une majoration plus fine si on sait majorer $df(tx + (1-t)y)(x-y)$ par une fonction $\phi_{x,y}(t)$ que l'on sait intégrer entre 0 et 1.

15.2.2.1 Application : caractérisation des applications constantes.

Corollaire 15.3 *Une application, définie sur un ouvert connexe est constante si et seulement si elle est de classe C^1 et sa différentielle est nulle.*

On pourrait, comme dans le cas d'une fonction d'une variable réelle, caractériser les fonctions lipschitziennes.

15.2.3 Point critiques

15.2.3.1 Points critiques ou stationnaires d'une fonction numérique

Définition 15.3 *Un point critique d'une fonction numérique différentiable est un point où la différentielle de cette fonction s'annule.*

15.2.3.2 Extremums locaux

Une fonction numérique définie sur une partie non vide d'un espace vectoriel normé E, atteint en un point a de X un maximum local si et seulement si il existe un voisinage V de a tel que pour tout x de V on ait $f(x) \leq f(a)$. Si pour tout x de V distinct de a on a $f(x) < a$ on parle de maximum local strict. On définirait de même les notions de minimum local et de mimimum local strict. Si on est dans une de ces situations, on dit que f atteint en a un extremum strict. Si $f(x) \leq f(a)$ (resp. $f(x) < a$, $f(x) \geq a$, $f(x) > a$) pour tout x de X, f atteint en a un maximum global (resp. maximum global strict, etc...).

15.2.3.3 Caractérisation par une condition nécessaire des extremums locaux

Proposition 15.6 *Si une fonction numérique f définie sur un ouvert U admet un extremum local en a alors a est un point critique de f.*

Cette proposition souffre de deux faiblesses pour être mise directement en application pour trouver les extremums d'une fonction numérique. Elle est nécessaire mais non suffisante (le point 0 est un point critque de $x \mapsto x^3$, où la fonction n'atteint pas d'extremum local) et elle ne s'applique que sur des ouverts, alors que le théorème qui nous permet d'affirmer l'existence d'extremums est la continuité sur un compact.

15.3 Dérivée partielle d'ordre $k \geq 2$

15.3.1 Définition

Si f est une fonction définie sur un ouvert U admettant une dérivée partielle $g = \dfrac{\partial f}{\partial x_j}$ sur U, et cette dérivée partielle admet une dérivée partielle $\dfrac{\partial g}{\partial x_i}$, on dit que f admet une dérivée partielle d'ordre 2, notée $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

Si f est différentiable et si sa différentielle est aussi différentiable (on dit alors que f est deux fois différentiable) alors f admet des dérivées partielles par rapport à tout couple d'indice. Mais la réciproque est fausse.

La différentielle seconde d'une application numérique peut être identifiée à une forme bilinéaire.

Une application est de classe C^2 si elle est deux fois différentiable si elle est deux fois différentiable et si sa différentielle seconde est continue.

La caractérisation suivante peut tenir lieu de définition.

Définition 15.4 *Une fonction f définie sur un ouvert U est de classe C^2 si et seulemnt si elle admet des dérivées partielles secondes par rapport à tout couple de variables et ces dérivées partielles sont continues.*

15.3.2 Théorème de Schwarz

Si f est de classe C^2 sur U, alors pour tout couple de variables on a :

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

La différentielle seconde d'une fonction numérique de classe C^2 est donc un forme bilinéaire symétrique.

On peut généraliser cette notion de dérivée partielle d'ordre 2 à la notion de dérivée partielle d'ordre k. Une fonction sera de classe C^k si elle admet des dérivées partielles d'ordre k par rapport à un k-uplet quelconque de variables et si celles-ci sont continues.

Le théorème de Schwarz se généralise pour affirmer que le calcul d'une la dérivée partielle d'ordre k, d'une fonction de classe C^k, ne dépend pas de l'ordre dans lequel sont calculées les dérivées partielles successives.

Les formules donnant les dérivées partielles des sommes, produit externe ou interne et composée de fonctions de classe C^1, nous permettent de démontrer par récurrence la stabilité de l'ensemble des applications de classe C^k pour ces opérations.

En particulier : $C^k(U,F)$, ensemble des applications de classe C^k de U vers F, est un sous espace vectoriel de $A(U,F)$, et $C^k(U,\mathbb{K}) = C^k(U)$ une sous-algèbre de $A(U,K)$.

Un exemple : L'application D_h, dérivée suivant un vecteur fixé, est une application linéaire de $C^k(U,F)$ vers $C^{k-1}(U,F)$, en effet $D_h f = \sum_i h_i \dfrac{\partial f}{\partial x_i}$.

Une application qui est de classe C^k pour tout k est dite de classe C^∞. (Notations $C^\infty(U,F)$ et $C^\infty(U)$).

15.4 Coordonnées polaires

On se place dans le plan affine euclidien $\mathbb{R}^2$, rapporté au repère $(O,\overrightarrow{e_1},\overrightarrow{e_2})$(ou plus génralement dans tout plan affine euclidien dont $(O,\overrightarrow{e_1},\overrightarrow{e_2})$ est un repère orthonormé). Si θ est un réel on définit deux vecteurs

$$\left\{\begin{array}{rcrcrcrc} \overrightarrow{u}(\theta) & = & \cos\theta & \overrightarrow{e_1} & + & \sin\theta & \overrightarrow{e_2} \\ \overrightarrow{v}(\theta) & = & -\sin\theta & \overrightarrow{e_1} & + & \cos\theta & \overrightarrow{e_2} \end{array}\right. .$$

Le repère $(O,\overrightarrow{u}(\theta),\overrightarrow{v}(\theta))$, plus simplement noté $(O,\overrightarrow{u},\overrightarrow{v})$, est un repère orthonormé que l'on appelle le repère polaire (d'angle θ).

On a facilement

$$\frac{d\overrightarrow{u}(\theta)}{d\theta} = \overrightarrow{v}(\theta), \quad \frac{d\overrightarrow{v}(\theta)}{d\theta} = -\overrightarrow{u}(\theta).$$

Pour tout point M de $\mathbb{R}^2$ il existe un couple (ρ,θ), dans $\mathbb{R}^+ \times \mathbb{R}$ tel que

$$\overrightarrow{OM} = \rho\overrightarrow{u}(\theta)$$

, un tel couple s'appelle des cordonnées polaires de M. $\rho = OM$ est uniquement déterminé, si M est distinct de O, θ est déterminé à un multiple de 2π près, si $M = O$, alors toute valeur de θ convient.

On peut montrer que l'application $(\rho,\theta) \mapsto (\rho\cos\theta, \rho\sin\theta)$ est un difféomorphisme de $\mathbb{R}^{*+}\times]-\pi,\pi[$ sur $\mathbb{R}^2$ privé de la demi-droite $\mathbb{R}^- \times \{0\}$.

15.4.0.1 Application : calcul du gradient en coordonnées polaires.

On pose $F(\rho,\theta) = f(\rho\cos\theta,\rho\sin\theta)$, avec cette notation, si f est de classe C^1 :

$$\operatorname{grad} f(\rho\cos\theta,\rho\sin\theta) = \frac{\partial F}{\partial\rho}\overrightarrow{u} + \frac{1}{r}\frac{\partial F}{\partial\theta}\overrightarrow{v}.$$