Fonctions de plusieurs variables

November 1, 2004

1 Différentiabilité

1.1 Motivation

Pour une fonction d'une variable f, définie au voisinage de 0, être dérivable en 0, c'est admettre un développement limité à l'ordre 1,

$$f(x) = b + ax + x\epsilon(x).$$

Alors $b = f(0)$ et $a = f'(0)$.

Interprétation géométrique. La courbe représentative de f possède en $(0, a)$ une tangente, la droite d'équation $y = b + ax$.

On veut faire pareil pour une fonction de deux variables. La courbe représentative est remplacée par une surface représentative d'équation $z = f(x, y)$, la droite tangente par un plan tangent d'équation $z = c + ax + by$. La tangence s'exprime en disant que la distance entre le point $(x, y, f(x, y))$ de la surface et le point $(x, y, c + ax + by)$ du plan est petite devant la distance de (x, y) à l'origine.

Exemple 1.1 $f(x, y) = x^2 + y^2$.

1.2 Différentiabilité d'une fonction de deux variables

Définition 1.2 *Soit f une fonction de deux variables, définie au voisinage de $(0, 0)$. On dit que f est* différentiable *en $(0, 0)$ si elle admet un* développement limité à l'ordre 1*, i.e. si on peut écrire*

$$f(x, y) = c + ax + by + \sqrt{x^2 + y^2}\,\epsilon(x, y),$$

où $\epsilon(x, y)$ tend vers 0 lorsque x et y tendent vers 0.

Dans ce cas, f admet des dérivées partielles en $(0, 0)$, et

$$c = f(0, 0), \quad a = \frac{\partial f}{\partial x}(0, 0), \quad \frac{\partial f}{\partial y}(0, 0).$$

La différentiabilité de f en un point quelconque (x_0, y_0) se traduit par le développement limité

$$f(x_0 + u, y_0 + v) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)u + \frac{\partial f}{\partial y}(x_0, y_0)v + \sqrt{u^2 + v^2}\,\epsilon(u, v),$$

où $\epsilon(u, v)$ tend vers 0 lorsque u et v tendent vers 0.

Exemple 1.3 $f(x, y) = x(2 - x + y) + y(1 - x - y)$ *est différentiable à l'origine.*

En effet,

$$\begin{aligned} f(x, y) &= 2x + y - x^2 - y^2 \\ &= 2x + y + \sqrt{x^2 + y^2}\,\epsilon(x, y), \end{aligned}$$

où

$$\epsilon(x,y) = -\sqrt{x^2+y^2}$$

tend vers 0 quand x et y tendent vers 0.

Théorème 1 *Soit f une fonction de deux variables définie au voisinage de $(0,0)$. Si les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont définies au voisinage de $(0,0)$ et continues en $(0,0)$, alors f est différentiable en $(0,0)$, et son développement limité à l'ordre 1 s'écrit*

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

Exemple 1.4 *$f(x,y) = x(2-x+y) + y(1-x-y)$ est différentiable en tout point.*

En effet, on n'a qu'a utiliser le théorème 1. On peut aussi calculer directement

$$\begin{aligned} f(x_0+u, y_0+v) &= 2x_0 + 2u + y_0 + v - x_0^2 - 2x_0u - u^2 - y_0^2 - 2y_0v - v^2 \\ &= 2x_0 + y_0 - x_0^2 - y_0^2 + (2-2x_0)u + (1-2y_0)v - u^2 - v^2 \\ &= 2x_0 + y_0 - x_0^2 - y_0^2 + (2-2x_0)u + (1-2y_0)v + \sqrt{u^2+v^2}\epsilon(u,v). \end{aligned}$$

1.3 Gradient

Définition 1.5 *Soit f une fonction de deux variables, différentiable tout point d'un domaine D. Son* gradient *est le champ de vecteurs défini sur D par*

$$\nabla f : (x,y) \mapsto \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{pmatrix}.$$

Exemple 1.6 *Le gradient de la fonction définie sur $\mathbf{R}^2$ par $f(x,y) = x^2$ est le champ de vecteurs horizontal $\nabla_{(x,y)} f = \begin{pmatrix} 2x \\ 0 \end{pmatrix}$.*

1.4 Interprétation du développement limité

Proposition 1.7 *Si f est différentiable en P, alors pour toute droite $t \mapsto P + tv$ passant par P, la fonction $t \mapsto f(P+tv)$ est dérivable, et*

$$\frac{d}{dt} f(P+tv)_{|t=0} = \nabla_P f \cdot v.$$

On verra plus loin (théorème 2) que cette formule est vraie pour toute courbe, et non seulement les droites, sous la forme

$$\frac{d}{dt} f(c(t)) = \nabla_{c(t)} f \cdot c'(t).$$

1.5 Lignes de niveau

Définition 1.8 *On appelle* lignes de niveau *de f les ensembles de la forme $L_w = \{(x,y)\,;\, f(x,y) = w\}$.*

Exemple 1.9 *Les lignes de niveau de la fonction $f(x,y) = x^2+y^2$ sont des cercles concentriques. Celles de la fonction $f(x,y) = xy$ sont des hyperboles, à l'exception de la ligne de niveau 0, qui est la réunion de deux droites.*

Proposition 1.10 *Le gradient d'une fonction est un vecteur perpendiculaire aux lignes de niveau, pointant dans la direction dans laquelle la fonction augmente. Sa longueur est d'autant plus grande que la fonction varie rapidement, i.e. que les lignes de niveau sont rapprochées. Le gradient indique la direction de plus grande pente.*

Preuve. Soit $t \mapsto c(t)$ une ligne de niveau. Alors $t \mapsto f(c(t))$ est constante, donc

$$0 = \frac{d}{dt} f(c(t)) = \nabla_{c(t)} f \cdot c'(t),$$

ce qui montre que le gradient est orthogonal à la tangente à la ligne de niveau.

Lorsque l'on se déplace dans la direction du gradient, par exemple, par $t \mapsto c(t) = P + t\nabla_P f$,

$$\frac{d}{dt} f(c(t))_{|t=0} = \nabla_P f \cdot c'(0) = \| \nabla_P f \|^2 > 0,$$

donc f augmente, d'autant plus vite que $\| \nabla_P f \|$ est grand.

Soit v un vecteur unitaire. Alors

$$\frac{d}{dt} f(P + tv)_{|t=0} = \nabla_P f \cdot v$$

est maximum lorsque v est colinéaire et de même sens que $\nabla_P f$, donc $\nabla_P f$ indique la direction de plus grande pente. ∎

1.6 Généralisation

De la même façon, on peut parler de développement limité et de différentiabilité pour une fonction de n variables (remplacer $\sqrt{x^2 + y^2}$ par $\sqrt{x_1^2 + \cdots + x_n^2}$), puis pour une application $\mathbf{R}^n \to \mathbf{R}^p$. Dans ce cas, les coefficients du développement limité sont des vecteurs de $\mathbf{R}^p$.

Exemple 1.11 *Soit I un intervalle de $\mathbf{R}$ et $c : I \to \mathbf{R}^2$ une courbe. Calculer un développement limité de c en 0, c'est calculer des développements limités des fonctions coordonnées $x(t) = a_0 + a_1 t + t\epsilon(t)$, $y(t) = b_0 + b_1 t + t\epsilon(t)$, et former le développement limité vectoriel*

$$c(t) = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + t \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + t\epsilon(t).$$

Proposition 1.12 *Une application $F = (f_1, \ldots, f_p) : \mathbf{R}^n \to \mathbf{R}^p$ est différentiable si et seulement si chacune de ses composantes l'est.*

1.7 La différentielle

Définition 1.13 *Soit $F := (f_1, \ldots, f_p) : \mathbf{R}^n \to \mathbf{R}^p$ une application différentiable en P. Sa* différentielle *en P est l'application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^p$ qui apparaît comme le terme non constant du développement limité à l'ordre 1 en P. Sa matrice, appelée* matrice jacobienne, *a pour coefficients les dérivées partielles,*

$$J_f(P) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{pmatrix}.$$

Exemple 1.14 *Si A est une matrice, alors l'application linéaire $f_A : \mathbf{R}^n \to \mathbf{R}^p$ qu'elle définit est différentiable, et sa matrice jacobienne est A en n'importe quel point.*

Exemple 1.15 *Soit $f(x, y) = 2x + y - x^2 - y^2$. Sa matrice jacobienne est*

$$\begin{pmatrix} 2 - 2x & 1 - 2y \end{pmatrix}.$$

Autrement dit, la matrice jacobienne d'une fonction, c'est son gradient vu comme un vecteur ligne.

Exemple 1.16 *Soit* $F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$. *Sa matrice jacobienne est* $\begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}$.

Autrement dit, la matrice jacobienne d'une courbe, c'est sa dérivée vue comme un vecteur colonne.

Exemple 1.17 *Soit* $F(r, \theta) = (r\cos(\theta), r\sin(\theta))$. *Sa matrice jacobienne est*

$$\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

1.8 Matrice jacobienne d'une fonction composée

Il s'agit de généraliser la formule

$$(g \circ f)' = (g' \circ f)f'.$$

Théorème 2 *Soient* $f : \mathbf{R}^n \to \mathbf{R}^p$ *et* $g : \mathbf{R}^p \to \mathbf{R}^q$ *des applications. On suppose* f *différentiable en* P *et* g *différentiable en* $f(P)$. *Alors* $g \circ f$ *est différentiable en* P, *et*

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

Preuve. Si $v \in \mathbf{R}^n$,

$$f(P + v) = f(P) + J_f(P)v + \parallel v \parallel \epsilon(v).$$

On pose $w = f(P + v) - f(v)$. Alors

$$g(f(P) + w) = g(f(P)) + J_g(f(P))w + \parallel w \parallel \epsilon(w).$$

Autrement dit,

$$\begin{aligned} g \circ f(P + v) &= g \circ f(P) + J_g(f(P))(J_f(P)v + \parallel v \parallel \epsilon(v)) + \parallel w \parallel \epsilon(w) \\ &= g \circ f(P) + J_g(f(P))J_f(P)v + \parallel v \parallel \epsilon(v), \end{aligned}$$

car $\parallel w \parallel / \parallel v \parallel$ est borné. ∎

Corollaire 1.18 *Soit* I *un intervalle de* $\mathbf{R}$, *soit* $c : I \to \mathbf{R}^2$ *une courbe dans le plan. Soit* $f : \mathbf{R}^2 \to \mathbf{R}$ *une fonction sur le plan. Alors*

$$(f \circ c)'(t) = J_g c'(t) = \nabla_{c(t)} f \cdot c'(t) = \frac{\partial f}{\partial x}(c(t))x'(t) + \frac{\partial f}{\partial y}(c(t))y'(t).$$

Corollaire 1.19 *Soit* $f : \mathbf{R}^2 \to \mathbf{R}$ *une fonction sur le plan. Soit* $g : \mathbf{R} \to \mathbf{R}$ *une fonction d'une variable. Alors*

$$J_{g\circ f} = g'(f)J_f, \quad i.e. \quad \nabla_P g \circ f = g'(f(P))\nabla_P f.$$

Corollaire 1.20 *Soit* $F : \mathbf{R}^2 \to \mathbf{R}^2$, $F(r, \theta) = (r\cos(\theta), r\sin(\theta))$, *le changement de coordonnées polaires. Soit* $c : \mathbf{R} \to \mathbf{R}^2$ *une courbe paramétrée, vue en coordonnées cartésiennes* $(x(t), y(t))$ *ou polaires* $(r(t), \theta(t))$. *Alors la vitesse en coordonnées cartésiennes s'obtient en appliquant la matrice jacobienne de* F *à la dérivée des coordonnées polaires,*

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \begin{pmatrix} r' \\ \theta' \end{pmatrix} = r'e_r + \theta' r e_\theta.$$

1.9 Condition d'extremum

Proposition 1.21 *Soit f une fonction à valeurs réelles définie au voisinage d'un point P de $\mathbf{R}^n$. Si P est un minimum local (resp. maximum local) de f, alors le gradient de f s'annule en P.*

Preuve. Cas $n = 2$. Soit $P = (x_0, y_0)$. A fortiori, x_0 est un minimum local (resp. maximum local) de la fonction $x \mapsto f(x, y_0)$, donc sa dérivée en x_0 est nulle. Or celle-ci vaut $\frac{\partial f}{\partial x}(P)$. De même, $\frac{\partial f}{\partial x}(P) = 0$, donc $\nabla_P f = 0$. ■

Remarque 1.22 *En général, la réciproque est fausse.*

On peut donner des conditions suivantes plus fortes, faisant intervenir les dérivées secondes. C'est l'objet du paragraphe suivant.

2 Développement limité à l'ordre 2

2.1 Motivation

On s'intéresse au mouvement dans un champ de forces dérivant d'un potentiel V. Les positions d'équilibre correspondent aux points où les dérivées partielles de V s'annulent. Pour qu'une position d'équilibre P soit *stable*, il vaut mieux que V possède un *minimum local strict* en P, i.e., que pour $v \neq 0$ assez petit, $V(P + v) > V(P)$.

Soit f une fonction d'une variable. Supposons que f admet un minimum en 0. Alors sa dérivée $f'(0)$ s'annule. La réciproque n'est pas vraie : la fonction définie sur $\mathbf{R}$ par $f(x) = x^3$ a une dérivée nulle en 0 mais n'admet pas de minimum local. Une condition suffisante fait intervenir la dérivée seconde.

Proposition 2.1 *Soit f une fonction d'une variable. Supposons que $f'(0) = 0$ et $f''(0) > 0$. Alors f possède un minimum local strict en 0 : pour $x \neq 0$ suffisamment petit, $f(x) > f(0)$.*

Preuve. Le développement limité de Taylor-Young donne

$$f(x) = f(0) + \frac{1}{2}f''(0)x^2 + x^2\epsilon(x).$$

Alors

$$\frac{f(x) - f(0)}{x^2} = \frac{1}{2}f''(0) + \epsilon(x) > 0$$

pour x assez petit. ■

On peut aussi parler de développement limité à l'ordre 2 pour une fonction de plusieurs variables. C'est lié aux dérivées partielles secondes, cela donne un condition suffisante pour un minimum local strict.

2.2 Définition

Proposition 2.2 *Soit $m(x, y) = ax^r y^s$ un polynôme de degré $r+s$. Alors on peut écrire $m(x, y) = (\sqrt{x^2 + y^2})^{r+s-1}\epsilon(x, y)$ où $\epsilon(x, y)$ tend vers 0 quand x et y tendent vers 0*

Autrement dit, dès que $r+s \geq 2$, un monôme $ax^r y^s$ peut être mis dans le reste d'un développement limité à l'ordre 1. Il ne reste donc dans le développement limité à l'ordre 1 d'une fonction f que des termes de degré 0 (le terme constant $f(0, 0)$) et 1 (la différentielle de f en $(0, 0)$).

On va voir que les monômes $ax^r y^s$ tels que $r + s \geq 3$, peuvent être mis dans les restes des développements limités à l'ordre 2. Ceux-ci ne comportent donc que des termes de degrés 0, 1 et 2. Les termes de degré 2 sont de la forme $px^2 + rxy + sy^2$, où p, q et r sont des constantes. Cela motive la définition suivante.

Définition 2.3 *Soit f une fonction de deux variables définie au voisinage de 0. On dit que f admet un* développement limité à l'ordre 2 *en* $(0,0)$ *si on peut écrire*

$$f(x,y) = c + ax + by + px^2 + qxy + ry^2 + (x^2+y^2)\epsilon(x,y),$$

où $\epsilon(x,y)$ tend vers 0 lorsque x et y tendent vers 0.

Plus généralement, on dit que f admet un développement limité à l'ordre 2 en (x_0,y_0) si on peut écrire

$$f(x_0+u, y_0+v) = c + au + bv + pu^2 + quv + rv^2 + (u^2+v^2)\epsilon(u,v),$$

où $\epsilon(u,v)$ tend vers 0 lorsque u et v tendent vers 0.

Théorème 3 (Développement limité de Taylor-Young). *Soit f une fonction de deux variables définie au voisinage de 0. On suppose que f admet des dérivées partielles secondes $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y^2}$, et que celles-ci sont continues au voisinage de 0. Alors f admet un développement limité à l'ordre 2,*

$$\begin{aligned} f(x,y) &= f(0,0) + \frac{\partial f}{\partial x}(0,0)x + \frac{\partial f}{\partial y}(0,0)y + \frac{1}{2}(\frac{\partial^2 f}{\partial x^2}(0,0)x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0,0)xy + \frac{\partial^2 f}{\partial y^2}(0,0)y^2) \\ &\quad +(x^2+y^2)\epsilon(x,y). \end{aligned}$$

Autrement dit, la plupart des fonctions qu'on rencontrera admetteront un développement limité.

Exemple 2.4 $f(x,y) = -\cos(x)\cos(y)$ *admet en* $(0,0)$ *le développement limité*

$$\begin{aligned} f(x,y) &= -(1 - \frac{1}{2}x^2 + x^2\epsilon(x))(1 - \frac{1}{2}y^2 + y^2\epsilon(y)) \\ &= -1 + \frac{1}{2}x^2 + \frac{1}{2}y^2 + (x^2+y^2)\epsilon(x,y)). \end{aligned}$$

En $(\frac{\pi}{2}, \frac{\pi}{2})$, elle admet le développement limité

$$\begin{aligned} f(\frac{\pi}{2}+u, \frac{\pi}{2}+v) &= -\sin(u)\sin(v) \\ &= -(u+u^2\epsilon(u))(v+v^2\epsilon(v)) \\ &= -uv + (u^2+v^2)\epsilon(u,v). \end{aligned}$$

Dans les deux cas, on reconnaît les dérivées partielles secondes dans les coefficients de u^2, uv et v^2.

2.3 Signe

Pour une fonction d'une variable de la forme px^2, le signe ne dépend que du signe de p. Pour une fonction de deux variables de la forme $px^2 + qxy + ry^2$, l'étude du signe se ramène à celui du trinôme du second degré $Z \mapsto pZ^2 + qZ + r$. En effet, si on pose $Z = x/y$,

$$px^2 + qxy + ry^2 = x^2(pZ^2 + qZ + r).$$

Par conséquent,

Proposition 2.5

- *Si $q^2 - 4pr < 0$ et $p > 0$, alors pour tout $(x,y) \neq (0,0)$, $px^2 + rxy + sy^2 > 0$.*
- *Si $q^2 - 4pr = 0$, $p \geq 0$ et $r \geq 0$, alors pour tout (x,y), $px^2 + qxy + ry^2 \geq 0$.*
- *Si $q^2 - 4pr > 0$, la fonction $px^2 + qxy + ry^2$ prend les deux signes au voisinage de 0.*

Théorème 4 *Soit f une fonction de deux variables définie au voisinage de 0. On suppose que f admet un développement limité à l'ordre 2 au voisinage de $(0,0)$, de la forme*

$$f(x,y) = c + ax + by + px^2 + qxy + ry^2 + (x^2+y^2)\epsilon(x,y).$$

- *Si* $(0,0)$ *est un minimum local pour* f*, alors* $a=b=0$*,* $q^2-4pr\leq 0$*,* $p\geq 0$ *et* $s\geq 0$*.*
- *De même, si* $(0,0)$ *est un maximum local pour* f*, alors* $a=b=0$*,* $q^2-4pr\leq 0$*,* $p\leq 0$ *et* $s\leq 0$*.*
- *Réciproquement, si* $a=b=0$*,* $q^2-4pr<0$ *et* $p>0$*, alors* $(0,0)$ *est un minimum local pour* f*.*
- *De même, si* $a=b=0$*,* $q^2-4pr<0$ *et* $p<0$*, alors* $(0,0)$ *est un maximum local pour* f*.*

Exemple 2.6 *La fonction* $f(x,y)=-\cos(x)\cos(y)$ *de l'exemple 2.4 admet en* $(0,0)$ *un minimum local strict. En revanche, en* $(\frac{\pi}{2},\frac{\pi}{2})$*, il ne s'agit pas d'un minimum local. Si on interprète* f *comme le relief d'une table bosselée, une bille qui roule sur la table s'arrêtera dans un creux (par exemple, en* $(0,0)$*), mais pas dans un col comme* $(\frac{\pi}{2},\frac{\pi}{2})$*.*

Exemple 2.7 *On s'intéresse aux boîtes en forme de parallépipède. On cherche, parmi les boîtes de contenance donnée 1, à minimiser l'aire. Montrer que l'aire atteint un minimum local pour la boîte cubique.*

Notons x et y les longueurs de deux des côtés. Si la contenance vaut 1, alors la hauteur vaut $z=\frac{1}{xy}$. L'aire de la boîte, somme des aires des 6 faces, vaut

$$f(x,y)=2xy+2yz+2zx=2xy+\frac{2}{x}+\frac{2}{y}.$$

La boîte cubique correspond à $x=y=1$. On applique le théorème 3 ou on développe

$$\begin{aligned}
f(1+u,1+v) &= 2(1+u)(1+v)+\frac{2}{1+u}+\frac{2}{1+v} \\
&= 2+2u+2v+2uv+2-2u+2u^2+2-2v+2v^2+u^2\epsilon(u)+v^2\epsilon(v) \\
&= 6+2u^2+2uv+2v^2+(u^2+v^2)\epsilon(u,v).
\end{aligned}$$

Le discriminant $q^2-4pr=-12<0$, donc le critère 4 s'applique, et la boîte cubique est bien un minimum local de l'aire.