Chapitre 2

Fonctions différentiables

Sommaire

2.1 Introduction

Soit $f :]a,b[\longrightarrow \mathbb{R}$ dérivable au point $x_0 \in]a,b[$. On a par définition

$$f'(x_0) = \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(x_0+h) - f(x_0)}{h}.$$

Soit ε la fonction définie au voisinage de 0 par $\varepsilon(h) = \dfrac{f(x_0+h) - f(x_0)}{h} - f'(x_0)$. On alors

$$f(x_0+h) - f(x_0) - hf'(x_0) = h\varepsilon(h) \text{ avec } \lim_{h\to 0} \varepsilon(h) = 0.$$

La dérivabilité de la fonction f en x_0 revient à l'existence d'un nombre $f'(x_0)$ et une fonction ε définie au voisinage de 0 vérifiant $\lim_{h\to 0} \varepsilon(h) = 0$.

Lorsque h est assez petit, on peut écrire $f(x_0+h) \approx f(x_0) + hf'(x_0)$. La notion de dérivée en un point permet d'approximer les valeurs de la fonction en des points proches de celui-là.

Nous allons voir comment ceci se généralise au cas d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$ moyennant l'introduction de la notion de différentielle.

2.2 Définitions et propriétés élémentaires

Dans toute la suite de ce chapitre U désignera un ouvert non vide de $\mathbb{R}^n$.

Définition 53

Soit f une fonction de U dans $\mathbb{R}$ et a un point de U. On dit que f est différentiable en a s'il existe une forme linéaire u_a sur $\mathbb{R}^n$ (i.e. une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$) telle que

$$\lim_{\substack{h\to 0\\ h\neq 0, a+h\in U}} \frac{f(a+h)-f(a)-u_a(h)}{\|h\|} = 0,$$

où $\|.\|$ est une norme quelconque sur $\mathbb{R}^n$.

Si f est différentiable en tout point de U, on dit que f est différentiable sur U.

Remarque 54

La différentiabilité de f en a est équivalente à l'existence d'une forme linéaire u_a sur $\mathbb{R}^n$ et une fonction ε telle que

$$f(a+h)-f(a)-u_a(h) = \|h\|\,\varepsilon(h) \text{ avec } \lim_{h\to 0}\varepsilon(h) = 0.$$

Proposition 55

La forme linéaire u_a introduite dans le définition précédente, quand elle existe, est unique.

Démonstration. Soit $\{e_1, ..., e_n\}$ la base canonique de $\mathbb{R}^n$. Supposons qu'il existe deux formes linéaires u_a et v_a vérifiant la définition 53. Il vient alors

$$f(a+h)-f(a)-u_a(h) = \|h\|\,\varepsilon_1(h) \text{ avec } \lim_{h\to 0}\varepsilon_1(h) = 0,$$
$$f(a+h)-f(a)-v_a(h) = \|h\|\,\varepsilon_2(h) \text{ avec } \lim_{h\to 0}\varepsilon_2(h) = 0.$$

En posant $w_a = u_a - v_a$ et $\varepsilon = \varepsilon_1 - \varepsilon_2$, il vient $w(h) = \|h\| \varepsilon(h)$ avec $\lim\limits_{h\to 0} \varepsilon(h) = 0$.

Pour $h = t\ e_i, t > 0$, on obtient $w(t\ e_i) = \|t\ e_i\| \varepsilon(t\ e_i)$ avec $\lim\limits_{t\to 0} \varepsilon(t\ e_i) = 0$.

D'où, en utilisant la linéarité de w et en simplifiant

$$w(e_i) = \|e_i\| \varepsilon(t\ e_i) \text{ avec } \lim_{t\to 0} \varepsilon(t\ e_i) = 0.$$

Ce qui donne $w(e_i) = 0, i = 1, ..., n$. Alors $w_a = 0$ et donc $u_a = v_a$. ∎

Définition 56

La forme linéaire u_a s'appelle différentielle de f en a et se note $Df(a)$.

Remarque 57

Soit $f = (f_1, ..., f_m)$ une fonction définie au voisinage d'un point $a \in \mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$.
La fonction f est différentiable en a si et seulement si chaque composante $f_j, j = 1, ..., m$ est différentiable en a.

Définition 58

Soit f une fonction de U dans $\mathbb{R}$ différentiable sur U.

On appelle fonction différentielle de f la fonction notée Df définie par

$$\begin{array}{rccl} Df: & U & \longrightarrow & (\mathbb{R}^n)^* \\ & x & \longmapsto & Df(x), \end{array}$$

où $(\mathbb{R}^n)^*$ désigne le dual de $\mathbb{R}^n$ i.e. l'espace vectoriel des formes linéaires sur $\mathbb{R}^n$.

Exemple 11

Soit u une forme linéaire sur $\mathbb{R}^n$. On a $u(x+h) = u(x) + u(h)$, pour tout $x, h \in \mathbb{R}^n$. D'où

$$\frac{u(a+h) - u(a) - u(h)}{\|h\|} = 0.$$

Il est alors clair (définition 53) que u est différentiable sur $\mathbb{R}^n$ et $Du(x) = u,\ \forall x \in \mathbb{R}^n$.

La fonction différentielle de u est alors la fonction constante définie par

$$\begin{array}{rccl} Du: & \mathbb{R}^n & \longrightarrow & (\mathbb{R}^n)^* \\ & x & \longmapsto & Du(x) = u. \end{array}$$ ▪

Exemple 12

Soit I un intervalle ouvert de $\mathbb{R}$ et f une fonction de I dans $\mathbb{R}$.

Montrons que la dérivabilité de f sur I implique sa différentiabilité et réciproquement.

Supposons f dérivable sur I et soit $x \in I$. On a alors

$$\lim_{\substack{h \to 0 \\ h \neq 0, x+h \in I}} \frac{f(x+h) - f(x)}{h} = f'(x).$$

D'où

$$\lim_{\substack{h \to 0 \\ h \neq 0, x+h \in I}} \frac{f(x+h) - f(x) - f'(x)h}{|h|} = 0.$$

Comme l'application

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ h & \longmapsto & f'(x)h. \end{array}$$

est linéaire, il découle alors que f est différentiable en x et $Df(x)(h) = f'(x)h, \forall x \in I,\ \forall h \in \mathbb{R}$.

Comme $Df(x)$ est une application linéaire on a $Df(x)(h) = hDf(x)(1) = f'(x)h,\ \forall h \in \mathbb{R}$.

D'où $Df(x)(1) = f'(x),\ \forall x \in I$.

Réciproquement, supposons f différentiable sur I, i.e.

$$\lim_{\substack{h \to 0 \\ h \neq 0, x+h \in I}} \frac{f(x+h) - f(x) - Df(x)(h)}{|h|} = 0.$$

Ce qui donne

$$\lim_{\substack{h \to 0 \\ h \neq 0, x+h \in I}} \frac{f(x+h) - f(x) - Df(x)(h)}{h} = 0.$$

Comme $Df(x)(h) = hDf(x)(1)$, on tire $\displaystyle\lim_{\substack{h \to 0 \\ h \neq 0, x+h \in I}} \frac{f(x+h) - f(x)}{h} = Df(x)(1)$.

Ce qui prouve que f est dérivable en x et que $f'(x) = Df(x)(1)$. ■

Remarque 59

Soit $f = (f_1, ..., f_m)$ une fonction de $U \subset \mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$ différentiable en $a \in U$.

La fonction différentielle de f en a notée $Df(a)$ définie par

$$\begin{array}{cccc} Df(a): & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & h & \longmapsto & Df(a)(h) = (Df_1(a)(h), ..., Df_m(a)(h)). \end{array}$$

Théorème 60

Si f est une fonction de U dans $\mathbb{R}$ différentiable en $a \in U$, elle est alors continue en a.

Démonstration. Si f est différentiable en a, on a alors

$$f(a+h) = f(a) + Df(a)(h) + \|h\|\,\varepsilon(h) \text{ avec } \lim_{h\to 0} \varepsilon(h) = 0.$$

Comme $\lim_{h\to 0} Df(a)(h) = Df(a)(0) = 0$, il vient $\lim_{h\to 0} f(a+h) = f(a)$. Il s'ensuit que f est continue en a. ■

Remarque 61

1. La réciproque de ce théorème est fausse.
2. Ce théorème montre que la non continuité de f en a entraîne sa non différentiabilité (c'est souvent sous cette forme qu'il est utilisé).

Les opérations algébriques sur les fonctions différentiables concernant sommes, différences, produits, quotients et multiplication par un scalaire sont les mêmes que celles utilisées dans le cas des fonctions dérivables d'une variable réelle.

2.3 Dérivées partielles, gradient et matrice jacobienne

Soit f une fonction définie au voisinage d'un point $a = (a_1, ..., a_n) \in \mathbb{R}^n$ à valeurs dans $\mathbb{R}$.
Pour $\delta > 0$ assez petit et i fixé entre 1 et n désignons par g_i la fonction réelle à une variable réelle, définie par

$$\begin{array}{rrcl} g_i: &]a_i - \delta, a_i + \delta[& \longrightarrow & \mathbb{R} \\ & t & \longmapsto & g_i(t) = f(a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n), \end{array}$$

(toutes les variables dans f sont fixées sauf la i-ième).

Définition 62

On dit que la fonction $f : U \to \mathbb{R}$ admet une dérivée partielle par rapport à la i-ième variable au point a si la fonction g_i introduite ci-dessus est dérivable au point a_i.
On appelle alors dérivée partielle de f par rapport à la i-ième variable, au point a le nombre noté $\dfrac{\partial f}{\partial x_i}(a)$ défini par

$$\frac{\partial f}{\partial x_i}(a) = g_i'(a_i) = \lim_{\substack{h\to 0 \\ h\neq 0}} \frac{f(a_1, ..., a_{i-1}, a_i + h, a_{i+1}, ..., a_n) - f(a)}{h}.$$

Le calcul de $\dfrac{\partial f}{\partial x_i}$ consiste à ne dériver l'expression de f que par rapport à x_i. Notons que les fonctions dérivées partielles $\dfrac{\partial f}{\partial x_i}$ sont aussi des fonctions de n variables à valeurs dans $\mathbb{R}$.

Exemple 13

La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \sin(xy)$ admet des dérivées partielles par rapport aux variables x et y en tout point (x, y) de $\mathbb{R}^2$ et on a

$$\frac{\partial f}{\partial x}(x, y) = y\cos(xy), \quad \frac{\partial f}{\partial y}(x, y) = x\cos(xy), \ \forall (x, y) \in \mathbb{R}^2. \quad \blacksquare$$

Proposition 63

Si une fonction $f : U \to \mathbb{R}$ est différentiable en a, elle admet en a des dérivées partielles par rapport à toutes les variables et on a

$$\begin{aligned} &\frac{\partial f}{\partial x_i}(a) = Df(a)(e_i), \quad i = 1, ..., n, \\ &Df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, h_i, \quad \forall h = (h_1, ..., h_n) \in \mathbb{R}^n, \end{aligned}$$

$\{e_1, ..., e_n\}$ étant la base canonique de $\mathbb{R}^n$.

Démonstration. L'hypothèse de différentiabilité de f en a s'écrit

$$f(a + h) = f(a) + Df(a)(h) + \|h\|\,\varepsilon(h) \ \text{ avec } \lim_{h\to 0} \varepsilon(h) = 0.$$

Pour $h = h_i e_i$, la relation précédente donne

$$f(a + h_i e_i) = f(a) + Df(a)(h_i e_i) + \|h_i e_i\|\,\varepsilon(h_i) \ \text{ avec } \lim_{h_i\to 0} \varepsilon(h_i) = 0,$$

ou encore, en tenant compte du fait que $Df(a)$ est une forme linéaire

$$\frac{f(a + h_i e_i) - f(a)}{h_i} = Df(a)(e_i) + \frac{|h_i|}{h_i}\|e_i\|\,\varepsilon(h_i) \ \text{ avec } \lim_{h_i\to 0} \varepsilon(h_i) = 0.$$

Ce qui implique que

$$\lim_{h_i\to 0} \frac{f(a_1, ..., a_{i-1}, a_i + h_i, a_{i+1}, ..., a_n) - f(a)}{h_i} = Df(a)(e_i).$$

La fonction f admet donc une dérivée partielle par rapport à x_i en a et on a

$$\frac{\partial f}{\partial x_i}(a) = Df(a)(e_i), \quad i = 1, ..., n.$$

D'où

$$Df(a)(h) = Df(a)\left(\sum_{i=1}^{n} h_i e_i\right) = \sum_{i=1}^{n} h_i Df(a)(e_i) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, h_i. \quad \blacksquare$$

Définition 64 (gradient)

Si f est une fonction de U dans $\mathbb{R}$ admettant des dérivées partielles en a, on appelle *gradient* de f en a le vecteur de $\mathbb{R}^n$ noté $\operatorname{grad} f(a)$ ou $\nabla f(a)$ défini par

$$\operatorname{grad} f(a) = \left(\frac{\partial f}{\partial x_i}(a)\right)_{1\leq i\leq n} = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), ..., \frac{\partial f}{\partial x_2}(a)\right).$$

Définition 65 (matrice jacobienne)

Si $F = (F_1, ..., F_m)$ est une fonction de U à valeurs dans $\mathbb{R}^m$ admettant des dérivées partielles en a, on appelle *matrice jacobienne* de F en a la matrice à m lignes et n colonnes notée $J_F(a)$ définie par

$$J_F(a) = \left(\frac{\partial F_i}{\partial x_j}(a)\right)_{1\leq i\leq m, 1\leq j\leq n} = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(a) & ... & \frac{\partial F_1}{\partial x_j}(a) & ... & \frac{\partial F_1}{\partial x_n}(a) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_i}{\partial x_1}(a) & ... & \frac{\partial F_i}{\partial x_j}(a) & ... & \frac{\partial F_i}{\partial x_n}(a) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(a) & ... & \frac{\partial F_m}{\partial x_j}(a) & ... & \frac{\partial F_m}{\partial x_n}(a) \end{bmatrix}.$$

Cette matrice est également notée par $\dfrac{\partial (F_1, ..., F_m)}{\partial (x_1, ..., x_n)}(a)$ ou $\dfrac{D(F_1, ..., F_m)}{D(x_1, ..., x_n)}(a)$.

À noter que certains ouvrages définissent la matrice jacobienne comme la transposée de la matrice ci-dessus.

Remarque 66

Avec les deux notions précédentes, les différentielles peuvent être écrites sous la forme

$$Df(a)(h) = \langle \operatorname{grad} f(a), h\rangle \text{ et } DF(a)(h) = J_F(a)\,h,\ \forall h = (h_1, ..., h_n) \in \mathbb{R}^n,$$

où le symbole $\langle\ ,\ \rangle$ désigne le produit scalaire euclidien de $\mathbb{R}^n$ qui est défini par

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i,\ \forall x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in \mathbb{R}^n.$$

Définition 67

Soit f une fonction de U dans $\mathbb{R}$. On dit que f est de *classe* $\mathcal{C}^1$ sur U si ses dérivées partielles $\dfrac{\partial f}{\partial x_i}(x), i = 1, ..., n$ existent en tout point x de U et si les fonctions dérivées partielles

$$\begin{array}{rccc} \dfrac{\partial f}{\partial x_i}: & U & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \dfrac{\partial f}{\partial x_i}(x) \end{array}$$

sont continues.

Exemple 14

La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \mathrm{Log}(x^2 + y^2 + 1)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ puisque

1. Ses dérivées partielles existent en tout point (x, y) de $\mathbb{R}^2$:
$$\frac{\partial f}{\partial x}(x, y) = \frac{2x}{x^2 + y^2 + 1}, \quad \frac{\partial f}{\partial y}(x, y) = \frac{2y}{x^2 + y^2 + 1}.$$
2. Les fonctions dérivées $(x, y) \mapsto \dfrac{\partial f}{\partial x}$ et $(x, y) \mapsto \dfrac{\partial f}{\partial y}$ sont continues sur $\mathbb{R}^2$. ■

Exemple 15

Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \begin{cases} x^2 y \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

Il est facile de vérifier que $\dfrac{\partial f}{\partial x}(x, y)$ et $\dfrac{\partial f}{\partial y}(x, y)$ existent sur $\mathbb{R}^2$ et sont données par

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} y\left(2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}, \quad \frac{\partial f}{\partial y}(x, y) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

Montrons que f n'est pas continue aux points $(0, y_0)$, $y_0 \in \mathbb{R}^*$.
Pour cela considérons la suite de points $\left(\frac{1}{2\pi n}, y_0\right) \in \mathbb{R}^2$. Cette suite tend vers $(0, y_0)$ quand n tend vers l'infini, cependant

$$\lim_{n \to +\infty} \frac{\partial f}{\partial x}\left(\frac{1}{2\pi n}, y_0\right) = -y_0 \neq \frac{\partial f}{\partial x}(0, y_0) = 0.$$

Donc la fonction f n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. ■

Proposition 68

Soit f une fonction de U dans $\mathbb{R}$. Si f est de classe $\mathcal{C}^1$ sur U alors f est différentiable sur U.

Démonstration. Soit $a = (a_1, ..., a_n) \in U$. U étant ouvert, il existe une boule ouverte $B(a, r) \subset U$, ce qui signifie que

$$\forall h \in \mathbb{R}^n, \ \|h\| < r \text{ implique que } a + h \in B(a, r) \subset U \text{ où } \|h\| = \sum_{i=1}^{n} |h_i| .$$

Soit h vérifiant $\|h\| < r$. Posons $b_i = a + (h_1, ..., h_i, 0, ..., 0)\,,\ i = 1, ..., n$.
Alors $b_i \in B(a, r)$ car $\|b_i - a\| = \|(h_1, ..., h_i, 0, ..., 0)\| = \sum\limits_{j=1}^{i} |h_j| \leq \|h\| < r$.
On a

$$\begin{aligned} f(a+h) - f(a) &= f(b_n) - f(a) \\ &= f(b_n) - f(b_{n-1}) + f(b_{n-1}) - f(b_{n-2}) + ...f(b_1) - f(a) \\ &= \sum_{i=1}^{n} \left(f(b_i) - f(b_{i-1})\right) \text{ avec } b_0 = a. \end{aligned}$$

Comme la fonction f admet des dérivées partielles sur U, alors par application du théorème des accroissements finis à la fonction

$$t \longmapsto f(a_1 + h_1, ..., a_{i-1} + h_{i-1}, \ t\ , a_{i+1}, ..., a_n)\,,$$

on aura

$$\begin{aligned} f(b_i) - f(b_{i-1}) &= f(a_1 + h_1, ..., a_i + h_i, a_{i+1}, ..., a_n) - f(a_1 + h_1, ..., a_{i-1} + h_{i-1}, a_i, ..., a_n) \\ &= h_i \frac{\partial f}{\partial x_i}(c_i)\,, \end{aligned}$$

où $c_i = a + (h_1, ..., h_{i-1}, \theta_i h_i, 0, ..., 0)$ et $\theta_i \in \left]0, 1\right[\,,\ i = 1, ..., n$.
Ceci permet d'écrire

$$\begin{aligned} \left| f(a+h) - f(a) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) \right| &= \left| \sum_{i=1}^{n} h_i \left(\frac{\partial f}{\partial x_i}(c_i) - \frac{\partial f}{\partial x_i}(a) \right) \right| \\ &\leq \|h\| \max_{1 \leq i \leq n} \left| \frac{\partial f}{\partial x_i}(c_i) - \frac{\partial f}{\partial x_i}(a) \right|. \end{aligned}$$

Comme $\dfrac{\partial f}{\partial x_i}$ est continue, alors pour $\varepsilon > 0$, il existe $\delta_i > 0$ tel que

$$\forall x \in U \ \ \|x - a\| < \delta_i \Longrightarrow \left| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(a) \right| < \varepsilon.$$

Pour $x = c_i$ on a $\|x-a\| = \|c_i - a\| = \|(h_1, ..., h_{i-1}, \theta_i h_i, 0, ..., 0)\| \leq \|h\|$.
Si on prend $\delta = \min(r, \delta_1, ..., \delta_n)$, alors pour $\|h\| < \delta$ on aurait $\max\limits_{1\leq i\leq n} \left| \dfrac{\partial f}{\partial x_i}(c_i) - \dfrac{\partial f}{\partial x_i}(a) \right| < \varepsilon$.
Il s'ensuit alors que

$$\frac{\left| f(a+h) - f(a) - \sum\limits_{i=1}^{n} h_i \dfrac{\partial f}{\partial x_i}(a) \right|}{\|h\|} < \varepsilon \text{ tandis que } \|h\| < \delta.$$

Ce qui signifie que

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - \sum\limits_{i=1}^{n} h_i \dfrac{\partial f}{\partial x_i}(a)}{\|h\|} = 0.$$

Cette dernière relation jointe au fait que l'application

$$\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ h & \longmapsto & \sum\limits_{i=1}^{n} h_i \dfrac{\partial f}{\partial x_i}(a) \end{array}$$

est linéaire montre que f est différentiable en a et sa différentielle est

$$Df(a)(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a), \quad \forall h \in \mathbb{R}^n. \qquad \blacksquare$$

Remarque 69

La réciproque de cette proposition est fausse i.e. il existe des fonctions qui sont différentiables mais pas de classe $\mathcal{C}^1$. Autrement dit, l'ensemble des fonctions de classe $\mathcal{C}^1$ sur U est strictement inclus dans l'ensemble des fonctions différentiables.

Exemple 16

Reprenons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de l'exemple 15, qu'est définie par

$$f(x,y) = \begin{cases} x^2 y \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

On a vu que f n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. Posons $E = \{(0,y) \in \mathbb{R}^2,\ y \in \mathbb{R}\}$.
Sur $\mathbb{R}^2 \backslash E$ la fonction f est de classe $\mathcal{C}^1$ car ses dérivées partielles sont continues. Il reste à prouver la différentiabilité de f sur E.
On a

$$\frac{\left| f(h, y+k) - f(0,y) - \left(h\dfrac{\partial f}{\partial x}(0,y) + k\dfrac{\partial f}{\partial y}(0,y) \right) \right|}{|h| + |k|} = \begin{cases} \dfrac{h^2 (y+k) \sin\left(\frac{1}{h}\right)}{|h|+|k|} & \text{si } h \neq 0 \\ 0 & \text{si } h = 0 \end{cases}$$
$$\leq |h(y+k)|.$$

Ce qui implique que

$$\lim_{(h,k)\to(0,0)} \frac{\left|f(h,y+k)-f(0,y)-\left(h\dfrac{\partial f}{\partial x}(0,y)+k\dfrac{\partial f}{\partial y}(0,y)\right)\right|}{|h|+|k|}=0,$$

la fonction f est donc différentiable en tout point de E et $Df(0,y)=0$, $\forall y\in\mathbb{R}$. ∎

2.3.1 Interprétation géométrique de la différentielle

Soit f une fonction définie sur $U\subset\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, a un point de U, h un vecteur de $\mathbb{R}^n$. Supposons que f est différentiable en a. Alors on a

$$f(a+h)-f(a)=Df(a)(h)+\|h\|\,\varepsilon(h) \text{ avec } \lim_{h\to 0}\varepsilon(h)=0.$$

Lorsque $\|h\|$ est assez petit, on peut écrire $f(a+h)\approx f(a)+Df(a)(h)$. Ce qui signifie que f est peu différent de sa partie linéaire au voisinage de a.

$$f(x)\approx f(a)+\operatorname{grad} f(a)\cdot(x-a)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)(x_i-a_i).$$

Géométriquement, au voisinage de $(a,f(a))\in\mathbb{R}^{n+1}$, le graphe de f,

$$G_f=\left\{(x,f(x))\in\mathbb{R}^{n+1}\ ;\ \ x\in U\subset\mathbb{R}^n\right\}$$

diffère peu du plan de $\mathbb{R}^{n+1}$,

$$\{(x,y)\in\mathbb{R}^n\times\mathbb{R}\ ;\ \ x\in\mathbb{R}^n,\ y=f(a)+\operatorname{grad} f(a)\cdot(x-a)\}.$$

Ce plan est appelé le plan tangent de G_f en $(a,f(a))$ qui a pour équation

$$y=f(a)+\operatorname{grad} f(a)\cdot(x-a).$$

Ce plan est engendré par les vecteurs

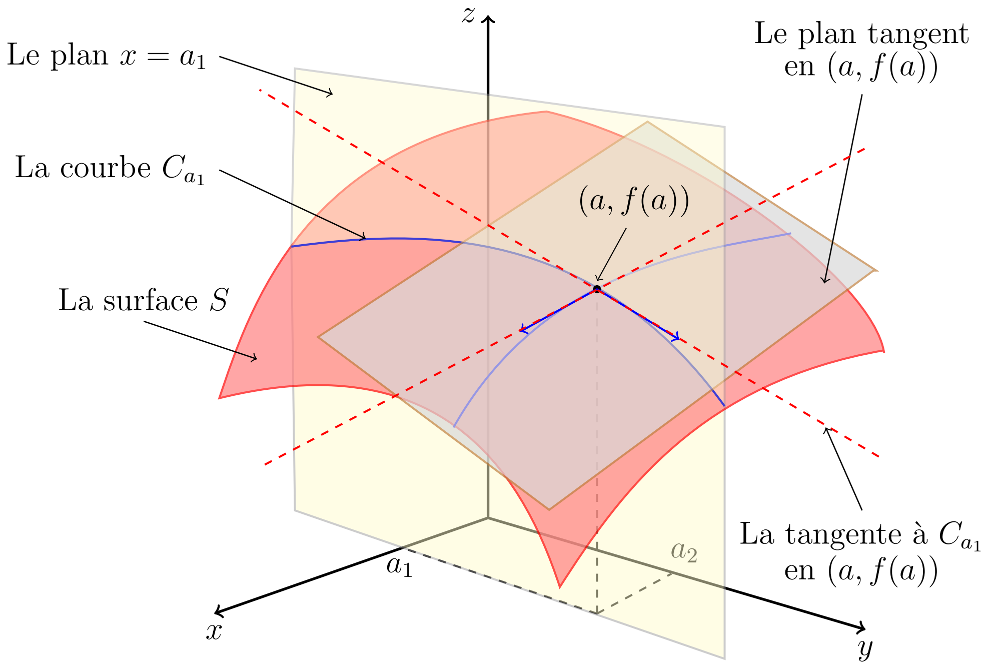
$$\frac{\partial}{\partial x_i}(x,f(x))(a),\, i=1...,n, \text{ ou encore } \left(e_i,\frac{\partial f}{\partial x_i}(a)\right), i=1...,n,$$

où $\{e_1,...,e_n\}$ est la base canonique de $\mathbb{R}^n$.

En dimension 2 le graphe de f est une surface S de $\mathbb{R}^3$ d'équation $z=f(x,y)$ et le plan tangent en $(a_1,a_2,f(a_1,a_2))$ a pour équation

$$z=f(a_1,a_2)+\frac{\partial f}{\partial x}(a)(x-a_1)+\frac{\partial f}{\partial y}(a)(y-a_2).$$

La section de la surface S par le plan d'équation $x=a_1$ est une courbe C_{a_1}. Cette courbe est le graphe de la fonction $y\mapsto z(y)=f(a_1,y)$, et $\dfrac{\partial f}{\partial y}(a_1,a_2)$ est la pente de sa tangente en $(a_1,a_2,f(a_1,a_2))$. De même pour la section de la surface S par le plan d'équation $y=a_2$.



2.4 Composition des fonctions différentiables

Soit n, m deux entiers non nuls, U un ouvert de $\mathbb{R}^n$, V un ouvert dans $\mathbb{R}^m$, F une fonction de U dans V, G une fonction de V dans $\mathbb{R}$, a un point de U.

$$\begin{tikzcd}
U \subset \mathbb{R}^n \arrow[r, "F"] \arrow[rr, bend left, "G \circ F"] & V \subset \mathbb{R}^m \arrow[r, "G"] & \mathbb{R} \\
\Cup & \Cup & \Cup \\
a & F(a) & (G \circ F)(a)
\end{tikzcd}$$

Théorème 70

Si F est différentiable en a et G différentiable en $F(a)$ alors la fonction $G \circ F$ est différentiable en a et on a

$$D(G \circ F)(a) = DG(F(a)) \circ DF(a).$$

Démonstration. Posons $b = F(a)$ et $H = G \circ F$.

La différentiabilité de F en a et de G en b signifie que

$$F(a+h) - F(a) = DF(a)(h) + \|h\| \varepsilon_1(h) \text{ avec } \lim_{h \to 0} \varepsilon_1(h) = 0,$$

$$G(b+k) - G(b) = DG(b)(k) + \|k\| \varepsilon_2(k) \text{ avec } \lim_{k \to 0} \varepsilon_2(k) = 0.$$

En prenant $k = k(h) = F(a+h) - F(a) = F(a+h) - b$, on déduit

$$\begin{aligned} H(a+h) - H(a) &= G(F(a+h)) - G(F(a)) \\ &= G(b+k) - G(b) = DG(b)(k) + \|k\| \varepsilon_2(k) \\ &= DG(b)\Big(DF(a)(h) + \|h\| \varepsilon_1(h)\Big) + \|k\| \varepsilon_2(k) \\ &= (DG(b) \circ DF(a))(h) + \|h\| \varepsilon(h), \end{aligned}$$

où $\varepsilon(h) = DG(b)(\varepsilon_1(h)) + \dfrac{\|k(h)\|}{\|h\|} \varepsilon_2(k(h))$.
Comme $DG(b)$ et $DF(a)$ sont des applications linéaires, alors $DG(b) \circ DF(a)$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$. Par suite, pour prouver que H est différentiable en a, il suffit de prouver que $\lim\limits_{h \to 0} \varepsilon(h) = 0$. Et alors on aura

$$DH(a)(h) = D(G \circ F)(a)(h) = (DG(F(a)) \circ DF(a))(h).$$

D'une part, $\dfrac{\|k(h)\|}{\|h\|}$ est borné au voisinage de $h = 0$ car

$$\begin{aligned} \frac{\|k(h)\|}{\|h\|} = \frac{\|DF(a)(h) + \|h\| \varepsilon_1(h)\|}{\|h\|} &= \left\| DF(a)\left(\frac{h}{\|h\|}\right) + \varepsilon_1(h) \right\| \\ &\leq \left\| DF(a)\left(\frac{h}{\|h\|}\right) \right\| + \|\varepsilon_1(h)\| \leq \|DF(a)\| + \|\varepsilon_1(h)\|. \end{aligned}$$

D'autre part on a $\lim\limits_{h \to 0} DG(b)(\varepsilon_1(h)) = 0$ et $\lim\limits_{h \to 0} \varepsilon_2(k(h))$ car $DG(b)$ est une application linéaire et F est une fonction continue puisqu'elle est différentiable.
Par conséquent $\lim\limits_{h \to 0} \varepsilon(h) = 0$. Ce qui termine la démonstration. ■

Corollaire 71

Si F et G sont différentiables sur U et V respectivement alors $G \circ F$ est différentiable sur U et ses dérivées partielles sont données par

$$\frac{\partial (G \circ F)}{\partial x_i}(x) = \sum_{j=1}^{m} \frac{\partial G}{\partial y_j}(F(x)) . \frac{\partial F_j}{\partial x_i}(x), \quad \forall x \in U, \ i = 1, ..., n.$$

Cette formule est appelée la règle de dérivation en chaîne.

Exemple 17 (Passage en coordonnées polaire)

Soit φ la fonction associée au changement de variables en coordonnées polaires et $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

$$\begin{tikzcd} \mathbb{R}_+^* \times [0, 2\pi[\arrow[r, "\varphi"] \arrow[rr, bend left, "g = f \circ \varphi"] & \mathbb{R}^2 \arrow[r, "f"] & \mathbb{R} \end{tikzcd}$$

$$\begin{array}{ccccc} (r, \theta) & \longmapsto & \varphi(r, \theta) = (x(r, \theta), y(r, \theta)) & \longmapsto & g(r, \theta) = f(x, y) \\ & & = (r\cos\theta, r\sin\theta) & & = f(r\cos\theta, r\sin\theta) \end{array}$$

Les fonctions composantes de φ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^* \times [0, 2\pi[$ car leurs dérivées partielles

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

sont de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^* \times [0, 2\pi[$. Il s'ensuit alors que la fonction $g = f \circ \varphi$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^* \times [0, 2\pi[$ et on a

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \frac{\partial f}{\partial x}(\varphi(r, \theta))\frac{\partial x}{\partial r}(r, \theta) + \frac{\partial f}{\partial y}(\varphi(r, \theta))\frac{\partial y}{\partial r}(r, \theta), \\ \frac{\partial g}{\partial \theta}(r, \theta) &= \frac{\partial f}{\partial x}(\varphi(r, \theta))\frac{\partial x}{\partial \theta}(r, \theta) + \frac{\partial f}{\partial y}(\varphi(r, \theta))\frac{\partial y}{\partial \theta}(r, \theta), \end{aligned}$$

ou encore

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \cos\theta\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta), \\ \frac{\partial g}{\partial \theta}(r, \theta) &= -r\sin\theta\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta\frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta). \end{aligned} \qquad \blacksquare$$

2.5 Dérivée suivant un vecteur - Dérivée directionnelle

Soit f une fonction de U dans $\mathbb{R}$, a un point de U, h un vecteur non nul de $\mathbb{R}^n$. U étant ouvert, il existe une boule ouverte $B(a, r) \subset U$, ce qui signifie que

$$\forall x \in \mathbb{R}^n, \ \|x\| < r \text{ implique que } a + x \in B(a, r) \subset U.$$

Soit r_0 un réel strictement positif vérifiant $\|r_0 h\| < r$. Posons $I_{r_0} =]-r_0, r_0[$.

Soit ainsi la fonction $\varphi : I_{r_0} \to \mathbb{R}$ définie par $\varphi(t) = f(a + t\,h)$.

Noter que $a + t\,h \in B(a, r) \subset U$ car $\|t\,h\| < \|r_0 h\| < r$.

Définition 72

On dit que f admet une dérivée en a suivant le vecteur h si la fonction φ précédente est dérivable en 0. Dans ce cas, le nombre $\varphi'(0)$ s'appelle dérivée de f en a suivant le vecteur h et se note $d_h f(a)$ qui est donnée par

$$d_h f(a) = \lim_{t\to 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t\to 0} \frac{f(a + t\,h) - f(a)}{t}.$$

La dérivée de f en a suivant le vecteur h mesure les variations de f lorsqu'on se déplace autour de a dans la direction du vecteur h. Si h est le vecteur nul, cette dérivée existe toujours et a une valeur nulle.

Remarque 73

Si $\|h\| = 1$, la dérivée précédente porte le nom de dérivée directionnelle de f en a suivant la direction h.

Proposition 74

Soit f une fonction de U dans $\mathbb{R}$. Si f est différentiable en a alors elle admet en a une dérivée suivant n'importe quel vecteur h non nul et on a

$$Df(a)(h) = d_h f(a).$$

Démonstration. Soit h un vecteur no nul de $\mathbb{R}^n$ et t un réel vérifiant $a + t\,h \in B(a, r) \subset U$. On a

$$f(a + t\,h) - f(a) = Df(a)(t\,h) + \|t\,h\|\,\varepsilon(t\,h) \text{ avec } \lim_{t\to 0} \varepsilon(t\,h) = 0.$$

En utilisant la linéarité de $Df(a)$ on en déduit que

$$\frac{f(a + t\,h) - f(a)}{t} = Df(a)(h) + \frac{|t|}{t}\|h\|\,\varepsilon(t\,h) \text{ avec } \lim_{t\to 0} \varepsilon(t\,h) = 0.$$

D'où $\lim_{t\to 0} \dfrac{f(a + t\,h) - f(a)}{t} = Df(a)(h)$, ce qui donne $d_h f(a) = Df(a)(h)$. ∎

Remarque 75

On voit en particulier que $\dfrac{\partial f}{\partial x_i}(a)$ correspond à la dérivée en a suivant le vecteur de la base canonique e_i.

$$\frac{\partial f}{\partial x_i}(a) = Df(a)(e_i) = d_{e_i} f(a).$$

La réciproque de la proposition 74 est fausse, il existe des fonctions qui admettent des dérivées suivant tout vecteur en un point mais ne sont pas différentiables en ce point comme le prouve l'exemple suivant.

Exemple 18

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}.$$

Pour $h = (h_1, h_2) \neq (0,0)$, on a

$$\lim_{t\to 0} \frac{f(t\,h_1, t\,h_2) - f(0,0)}{t} = \lim_{t\to 0} \frac{t^3 h_1^2 h_2}{t^3 (h_1^2 + h_2^2)} = \frac{h_1^2 h_2}{h_1^2 + h_2^2}.$$

Alors, la fonction f admet une dérivée suivant tout vecteur h en $(0,0)$ et on a

$$d_h f(0,0) = \frac{h_1^2 h_2}{h_1^2 + h_2^2}.$$

Montrons que f n'est pas différentiable en $(0,0)$. Un calcul simple montre que

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

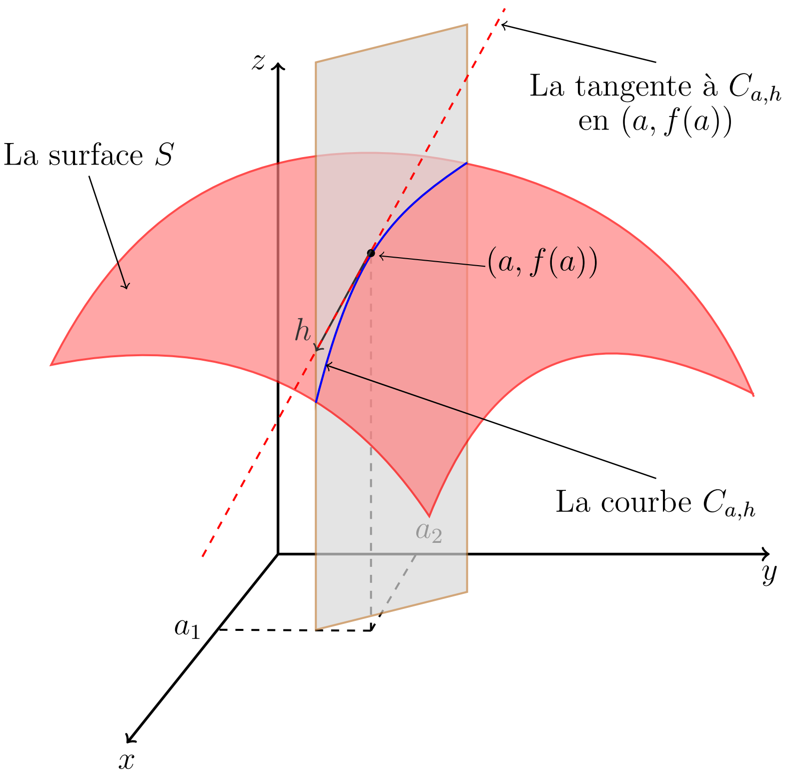
Si f était différentiable en $(0,0)$ on aurait $\lim\limits_{(x,y)\to(0,0)} \dfrac{f(x,y)}{\sqrt{x^2+y^2}} = 0.$
Montrons que cela est faux. On a pour $y = \lambda |x|,\ \lambda \neq 0,$

$$\frac{f(x, \lambda |x|)}{\sqrt{x^2 + \lambda^2 |x|^2}} = \frac{\lambda x^2 |x|}{\left(x^2 + \lambda^2 |x|^2\right)^{\frac{3}{2}}} = \frac{\lambda}{(1+\lambda^2)^{\frac{3}{2}}}.$$

Alors la limite de $\dfrac{f(x,y)}{\sqrt{x^2+y^2}}$ n'existe pas quand (x,y) tend vers $(0,0)$. La fonction f n'est donc pas différentiable en $(0,0)$. ▪

2.5.1 Interprétation géométrique en dimension 2

Soit f une fonction de $U \subset \mathbb{R}^2$ dans $\mathbb{R}$, différentiable en $a = (a_1, a_2) \in U$, $h = (h_1, h_2)$ un vecteur de $\mathbb{R}^2$. Le graphe de f est une surface S de $\mathbb{R}^3$ d'équation $z = f(x,y)$.
Si on coupe la surface S par le plan vertical contenant la droite passant par a et dirigée par h, on obtient la courbe paramétrée $C_{a,h} : t \mapsto (a + t\,h, f(a + t\,h))$. La dérivée $d_h f(a)$ de f au point a suivant le vecteur h est la pente de la tangente à la courbe $C_{a,h}$ en $(a, f(a))$.



L'inégalité de Cauchy-Schwarz permet d'écrire $|d_h f(a)| = |\langle \operatorname{grad} f(a), h \rangle| \leq \|\operatorname{grad} f(a)\| \, \|h\|$ avec égalité si et seulement si h est colinéaire au gradient de f en a.

La pente de la tangente en a est donc maximale en choisissant la direction du gradient en a, ce qui est à la base des méthodes de descente dans les problèmes de minimisation.