20 Calcul différentiel

« Les mathématiques sont un jeu que l'on exerce selon des règles simples en manipulant des symboles ou des concepts qui n'ont en soi, aucune importance particulière. »

David Hilbert (1862 – 1943)

Plan de cours

I Premiers pas avec les fonctions de deux variables à valeurs réelles 1
II Applications différentiables 8

Dans tout ce chapitre, E et F désignent des espaces vectoriels normés de dimension finie. On cherche à étendre la notion de dérivabilité (relative aux fonctions d'une variable réelle) aux fonctions plus généralement définies sur un ouvert $\mathscr{U}$ de E et à valeurs dans F. Le choix de travailler sur un ouvert permet de s'assurer que si $a \in \mathscr{U}$, alors, pour $\|h\|$ suffisamment petite, $a+h \in \mathscr{U}$: on donne ainsi un sens à l'écriture $f(a+h)-f(a)$.

Comme nous le verrons tout au long de ce chapitre, le calcul différentiel se construit autour d'un grand principe qu'il est utile d'avoir dès à présent en tête :

$$\begin{pmatrix}\text{accroissement}\\ \text{de la fonction}\end{pmatrix} = \begin{pmatrix}\text{terme linéaire par rapport}\\ \text{à l'accroissement de la variable}\end{pmatrix} + \begin{pmatrix}\text{terme}\\ \text{correctif}\end{pmatrix}$$

Ce qui se traduit, pour une « simple » fonction $f : \mathbb{R} \to \mathbb{R}$ dérivable, par : $f(a+h)-f(a) \underset{h\to 0}{=} f'(a)h + o(h)$.

I | Premiers pas avec les fonctions de deux variables à valeurs réelles

Débutons notre étude en la restreignant aux fonctions définies sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. On introduit à cet effet la fonction :

$$f : \left| \begin{array}{rcl} \mathscr{U} \subset \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array} \right.$$

Exercice 1

Représenter l'ensemble de définition des fonctions définies par :

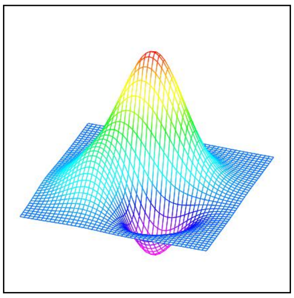
$$f_1(x,y) = \frac{xy}{\sqrt{x^2+y^2}} \ ; \quad f_2(x,y) = \frac{xy}{\sqrt{x^2+y^2-1}} \ ; \quad f_3(x,y) = \frac{xy}{\sqrt{x+y-1}}$$

La surface d'équation $z = f(x,y)$ est l'ensemble des points de coordonnées $(x,y,f(x,y))$ pour $(x,y) \in \mathscr{D}_f$.

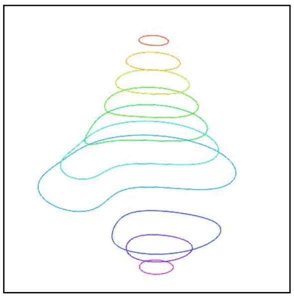
Définition 20.1 : Ligne de niveau

On appelle ligne de niveau de hauteur k (ou d'altitude k) la courbe d'équation $f(x,y) = k$.

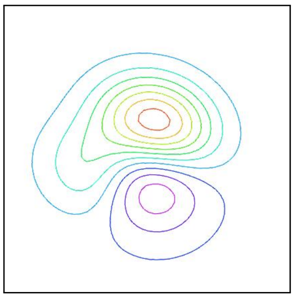
Les lignes de niveau permettent de visualiser la surface d'équation $z = f(x,y)$.



Exemple de surface



Lignes de niveau 3D



Lignes de niveau

A – Continuité

Par la suite, on munira $\mathbb{R}^2$ de la norme euclidienne notée $\|\cdot\|$. Rappelons que f est continue en $(x_0, y_0) \in \mathscr{U}$ si :

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall (x, y) \in \mathscr{U}, \quad \|(x, y) - (x_0, y_0)\| < \eta \Longrightarrow \left|f(x, y) - f(x_0, y_0)\right| < \varepsilon$$

La somme, le produit, la composée et le quotient dont le dénominateur ne s'annule pas de fonctions continues sont des fonctions continues.

Quelques remarques :

- Si f est continue, alors l'image de tout compact de $\mathbb{R}^2$ (i.e. un fermé borné) est un fermé borné de $\mathbb{R}$. Si A est un compact connexe par arcs de $\mathbb{R}^2$, $f(A)$ sera même un segment (sous réserve de continuité de f).
- Si f est continue alors les applications partielles $f_x : y \mapsto f(x, y)$ et $f_y : x \mapsto f(x, y)$ le sont également. Attention, la réciproque est fausse !

Exemple

Soit f la fonction définie par $f(x, y) = \dfrac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$. f n'est pas continue en 0 :

$$f(0, y) = 0 \xrightarrow[y \to 0]{} 0 \quad \text{alors que} \quad f(x, x) = \frac{1}{2} \xrightarrow[x \to 0]{} \frac{1}{2}$$

B – Dérivées partielles premières

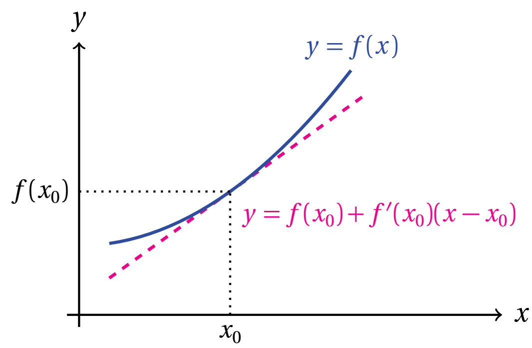
1 – Retour sur les fonctions d'une variable réelle

Si $f : \mathbb{R} \to \mathbb{R}$ est dérivable sur $\mathbb{R}$, la formule de Taylor-Young nous permet d'écrire en tout point $x_0 \in \mathbb{R}$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \mathrm{o}(x - x_0) \quad (\star)$$

L'équation de la tangente à la courbe au point d'abscisse x_0 est alors :

$$y = f(x_0) + f'(x_0)(x - x_0)$$



L'égalité $(\star)$ peut s'écrire :

$$f(x_0 + h) \underset{h \to 0}{=} \underbrace{f(x_0) + \underbrace{f'(x_0) \cdot h}_{\text{partie linéaire}}}_{\text{partie affine}} + \mathrm{o}(h)$$

L'application $h \mapsto f'(x_0) \cdot h$ est linéaire. On l'appelle différentielle de f en x_0 et on la note $\mathrm{d}f_{x_0}$. $\mathrm{d}f_{x_0} \in \mathscr{L}(\mathbb{R})$.

2 – Fonctions de deux variables : dérivées partielles, différentielle et plan tangent

On considère à nouveau une fonction f définie sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.

Définition 20.2

Sous réserve d'existence, on définit les dérivées partielles premières de f au point $(x_0, y_0) \in \mathscr{U}$ par :

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}$$

$\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont, comme f, des fonctions de deux variables à valeurs réelles.

Définition 20.3

On dit que f est de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ si les dérivées partielles existent et sont continues en tout point de $\mathscr{U}$.

Exemple

L'application $(x,y)\mapsto 3x^2+5x\sin(y)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et :

$$\forall (x,y)\in\mathbb{R}^2,\quad \frac{\partial f}{\partial x}(x,y)=6x+5\sin(y)\quad \text{et}\quad \frac{\partial f}{\partial x}(x,y)=5x\cos(y)$$

On admet provisoirement le résultat suivant.

Théorème 20.4 : Formule de Taylor-Young à l'ordre 1

Si f est de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ alors, pour tout $(x_0,y_0)\in\mathscr{U}$:

$$f(x_0+h,y_0+k)\underset{(h,k)\to(0,0)}{=} f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)\cdot h+\frac{\partial f}{\partial y}(x_0,y_0)\cdot k+\mathrm{o}(\|(h,k)\|)$$

La surface d'équation $z=f(x,y)$ admet alors un plan tangent en tout point (x_0,y_0,z_0) tel que $z_0=f(x_0,y_0)$. Celui-ci a pour équation :

$$z=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)\cdot(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)\cdot(y-y_0)$$

C'est bien une équation de plan, de la forme $ax+by+cz=d$ avec $a=\frac{\partial f}{\partial x}(x_0,y_0)$, $b=\frac{\partial f}{\partial y}(x_0,y_0)$ et $c=-1$.

Le vecteur $\left(\frac{\partial f}{\partial x}(x_0,y_0),\frac{\partial f}{\partial y}(x_0,y_0),-1\right)$ est normal au plan tangent.

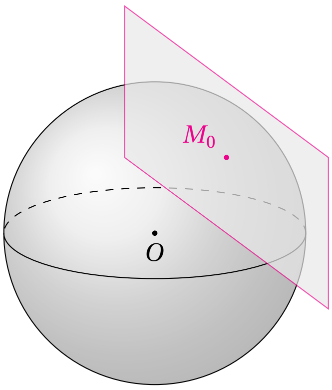
Exemple

Considérons la sphère de centre O et de rayon 1 qui a pour équation $x^2+y^2+z^2=1$.

On a alors $z=\pm\sqrt{1-x^2-y^2}$ et on pose :

$$f:(x,y)\mapsto\sqrt{1-x^2-y^2}$$

La surface représentative de f est une demi-sphère.



La sphère et son plan tangent en (x_0,y_0,z_0)

Déterminons une équation du plan tangent au point M_0 de coordonnées $(x_0,y_0,z_0)=(0,1/\sqrt{2},1/\sqrt{2})$.

Tout d'abord :

$$\frac{\partial f}{\partial x}(x_0,y_0)=\frac{-x_0}{\sqrt{1-x_0^2-y_0^2}}\quad\text{et}\quad\frac{\partial f}{\partial y}(x_0,y_0)=\frac{-y_0}{\sqrt{1-x_0^2-y_0^2}}$$

On a ici $x_0=0$ et $y_0=z_0=\frac{1}{\sqrt{2}}$ d'où :

$$z=\frac{1}{\sqrt{2}}+0\cdot(x-0)+(-1)\cdot\left(y-\frac{1}{\sqrt{2}}\right)\quad\text{c'est-à-dire :}\quad y+z=\sqrt{2}$$

Si f de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ et $(x_0, y_0) \in \mathscr{U}$, l'application

$$(h,k) \mapsto \frac{\partial f}{\partial x}(x_0, y_0) \cdot h + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k$$

est linéaire ; on l'appelle différentielle de f en (x_0, y_0). On la note $\mathrm{d}f_{(x_0,y_0)}$. On observera que $\mathrm{d}f_{(x_0,y_0)} \in \mathscr{L}(\mathbb{R}^2, \mathbb{R})$. Sa matrice représentative dans les bases canoniques de $\mathbb{R}^2$ et de $\mathbb{R}$ est :

$$\begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}$$

Définition 20.5 : Gradient

Soit f de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ et $(x_0, y_0) \in \mathscr{U}$. On appelle gradient de f au point (x_0, y_0) le vecteur $\left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)\right)$. On le notera $\overrightarrow{\mathrm{grad}}\, f(x_0, y_0)$ ou $\nabla f(x_0, y_0)$.

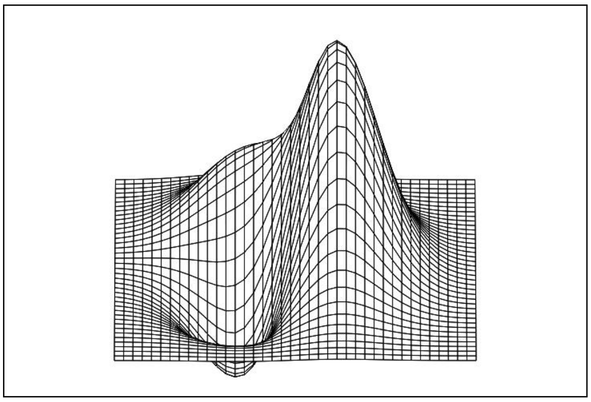
Autant pour la différentielle que pour le gradient, nos amis physiciens omettent parfois de préciser « (x_0, y_0) » mais il faut garder à l'esprit que l'on parle de la différentielle et du gradient en un point donné.

Proposition 20.6

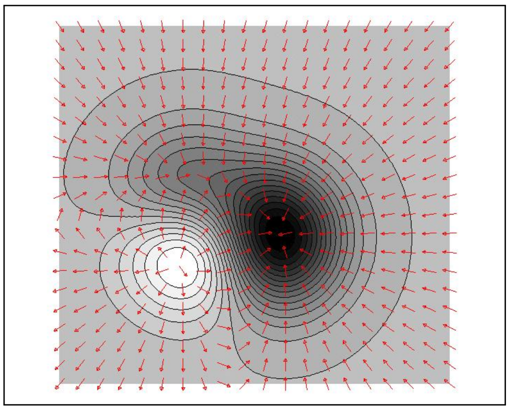
Le gradient au point M_0 est orthogonal à la ligne de niveau passant par M_0.

Le gradient indique en outre la ligne de plus grande pente.

Voici une illustration de cette propriété pour $f : (x, y) \mapsto \mathrm{e}^{-(x-1)^2 - (y+1/4)^2} + y\mathrm{e}^{-x^2-y^2} + 3$:



Représentation de la surface d'équation $z = f(x, y)$



Gradient et lignes de niveau associés

3 – Notations

Nous venons d'introduire la différentielle de f au point $(x_0, y_0) \in \mathscr{U}$:

$$\mathrm{d}f_{(x_0,y_0)} : (h,k) \mapsto \frac{\partial f}{\partial x}(x_0, y_0) \cdot h + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k$$

Si l'on pose $\mathrm{d}x : (h,k) \mapsto h$ et $\mathrm{d}y : (h,k) \mapsto k$, on peut alors écrire :

$$\mathrm{d}f_{(x_0,y_0)} = \frac{\partial f}{\partial x}(x_0, y_0)\,\mathrm{d}x + \frac{\partial f}{\partial y}(x_0, y_0)\,\mathrm{d}y$$

Et enfin, en notant $\mathrm{d}f$ l'application $(x, y) \mapsto \mathrm{d}f_{(x,y)}$ définie sur $\mathscr{U}$ et à valeurs dans $\mathscr{L}(\mathbb{R}^2, \mathbb{R})$:

$$\mathrm{d}f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y$$

4 – Dérivées partielles et composées : règle de la chaîne

▹ Pour $\mathscr{U}$ un ouvert de $\mathbb{R}^2$ et I un intervalle de $\mathbb{R}$, on considère les deux applications de classe $\mathscr{C}^1$:

$$\varphi : \left| \begin{array}{rcl} I & \longrightarrow & \mathscr{U} \\ t & \longmapsto & (x(t), y(t)) \end{array} \right. \qquad \text{et} \qquad f : \left| \begin{array}{rcl} \mathscr{U} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array} \right.$$

$f \circ \varphi : I \to \mathbb{R}$ est alors de classe $\mathscr{C}^1$ sur I et :

$$\forall t \in I \quad (f \circ \varphi)'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t))$$

Démonstration

φ est de classe $\mathscr{C}^1$ sur I, donc x et y sont dérivables et pour tout $t \in I$:

$$\begin{aligned}(f\circ\varphi)(t+h) - (f\circ\varphi)(t) &= f(x(t+h), y(t+h)) - f(x(t), y(t)) \\ &\underset{h\to 0}{=} f(x(t) + \underbrace{hx'(t) + \mathrm{o}(h)}_{=h'}, y(t) + \underbrace{hy'(t) + \mathrm{o}(h)}_{=k'}) - f(x(t), y(t))\end{aligned}$$

Or $f(x(t)+h', y(t)+k') \underset{h\to 0}{=} f(x(t), y(t)) + \dfrac{\partial f}{\partial x}(x(t), y(t)) \cdot h' + \dfrac{\partial f}{\partial y}(x(t), y(t)) \cdot k' + \mathrm{o}(\sqrt{h'^2 + k'^2})$. D'où,

$$(f\circ\varphi)(t+h) - (f\circ\varphi)(t) \underset{h\to 0}{=} \left[x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + hy'(t) \frac{\partial f}{\partial y}(x(t), y(t)) \right] \cdot h + \mathrm{o}(h)$$

■

Exercice 2

Montrer que $t \mapsto f(\cos(t), \sin(t))$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et calculer sa dérivée, avec $f \in \mathscr{C}^1(\mathbb{R}^2; \mathbb{R})$.

▹ Considérons désormais les applications $f : \mathbb{R}^2 \to \mathbb{R}$, $\varphi : \mathbb{R}^2 \to \mathbb{R}$ et $\psi : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et :

$$F : \left| \begin{array}{rcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(\varphi(x,y), \psi(x,y)) \end{array} \right.$$

Alors F est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et pour tout $(x,y) \in \mathbb{R}^2$,

$$\left| \begin{array}{l} \dfrac{\partial F}{\partial x}(x,y) = \dfrac{\partial \varphi}{\partial x}(x,y) \cdot \dfrac{\partial f}{\partial x}(\varphi(x,y), \psi(x,y)) + \dfrac{\partial \psi}{\partial x}(x,y) \cdot \dfrac{\partial f}{\partial y}(\varphi(x,y), \psi(x,y)) \\ \dfrac{\partial F}{\partial y}(x,y) = \dfrac{\partial \varphi}{\partial y}(x,y) \cdot \dfrac{\partial f}{\partial x}(\varphi(x,y), \psi(x,y)) + \dfrac{\partial \psi}{\partial y}(x,y) \cdot \dfrac{\partial f}{\partial y}(\varphi(x,y), \psi(x,y)) \end{array} \right.$$

Exemple

Si $f : \mathbb{R}^2 \to \mathbb{R}$ est de classe $\mathscr{C}^1$, il en va de même pour $F : (r, \theta) \mapsto f(r\cos(\theta), r\sin(\theta))$ et,

$$\forall (r,\theta) \in \mathbb{R}^2, \quad \frac{\partial F}{\partial r}(r,\theta) = \cos(\theta) \frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta)) + \sin(\theta) \frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta))$$

$$\forall (r,\theta) \in \mathbb{R}^2, \quad \frac{\partial F}{\partial \theta}(r,\theta) = -r\sin(\theta) \frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta)) + r\cos(\theta) \frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta))$$

C – Dérivées partielles secondes

Elles sont au nombre de 4 et définies à partir des dérivées partielles premières : $\dfrac{\partial^2 f}{\partial x^2}$, $\dfrac{\partial^2 f}{\partial x \partial y}$, $\dfrac{\partial^2 f}{\partial y \partial x}$ et $\dfrac{\partial^2 f}{\partial y^2}$.

Définition 20.7

On dit que f est de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ si les dérivées partielles secondes existent et sont continues en tout point de $\mathscr{U}$.

Théorème 20.8 : Théorème de Schwarz

Soit $f : \mathcal{U} \to \mathbb{R}^2$ une application de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^2$. Alors,

$$\forall (x_0, y_0) \in \mathcal{U}, \quad \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

D – Application à l'étude des extrema

Définition 20.9 : Extremum local

On dit que $f : \mathbb{R}^2 \to \mathbb{R}$ admet un minimum (resp. un maximum) local en (x_0, y_0) s'il existe un voisinage $\mathcal{U}$ de (x_0, y_0) tel que :

$$\forall (x, y) \in \mathcal{U}, \quad f(x, y) \geqslant f(x_0, y_0) \quad (\text{resp.} f(x, y) \leqslant f(x_0, y_0))$$

Définition 20.10 : Point critique

Soit f une fonction de classe $\mathscr{C}^1$ sur un **ouvert** $\mathcal{U}$ de $\mathbb{R}^2$. On dit que (x_0, y_0) est un point critique de f si :

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0 \quad \text{soit} \quad \overrightarrow{\text{grad}} f(x_0, y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix} = \overrightarrow{0}$$

Théorème 20.11 : Condition nécessaire d'existence

Si (x_0, y_0) est un extremum local de $f : \mathcal{U} \to \mathbb{R}$ de classe $\mathscr{C}^1$ alors (x_0, y_0) est un point critique de f.

Attention, $\mathcal{U}$ doit être **ouvert**.

Démonstration

Supposons que $f : \mathbb{R}^2 \to \mathbb{R}$ atteigne un minimum local en (x_0, y_0). Soit alors $\mathcal{U}$ un ouvert sur lequel :

$$\forall (x, y) \in \mathcal{U}, \quad f(x, y) - f(x_0, y_0) \geqslant 0$$

Pour h suffisamment proche de 0, on peut écrire $f(x_0 + h, y_0) - f(x_0, y_0) \geqslant 0$.

- Si $h \geqslant 0$, par passage à la limite, $\dfrac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \geqslant 0$ implique $\dfrac{\partial f}{\partial x}(x_0, y_0) \geqslant 0$.
- Si $h \leqslant 0$, par passage à la limite, $\dfrac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \leqslant 0$ implique $\dfrac{\partial f}{\partial x}(x_0, y_0) \leqslant 0$.

On obtient donc $\dfrac{\partial f}{\partial x}(x_0, y_0) = 0$. Idem pour $\dfrac{\partial f}{\partial y}(x_0, y_0)$, ce qui permet de conclure. ∎

Les extremums locaux sont donc à rechercher parmi les points critiques. Cependant, la réciproque est fausse ! Tout point critique ne correspond pas nécessairement à un extremum local.

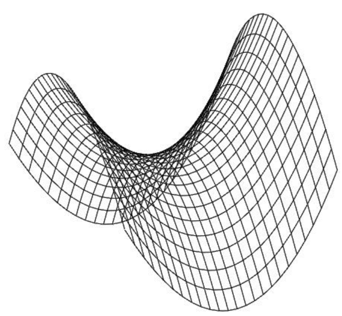
Exemple (Point selle)

Considérons la fonction $f : (x, y) \mapsto x^2 - y^2$.

$$\forall (x, y) \in \mathbb{R}^2, \quad \overrightarrow{\text{grad}} f(x, y) = (2x, -2y)$$

f admet $(0, 0)$ comme seul point critique mais ce point ne correspond ni à un maximum, ni à un minimum. En effet, $f(0, 0) = 0$ et pour tous x, y non nuls,

$$f(x, 0) = x^2 > f(0, 0) \text{ et } f(0, y) = -y^2 < f(0, 0)$$



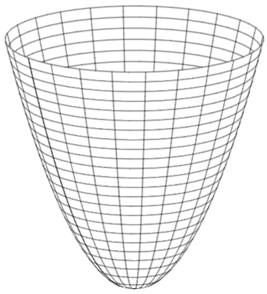
Point selle (ou point col)

Il sera donc nécessaire lors de l'étude d'une fonction de distinguer les points cols des points correspondant à des extrema (en revenant à la définition d'un extremum).

Exemple

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

$$f(x,y) = x^2 + xy + y^2 + 2x + 3y$$



Représentation de la fonction

- Recherche des points critiques
  f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ car polynomiale.

  $$\forall (x,y) \in \mathbb{R}^2, \quad \overrightarrow{\text{grad}}\, f(x,y) = (2x + y + 2, 2y + x + 3)$$

  Il y a un seul point critique : $M_0(-1/3, -4/3)$.
- Recherche des extrema
  Étudions la fonction au voisinage de M_0. $f(-1/3, -4/3) = -7/3$ et :

  $$f(-1/3 + h, -4/3 + k) - f(-1/3, -4/3) = h^2 + hk + k^2 = \frac{3}{4}(h+k)^2 + \frac{1}{4}(h-k)^2 \geqslant 0$$

  Autre possibilité pour vérifier que cette dernière quantité est bien positive : c'est un trinôme en h de discriminant négatif, donc de signe constant (ici positif puisque de coefficient dominant égal à 1).
  Ainsi, f atteint en M_0 un minimum que l'on visualise bien sur le paraboloïde ci-dessus.

Notons que $f(M_0) = -7/3$ est même un minimum global de la fonction.

Ne figure au programme que la seule condition nécessaire d'existence d'un extremum. Pour trouver une condition suffisante, nous pourrions être tentés d'exploiter le résultat suivant (malheureusement hors programme) :

Théorème 20.12 : Formule de Taylor-Young à l'ordre 2

Si f est de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ alors, pour tout $(x_0, y_0) \in \mathscr{U}$:

$$f(x_0+h, y_0+k) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\cdot h + \frac{\partial f}{\partial y}(x_0,y_0)\cdot k + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_0,y_0)\cdot h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\cdot hk + \frac{\partial^2 f}{\partial y^2}(x_0,y_0)\cdot k^2\right] + \mathrm{o}(h^2+k^2)$$

En un point critique, la formule de Taylor-Young à l'ordre 2 donne :

$$f(x_0+h, y_0+k) - f(x_0, y0) \underset{(h,k)\to(0,0)}{=} \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_0,y_0)\cdot h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\cdot hk + \frac{\partial^2 f}{\partial y^2}(x_0,y_0)\cdot k^2\right] + \mathrm{o}(h^2+k^2)$$

Au voisinage de (x_0, y_0), le signe de $f(x_0+h, y_0+k) - f(x_0, y0)$ est donné par :

$$\frac{\partial^2 f}{\partial x^2}(x_0,y_0)\cdot h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\cdot hk + \frac{\partial^2 f}{\partial y^2}(x_0,y_0)\cdot k^2$$

si cette dernière quantité est non nulle. On retrouve un trinôme dont on peut déterminer s'il est de signe constant ou non au voisinage de $(0,0)$.

E – Exemples de résolution d'équations aux dérivées partielles

Nous n'aborderons la résolution d'équations aux dérivées partielles (EDP) qu'à travers des exemples relativement simples et, à ce titre, très restreints. $\mathscr{U}$ désignera ici l'ouvert $I \times J$ où I et J sont deux intervalles ouverts et f une solution des équations aux dérivées partielles proposées. Nous justifierons dans la section suivante le choix d'un tel domaine.

- $\forall (x,y) \in \mathscr{U}, \quad \dfrac{\partial f}{\partial x}(x,y) = 0 : \quad \forall (x,y) \in \mathscr{U}, \quad f(x,y) = \varphi(y) \quad$ avec $\quad \varphi : J \to \mathbb{R}$
- $\forall (x,y) \in \mathscr{U}, \quad \dfrac{\partial f}{\partial x}(x,y) = g(x,y) : \quad \forall (x,y) \in \mathscr{U}, \quad f(x,y) = \displaystyle\int g(x,y)\,\mathrm{d}x + \varphi(y) \quad$ avec $\quad \varphi : J \to \mathbb{R}$

- $\forall (x,y)\in \mathscr{U},\quad \dfrac{\partial^2 f}{\partial x^2}(x,y)=0:\quad \forall (x,y)\in\mathscr{U},\quad f(x,y)=\varphi(y)x+\psi(y)\quad$ avec $\quad\varphi,\psi : J\to\mathbb{R}$
- Équation des ondes, dite des cordes vibrantes ou encore de d'Alembert :

$$\forall (x,t)\in\mathbb{R}^2,\quad \frac{\partial^2 f}{\partial x^2}(x,t)-\frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}(x,t)=0\quad (\star)\quad (c\neq 0)$$

On passe par le changement de variables de classe $\mathscr{C}^2$ et bijectif (car $c\neq 0$) : $\begin{cases} u=x-ct\\ v=x+ct\end{cases}$

Partant de $f(x,t)=g(u,v)$, on cherche une équation aux dérivées partielles vérifiée par g.

$$\frac{\partial f}{\partial x}=\frac{\partial g}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial g}{\partial v}\frac{\partial v}{\partial x}=\frac{\partial g}{\partial u}+\frac{\partial g}{\partial v}$$
$$\frac{\partial^2 f}{\partial x^2}=\left(\frac{\partial^2 g}{\partial u^2}\frac{\partial u}{\partial x}+\frac{\partial^2 g}{\partial v\partial u}\frac{\partial v}{\partial x}\right)+\left(\frac{\partial^2 g}{\partial u\partial v}\frac{\partial u}{\partial x}+\frac{\partial^2 g}{\partial v^2}\frac{\partial v}{\partial x}\right)=\frac{\partial^2 g}{\partial u^2}+2\frac{\partial^2 g}{\partial u\partial v}+\frac{\partial^2 g}{\partial v^2}$$

De même, on trouve :

$$\frac{\partial^2 f}{\partial t^2}=c^2\left(\frac{\partial^2 g}{\partial u^2}-2\frac{\partial^2 g}{\partial u\partial v}+\frac{\partial^2 g}{\partial v^2}\right)$$

L'équation $(\star)$ devient :

$$\forall (u,v)\in\mathbb{R}^2,\quad \frac{\partial^2 g}{\partial u\partial v}(u,v)=0$$

On trouve alors $g(u,v)=\varphi(u)+\psi(v)$ avec $\varphi,\psi:\mathbb{R}\to\mathbb{R}$, ce qui conduit à :

$$f(x,t)=\varphi(x-ct)+\psi(x+ct)$$

II | Applications différentiables

On considère désormais une application $f:\mathscr{U}\subset E\to F$ où E et F désignent deux espaces vectoriels normés sur $\mathbb{R}$ de dimensions respectives p et n et $\mathscr{U}$ un ouvert de E. Toutes les normes étant équivalentes en dimension finie, le choix d'une norme spécifique pour E ou F importera peu.

On parle de champ de vecteurs lorsque $f:E\to E$ et de champ scalaire lorsque $f:E\to\mathbb{R}$.

Nous travaillerons bien souvent – à quelques exceptions matricielles près – dans les espaces $E=\mathbb{R}^p$ et $F=\mathbb{R}^n$ munis de leur structure euclidienne canonique. L'on pourra ainsi noter :

$$f(x)=f(x_1,\ldots,x_p)=(f_1(x),\ldots,f_n(x))$$

Cette identification est toujours possible moyennant le choix d'une base de E et d'une base de F :

$$x=\sum_{j=1}^{p}x_je_j\quad\text{et}\quad f(x)=\sum_{i=1}^{n}f_i(x)e'_i$$

- On appelle applications partielles de f les applications : $x_j\mapsto f(x)$ pour $j\in[\![1,p]\!]$.
- On appelle applications composantes de f les applications $f_i : x\mapsto f_i(x)$ pour $i\in[\![1,n]\!]$.

Exemple

Soit $f:\left|\begin{array}{ccc}\mathbb{R}^2 & \longrightarrow & \mathbb{R}^3\\ (u,v) & \longmapsto & (\cos(u+v),u-v,u^2\mathrm{e}^v)\end{array}\right.$

- f possède deux applications partielles définies sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}^3$:

$$u\mapsto(\cos(u+v),u-v,u^2\mathrm{e}^v)\quad\text{et}\quad v\mapsto(\cos(u+v),u-v,u^2\mathrm{e}^v)$$

- f possède trois composantes définies sur $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$:

$$(u,v)\mapsto\cos(u+v),\quad (u,v)\mapsto u-v\quad\text{et}\quad (u,v)\mapsto u^2\mathrm{e}^v$$

A – Différentielle

Théorème / Définition 20.13 : Différentielle en un point

Une application $f : \mathscr{U} \subset E \to F$ est dite différentiable en $a \in \mathscr{U}$ s'il existe $\varphi \in \mathscr{L}(E,F)$ tel que :

$$f(a+h) \underset{h\to 0}{=} f(a) + \varphi(h) + \mathrm{o}(h)$$

Si l'application φ existe, elle est unique et est alors appelée *différentielle* de f au point a, mais aussi *application linéaire tangente* à f en a. On la note $\mathrm{d}f_a$ ou $\mathrm{d}f(a)$.

L'existence d'un développement limité à l'ordre 1 relève ainsi de la définition même de la différentiabilité :

$$f(a+h) \underset{h\to 0}{=} f(a) + \mathrm{d}f_a(h) + \mathrm{o}(h) \underset{h\to 0}{=} f(a) + \big[\mathrm{d}f(a)\big](h) + \mathrm{o}(h)$$

Au final, une fonction différentiable en un point est une fonction qui vérifie notre grand principe :

$$\begin{pmatrix}\text{accroissement}\\ \text{de la fonction}\end{pmatrix} = \begin{pmatrix}\text{terme linéaire par rapport}\\ \text{à l'accroissement de la variable}\end{pmatrix} + \begin{pmatrix}\text{terme}\\ \text{correctif}\end{pmatrix}$$

Il arrive parfois que l'on écrive $\mathrm{d}f_a(h) = \big[\mathrm{d}f(a)\big](h) = \mathrm{d}f(a)\cdot h$. Cette dernière notation fait sens quand on raisonne en termes matriciels : c'est le produit d'une matrice de taille $n \times p$ par un vecteur colonne de taille p.

Rappelons enfin que par définition, $\mathrm{o}(h) = \|h\|\varepsilon(h)$ où $\varepsilon : E \to F$ avec $\varepsilon(h) \xrightarrow[h\to 0_E]{} 0_F$.

Exemples

- Si $f : \mathbb{R} \to \mathbb{R}$ est dérivable en a alors f est différentiable en a et $\mathrm{d}f_a : h \mapsto f'(a)h$.
- Si $f : \mathscr{U} \subset E \to F$ est constante sur $\mathscr{U}$ alors, pour tout $a \in \mathscr{U}$, $\mathrm{d}f_a = 0_{\mathscr{L}(E,F)}$.
- Si $f : \mathscr{U} \subset E \to F$ est linéaire alors, pour tout $a \in \mathscr{U}$, $\mathrm{d}f_a = f$.

Proposition 20.14

Si $f : \mathscr{U} \subset E \to F$ est différentiable en $a \in \mathscr{U}$, alors f est continue en a.

Démonstration

Si f est différentiable en a, $f(a+h) - f(a) \underset{h\to 0}{=} \mathrm{d}f_a(h) + \mathrm{o}(h) \xrightarrow[h\to 0_E]{} \mathrm{d}f_a(0_E) = 0_F$, par continuité de $\mathrm{d}f_a$ en tant qu'application linéaire définie sur un espace normé de dimension finie. ■

Proposition 20.15 : Opérations algébriques

- Si $f, g : \mathscr{U} \subset E \to F$ sont différentiables en a, alors pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ est différentiable en a et :

$$\mathrm{d}(\lambda f + \mu g)_a = \lambda \mathrm{d}f_a + \mu \mathrm{d}g_a$$

- Si $f : \mathscr{U} \subset E \to F$ et $g : \mathscr{V} \subset F \to G$ avec $\mathscr{U}$ et $\mathscr{V}$ deux ouverts vérifiant $f(\mathscr{U}) \subset \mathscr{V}$ et E, F et G trois espaces normés de dimension finie. Si f est différentiable en a et g différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et :

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

Démonstration

- On montre que pour toute application linéaire φ, $\mathrm{o}(\varphi(h)) = \mathrm{o}(h)$ par lipschitzianité de φ et $\varphi(\mathrm{o}(h)) = \mathrm{o}(h)$.
- Prouvons la deuxième assertion. Par différentiabilité de f en a et de g en $f(a)$,

$$(g\circ f)(a+h) \underset{h\to 0}{=} g(f(a) + \underbrace{\mathrm{d}f_a(h) + \mathrm{o}(h)}_{=k}) = g(f(a)) + \mathrm{d}g_{f(a)}(k) + \mathrm{o}(k) = g(f(a)) + \big[\mathrm{d}g_{f(a)} \circ \mathrm{d}f_a\big](h) + \mathrm{o}(h)$$

Par unicité de la différentielle, $g\circ f$ est différentiable en a et $\mathrm{d}(g\circ f)_a = \mathrm{d}g_{f(a)} \circ \mathrm{d}f_a$. ■

Exercice 3

Soient $f, g : \mathscr{U} \subset E \to \mathbb{R}$ différentiables en a. Montrer que $f \times g$ est différentiable en a et déterminer la différentielle en ce point.

Définition 20.16 : Différentielle et application de classe $\mathscr{C}^1$

- L'application $f : \mathscr{U} \subset E \to F$ est dite différentiable sur $\mathscr{U}$ si f est différentiable en tout point de $\mathscr{U}$. On appelle alors différentielle de f l'application :

$$\mathrm{d}f : \left| \begin{array}{ccl} \mathscr{U} & \longrightarrow & \mathscr{L}(E,F) \\ a & \longmapsto & \mathrm{d}f(a) = \mathrm{d}f_a \end{array} \right.$$

- L'application $f : \mathscr{U} \subset E \to F$ est dite de classe $\mathscr{C}^1$ sur $\mathscr{U}$ si f est différentiable sur $\mathscr{U}$ et si sa différentielle $\mathrm{d}f$ est continue sur $\mathscr{U}$.

Si f est de classe $\mathscr{C}^1$ sur $\mathscr{U}$, on dira aussi que f est continûment différentiable sur $\mathscr{U}$.

Exemple

Si $f : E \to F$ est linéaire, f est différentiable et sa différentielle est constante : $\quad \forall a \in E, \quad \mathrm{d}f(a) = f$.

Exercice 4

Montrer que l'application $(x, y) \mapsto x^2 - y^2$ définie sur $\mathbb{R}^2$ est différentiable en tout point $(a, b) \in \mathbb{R}^2$ et calculer sa différentielle en (a, b). Faire alors le lien avec les dérivées partielles de f en (a, b).

B – Dérivée selon un vecteur et dérivées partielles

Définition 20.17 : Dérivée selon un vecteur

Soit $f : \mathscr{U} \subset E \to F$. Si la fonction de la variable réelle $\varphi : t \mapsto f(a + tu)$ avec $u \in E$ est dérivable en 0, on dit alors que f est dérivable en a selon le vecteur u et on pose :

$$D_u(f)(a) = \lim_{t \to 0} \frac{f(a + tu) - f(a)}{t}$$

Quand une fonction est différentiable, elle est dérivable dans toutes les directions.

Proposition 20.18

Si f est différentiable en a alors f est dérivable en a selon u pour tout vecteur $u \in E$ et $D_u(f)(a) = \mathrm{d}f_a(u)$.

Si E est muni d'une base $(e_1, \dots, e_p)$, on peut définir les dérivées partielles de f qui sont des dérivées selon des directions privilégiées.

Définition 20.19 : Dérivées partielles

Soit $f : \mathscr{U} \subset E \to F$. On suppose E muni d'une base $(e_1, \dots, e_p)$. Pour $j \in [\![1, p]\!]$, on appelle, sous réserve d'existence, dérivée partielle en a d'indice j la dérivée de f en a suivant e_j, c'est-à-dire :

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t} = \lim_{t \to 0} \frac{f(a_1, \dots, a_{j-1}, a_i + t, a_{j+1}, \dots, a_p) - f(a)}{t}$$

Si f est différentiable en a, alors les dérivées partielles existent et :

$$\forall j \in [\![1, p]\!], \quad \frac{\partial f}{\partial x_j}(a) = \mathrm{d}f_a(e_j)$$

Remarquons alors que par linéarité de $\mathrm{d}f_a$, $\mathrm{d}f_a(h) = \mathrm{d}f_a\left(\sum_{j=1}^{p} h_j e_j\right) = \sum_{j=1}^{p} h_j \mathrm{d}f_a(e_j) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a)$.

En notant $\mathrm{d}x_j$ les applications $h \mapsto h_j$,

$$\mathrm{d}f_a = \frac{\partial f}{\partial x_1}(a)\mathrm{d}x_1 + \dots + \frac{\partial f}{\partial x_p}(a)\mathrm{d}x_p = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a)\mathrm{d}x_j$$

Théorème 20.20 : Caractérisation des applications de classe $\mathscr{C}^1$

Soit $f : \mathscr{U} \subset E \to F$. f est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ si et seulement si les dérivées partielles de f existent et sont continues en tout point de $\mathscr{U}$.

Ce résultat (que l'on admet) est fondamental. Il permet de justifier facilement la différentiabilité d'une fonction, à la seule condition d'existence et de continuité des dérivées partielles (on obtient même en prime la continuité de la différentielle). L'expression de la différentielle à partir des dérivées partielles (cf. infra) en découle.

Si f est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ et $a \in \mathscr{U}$, alors :

$$f(a+h) \underset{h\to 0}{=} f(a) + \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a) + \mathrm{o}(h)$$

Exercice 5

Montrer que l'application $(x, y) \mapsto x^2 - y^2$ définie sur $\mathbb{R}^2$ est de classe $\mathscr{C}^1$ et déterminer sa différentielle.

Attention, les notations sont trompeuses ! $\dfrac{\partial f}{\partial y}$ ne désigne pas la dérivée d'une expression (ou d'une « grandeur ») par rapport à y mais la dérivée d'une fonction par rapport à sa seconde variable. C'est une simple convention. On note parfois cette dérivée partielle $\partial_2 f$ pour alléger les calculs et lever toute ambiguïté.

C – Représentation matricielle de la différentielle

On munit de nouveau nos espaces E et F des bases $(e_j)_{1\leqslant j\leqslant p}$ et $(e'_i)_{1\leqslant i\leqslant n}$. Pour toute fonction $f : E \to F$, on peut alors écrire, moyennant une petite identification :

$$f(x) = f(x_1, \dots, x_p) = (f_1(x), \dots, f_n(x))$$

On suppose f de classe $\mathscr{C}^1$. Les fonctions $\dfrac{\partial f_i}{\partial x_j}$ sont alors définies et continues pour tout $(i, j) \in [\![1, n]\!] \times [\![1, p]\!]$. Ce sont des fonctions définies sur $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}$.

Théorème / Définition 20.21 : Jacobienne

La matrice représentative de $\mathrm{d}f_x$ dans les bases $(e_j)_{1\leqslant j\leqslant p}$ et $(e'_i)_{1\leqslant i\leqslant n}$ est $\left(\dfrac{\partial f_i}{\partial x_j}(x)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$.

Elle est appelée matrice jacobienne[1] de f au point $x = (x_1, \dots, x_p)$. On la note généralement $J_f(x)$.

Ainsi, pour $f : \mathscr{U} \to F$ de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U} \subset E$ et $x \in \mathscr{U}$, la jacobienne de f au point x est :

$$J_f(x) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(x) & \dots & \dfrac{\partial f_1}{\partial x_p}(x) \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(x) & \dots & \dfrac{\partial f_n}{\partial x_p}(x) \end{pmatrix}$$

Exemple

Soit $f : (x, y, z) \mapsto (2x + yz, \cos(y) + x^2)$. L'application est de classe $\mathscr{C}^1$ sur $\mathbb{R}^3$ donc $\mathrm{d}f$ existe en tout point de $\mathbb{R}^3$. La jacobienne au point $(1, 0, 2)$ est :

$$J_f(1,0,2) = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

Ainsi, $\mathrm{d}f_{(1,0,2)} : \left| \begin{array}{rcl} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x, y, z) & \longmapsto & (2x + 2y, 2x) \end{array} \right.$

1. *Charles Gustave Jacob Jacobi* (1804-1851), mathématicien allemand qui a notamment travaillé sur les équations aux dérivées partielles et la théorie des nombres.

Soient $f : E \to F$ et $g : F \to G$, de classe $\mathscr{C}^1$ respectivement sur E et F. $g \circ f : E \to F$ est de classe $\mathscr{C}^1$ sur E et :

$$\forall x \in E \quad J_{g\circ f}(x) = J_g(f(x)) \times J_f(x) \quad \text{puisque} \quad \mathrm{d}(g \circ f)_x = \mathrm{d}g_{f(x)} \circ \mathrm{d}f_x$$

Ce produit matriciel nous permet de retrouver la fameuse règle de la chaîne (essayez!).

D – Gradient associé à une fonction numérique

On considère dans ce paragraphe uniquement des fonctions à valeurs dans $\mathbb{R}$.

1 – Définition

Soit donc $f : \mathscr{U} \subset E \to \mathbb{R}$ avec bien souvent $E = \mathbb{R}^p$. On suppose f de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$. D'après ce qui précède, en tout point $a \in \mathscr{U}$, $\mathrm{d}f_a$ est une forme linéaire et, dans la base canonique $\mathscr{B}$ de $\mathbb{R}^p$,

$$\mathrm{d}f_a \in \mathscr{L}(\mathbb{R}^p;\mathbb{R}) \quad \text{et} \quad J_f(a) = \mathrm{Mat}_{\mathscr{B}}(\mathrm{d}f_a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) & \cdots & \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}$$

On peut alors définir le gradient de f au point a au moyen de ses coordonnées dans la base $\mathscr{B}$ en posant :

$$\nabla f(a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}$$

En notant $\cdot$ le produit scalaire usuel de $\mathbb{R}^p$, pour tout $h \in \mathbb{R}^p$,

$$\nabla f(a) \cdot h = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a) = \sum_{j=1}^{p} h_j \mathrm{d}f_a(e_j) = \mathrm{d}f_a(h)$$

La différentiabilité de f au point a se traduit alors par l'égalité :

$$f(a+h) \underset{h\to 0}{=} f(a) + \nabla f(a) \cdot h + \mathrm{o}(h)$$

Il est également possible de définir le gradient de manière intrinsèque, sans l'usage des coordonnées et donc d'une base de E. Pour cela, démontrons le théorème suivant dit de représentation de Riesz.

Théorème 20.22 : Représentation des formes linéaires

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien. Pour tout forme linéaire φ, il existe un unique vecteur $u \in E$ tel que :

$$\forall x \in E, \quad \varphi(x) = \langle u, x \rangle$$

Démonstration

Pour tout $u \in E$, $x \mapsto \langle u, x \rangle$ est une bien forme linéaire. Justifions par ailleurs que l'application suivante est un isomorphisme entre E et $E^* = \mathscr{L}(E, \mathbb{R})$, appelé *dual* de E :

$$\xi : \left| \begin{array}{ccc} E & \longrightarrow & E^* \\ u & \longmapsto & [x \mapsto \langle u, x \rangle] \end{array} \right.$$

- La linéarité de ξ n'est pas excessivement difficile à obtenir.
- Soit $u \in \mathrm{Ker}(\xi)$. Cela signifie que $x \mapsto \langle u, x \rangle$ est l'application nulle de E.
  En évaluant en u, on obtient $\|u\|^2 = 0$ donc $u = 0_E$.
- E et E^* ont même dimension.

La bijectivité de ξ en découle. La surjectivité de ξ permet alors de conclure. ∎

En dimension finie, toute forme linéaire est donc issue d'un produit scalaire. Cela permet par exemple de montrer que toute forme linéaire sur $\mathscr{M}_n(\mathbb{R})$ est de la forme $M \mapsto \mathrm{Tr}(AM)$. Dans notre cas, elle sert à établir l'existence du gradient via la forme linéaire $\mathrm{d}f_a$, une fois E muni d'un produit scalaire.

Définition 20.23

Soit $f : \mathscr{U} \subset E \to \mathbb{R}$ différentiable en a. L'application $\mathrm{d}f_a$ étant une forme linéaire, il existe un vecteur appelé gradient de f en a – et noté $\nabla f(a)$ – tel que pour tout $h \in E$,

$$\mathrm{d}f_a(h) = \nabla f(a) \cdot h$$

2 – Gradient et ligne de niveaux

On suppose toujours $f : \mathscr{U} \subset E \to \mathbb{R}$ différentiable en a. Pour h proche de 0, l'inégalité de Cauchy-Schwarz nous permet d'écrire :

$$|f(a+h) - f(a)| \approx \left|\nabla f(a) \cdot h\right| \leqslant \|\nabla f(a)\| \times \|h\|$$

L'écart absolu $f(a+h) - f(a)$ est maximal dans le cas d'égalité $\left|\nabla f(a) \cdot h\right| = \|\nabla f(a)\| \times \|h\|$, c'est-à-dire lorsque la direction h est colinéaire au gradient. Comme annoncé lors de l'étude des fonctions de deux variables, le gradient indique bien la ligne de plus grande pente, là où les variations sont les plus importantes.

Ajoutons que si l'on suppose toute ligne de niveau décrite par un paramétrage γ de classe $\mathscr{C}^1$, $f \circ \gamma$ est constante. En dérivant, l'on obtient $\nabla f(\gamma(t)) \cdot \gamma'(t) = 0$, ce qui assure l'orthogonalité du gradient aux vecteurs qui dirigent les tangentes aux lignes de niveau (voir page 4).

3 – Recherche d'extrema

On généralise là encore des résultats obtenus pour les fonctions de deux variables à valeurs réelles.

Définition 20.24 : Point critique

Soit $f : \mathscr{U} \subset E \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$. On dit que $a \in \mathscr{U}$ est un point critique de f si :

$$\mathrm{d}f_a = 0_{\mathscr{L}(E,\mathbb{R})} \quad \text{c'est-à-dire} \quad \nabla f(a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Il n'est guère surprenant qu'un extremum soit atteint en un point de l'ouvert seulement là où les dérivées suivant toutes les directions vont s'annuler.

Théorème 20.25 : Condition nécessaire d'existence d'un extremum

Si $f : \mathscr{U} \subset E \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$ admet un extremum au point $a \in \mathscr{U}$ alors x est un point critique. Cela revient à dire que $\nabla f(a) = \vec{0}$.

La démonstration est identique au cas déjà étudié ($E = \mathbb{R}^2$).

E – Dérivées le long d'un arc et vecteurs tangents

Soient $f : \mathscr{U} \subset E \to F$ et $\gamma : I \subset \mathbb{R} \to E$ deux applications supposées de classe $\mathscr{C}^1$ respectivement sur l'espace normé E et l'intervalle I.

L'application composée $f \circ \gamma$ est elle-même définie sur I et à valeurs dans F. Géométriquement, c'est l'image de l'arc γ par la transformation f. Si $M(t_0)$ est régulier, c'est-à-dire si $\gamma'(t_0) \neq 0_E$, $\gamma'(t_0)$ dirige la tangente à la courbe γ en $M(t_0)$. Qu'en est-il pour la courbe $f \circ \gamma$?

Proposition 20.26 : Dérivation le long d'un arc

Si $f : \mathscr{U} \subset E \to F$ et $\gamma : I \subset \mathbb{R} \to E$ sont de classe $\mathscr{C}^1$, alors $f \circ \gamma$ est de classe $\mathscr{C}^1$ sur I et :

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

Démonstration

C'est le résultat d'une simple différentiation de composée d'applications différentiables :

$$\mathrm{d}(f\circ\gamma)_t = \mathrm{d}f_{\gamma(t)}\circ \mathrm{d}\gamma_t \quad \text{soit} \quad (f\circ\gamma)'(t) = \mathrm{d}f_{\gamma(t)}(\gamma'(t)) = \nabla f(\gamma(t))\cdot\gamma'(t)$$

■

On peut donc conclure. La tangente à la courbe $f\circ\gamma$ en t_0 est dirigée par $(f\circ\gamma)'(t_0) = \mathrm{d}f_{\gamma(t_0)}(\gamma'(t_0))$ si celui-ci est non nul. Ce dernier vecteur n'est rien d'autre que l'image du vecteur qui dirige la tangente à γ en t_0 par l'application linéaire $\mathrm{d}f_{\gamma(t_0)}$.

Proposition 20.27 : Intégration le long d'un arc

Si $f : \mathscr{U}\subset E\to F$ et $\gamma : [0,1]\subset\mathbb{R}\to E$ sont de classe $\mathscr{C}^1$ et si $\gamma(0)=a$ et $\gamma(1)=b$, alors :

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

Démonstration

La preuve est immédiate puisque l'on peut intégrer la fonction continue $(f\circ\gamma)'$ sur $[0,1]$:

$$\int_0^1 \mathrm{d}f_{\gamma(t)}(\gamma'(t))\,\mathrm{d}t = \int_0^1 (f\circ\gamma)'(t)\,\mathrm{d}t = f(\gamma(1)) - f(\gamma(0)) = f(b)-f(a)$$

■

On notera que le résultat ne dépend pas du chemin choisi.

Corollaire 20.28 : Caractérisation des fonctions constantes

Soient $\mathscr{U}$ un ouvert connexe par arcs et $f : \mathscr{U}\subset E\to F$. Alors,

$$f \text{ est constante sur } \mathscr{U} \text{ si et seulement si pour tout } a\in\mathscr{U},\ \mathrm{d}f_a = 0_{\mathscr{L}(E,F)}$$

Démonstration

L'implication est simple. La réciproque est obtenue grâce à la forme intégrale prouvée ci-dessus. Attention, cette dernière n'est valable que pour des chemins de classe $\mathscr{C}^1$. Conformément au programme, on restreindra la démonstration au cas où l'ouvert $\mathscr{U}$ est supposé convexe.

Soient $a, b\in\mathscr{U}$. On considère la fonction γ définie sur $[0,1]$ par $\gamma(t) = (1-t)a + tb$. γ est bien un chemin de $\mathscr{U}$ de classe $\mathscr{C}^1$. Par nullité de la différentielle,

$$f(b)-f(a) = \int_0^1 \mathrm{d}f_{\gamma(t)}(\gamma'(t))\,\mathrm{d}t = \int_0^1 0\,\mathrm{d}t = 0$$

Donc $f(a)=f(b)$. f est bien constante. ■

On généralise ainsi un résultat bien connu depuis (au moins) l'an dernier : la fonction $f\in\mathscr{C}^1(I,\mathbb{R})$ est constante sur l'intervalle I si et seulement si f' est nulle sur l'intervalle I. Rappelons que les parties de $\mathbb{R}$ connexes par arcs sont les intervalles.

On cherche maintenant à étendre la notion de vecteur tangent (introduite dans le cadre des arcs paramétrés) à celle de vecteur tangent à une partie d'un espace vectoriel. On passera là encore par des courbes : sera vecteur tangent à une partie X en un point x tout vecteur tangent à une courbe γ de X passant par ce point x.

Définition 20.29

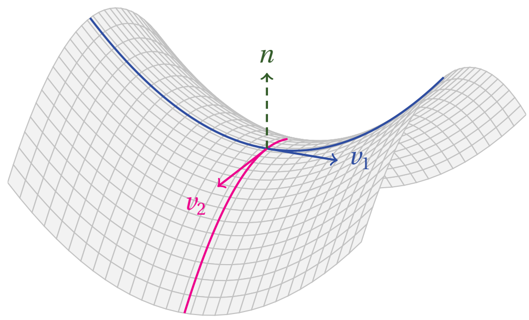
Si X est une partie de E et x un point de X, un vecteur v de E est tangent à X en x s'il existe $\varepsilon>0$ et un arc γ défini sur $]-\varepsilon,\varepsilon[$ dérivable en 0 à valeurs dans X, tels que $\gamma(0)=x, \gamma'(0)=v$.

Soit $\mathscr{S}$ la surface d'équation $z = f(x,y)$ où $f : \mathbb{R}^2 \to \mathbb{R}$ est supposée de classe $\mathscr{C}^1$. Notons alors $M_0(x_0, y_0, z_0)$ un point de $\mathscr{S}$ en lequel on cherche à déterminer le plan tangent.
On construit un arc à valeurs dans $\mathscr{S}$ en considérant deux fonctions $x, y : I \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur I vérifiant $x(t_0) = x_0$ et $y(t_0) = y_0$ pour un certain $t_0 \in I$ et l'application γ définie par :

$$\forall t \in I, \quad \gamma(t) = (x(t), y(t), f(x(t), y(t)))$$

Cette application γ est bien de classe $\mathscr{C}^1$ sur I, à valeurs dans $\mathscr{S}$ et vérifie $\gamma(t_0) = M_0$. En dérivant γ et en évaluant au point t_0, on obtient :

$$\gamma'(t_0) = \left(x'(t_0), y'(t_0), x'(t_0)\frac{\partial f}{\partial x}(x_0, y_0) + y'(t_0)\frac{\partial f}{\partial y}(x_0, y_0)\right)$$



Représentation de deux courbes tracées le long d'une surface et des vecteurs tangents associés.

Il ne reste plus qu'à vérifier que ce vecteur tangent à la courbe en M_0 est, quel que soit l'arc γ considéré, orthogonal au vecteur :

$$n = \begin{pmatrix} \dfrac{\partial f}{\partial x}(x_0, y_0) \\ \dfrac{\partial f}{\partial y}(x_0, y_0) \\ -1 \end{pmatrix}$$

Le plan tangent à $\mathscr{S}$ en M_0 peut alors être défini comme l'unique plan passant par M_0 et de vecteur normal n.
On retrouve l'équation déjà donnée par troncature à l'ordre 1 du développement limité de f en (x_0, y_0) :

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0)$$

Un bémol cependant : il faudrait vérifier que tout arc de classe $\mathscr{C}^1$ passant par M_0 admet un paramétrage de la forme $\gamma(t) = (x(t), y(t), f(x(t), y(t)))$. Il faudrait pour cela faire appel à un grand théorème du calcul différentiel : le théorème des fonctions implicites.

F – Applications de classe $\mathscr{C}^k$

Ce dernier paragraphe a seulement pour vocation de définir les dérivées partielles d'ordres supérieurs d'une fonction. Il suffit tout bonnement de... dériver les dérivées !
On notera, sous réserve d'existence, $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ ou $\partial_i \partial_j f$ la dérivée partielle d'indice i de la dérivée partielle $\dfrac{\partial f}{\partial x_j}$.
On définit plus généralement par récurrence les dérivées partielles d'ordre k pour tout $k \in \mathbb{N}^*$:

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \cdots \partial x_{i_1}} \right)$$

Définition 20.30 : Application de classe $\mathscr{C}^k$

Une application est dite de classe $\mathscr{C}^k$ sur un ouvert $\mathscr{U}$ si toutes ses dérivées partielles d'ordre k existent et sont continues.

Les sommes et composées de fonctions de classe $\mathscr{C}^k$ sont encore de classe $\mathscr{C}^k$, les fonctions polynomiales sont de classe $\mathscr{C}^\infty$...

Théorème 20.31 : Théorème de Schwarz

Soit $f : \mathscr{U} \subset E \to F$ une application de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$. Alors,

$$\forall (x_0, y_0) \in \mathscr{U}, \quad \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$