18 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

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 :

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

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

I | Premiers pas informels avec les fonctions numériques de deux variables

Débutons notre étude en la restreignant aux fonctions définies sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. Ces dernières peuvent en effet être représentées graphiquement et ainsi offrir des images mentales qui contribueront utilement à l'étude plus générale des fonctions de E dans F. Soit donc la fonction :

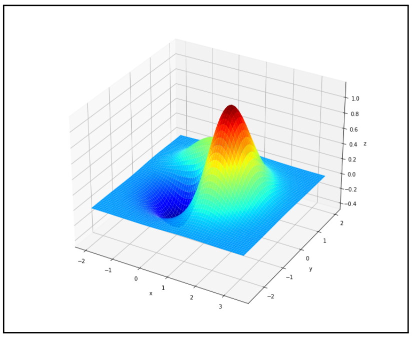
$$f : \left|\begin{array}{rcl} \mathscr{U} \subset \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array}\right.$$

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

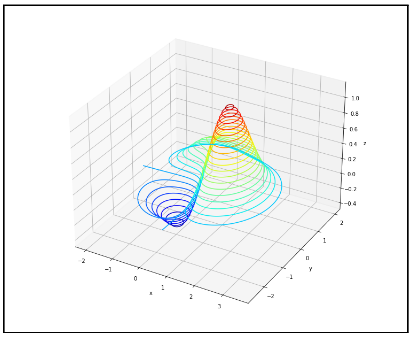
Définition 18.1 : Ligne de niveau

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

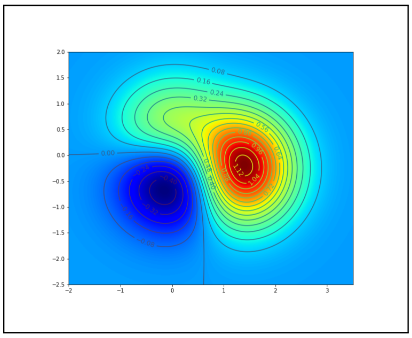
Les lignes de niveau, intersection de $\mathscr{S}$ et des plans d'équations $z = k$, permettent de visualiser la surface.



Exemple de surface



Lignes de niveau 3D



Lignes de niveau

A – Limites et continuité

On munit $\mathbb{R}^2$ de la norme euclidienne notée $\|\cdot\|$. Rappelons que f admet une limite $\ell \in \mathbb{R}$ en $(x_0, y_0) \in \mathscr{U}$ si :

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

Pour calculer une telle limite en (x_0, y_0), il ne suffit pas de faire tendre x vers x_0 *puis* y vers y_0 ou vice-versa.

Exemple

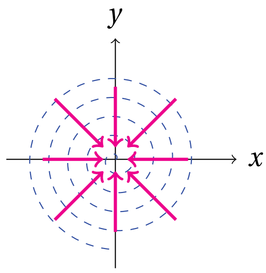
La quantité $f(x,y) = \dfrac{2xy}{x^2+y^2}$ n'admet pas de limite en $(0,0)$:

$$\lim_{y\to 0^*} \lim_{x \to 0^*} f(x,y) = \lim_{y \to 0^*} f(0,y) = 0 \quad \text{et} \quad \lim_{x\to 0^*} \lim_{y \to 0^*} f(x,y) = \lim_{y \to 0^*} f(x,0) = 0$$

et pourtant, $f(x,x) = 1 \xrightarrow[x\to 0^*]{} 1$. Plus généralement,

$$f(r\cos(\theta), r\sin(\theta)) = 2\cos(\theta)\sin(\theta) \xrightarrow[r\to 0^*]{} 2\cos(\theta)\sin(\theta)$$

L'existence de limites différentes selon la direction définie par l'angle θ choisi justifie la non-existence d'une limite en $(0,0)$.



Attention, l'existence d'une limite commune dans toutes les directions ne garantit pas l'existence d'une limite.

Exercice 1

Trouver les limites en $(0,0)$, lorsqu'elles existent, de $g(x,y) = \dfrac{x^2 y}{x^2+y^2}$ et $h(x,y) = \dfrac{x^2 y}{x^4+y^2}$.

Rappelons de plus que l'application f est dite continue en $(x_0,y_0) \in \mathscr{U}$ si $\displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$:

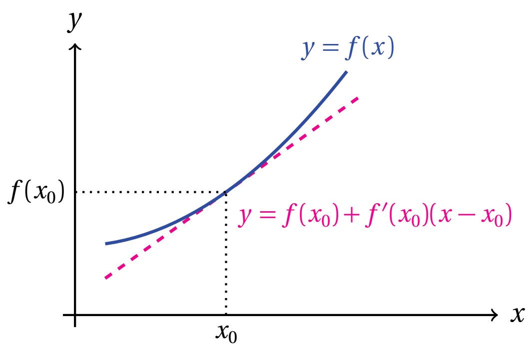
$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall (x,y) \in \mathscr{U}, \quad \|(x,y)-(x_0,y_0)\| < \delta \implies \left|f(x,y)-f(x_0,y_0)\right| < \varepsilon$$

On dit que f est continue sur $\mathscr{U}$ si f est continue en tout point de $\mathscr{U}$. La somme, le produit, la composée et le quotient dont le dénominateur ne s'annule pas de fonctions continues sont des fonctions continues.

Pour les mêmes raisons que celles évoquées pour les limites, 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 mais la réciproque est fausse.

B – Dérivées partielles premières

1 – Retour sur les fonctions de la 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 $\mathbb{R}^2$, à valeurs dans $\mathbb{R}$ et $(x_0, y_0) \in \mathscr{U}$.

Définition 18.2 : Dérivées partielles

Lorsque les limites suivantes existent et sont finies, on les appelle dérivées partielles premières de f au point $(x_0, y_0) \in \mathscr{U}$:

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

$\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$, également notées $\partial_1 f$ et $\partial_2 f$, sont, à l'image de f, des fonctions de deux variables à valeurs réelles.

Ces dérivées partielles ne sont que des cas particuliers de ce qu'on appelle les dérivées directionnelles de f.

Définition 18.3 : Dérivées directionnelles

Soit $u(a, b) \in \mathbb{R}^2$. Si la fonction de la variable réelle $t \mapsto f(x_0 + ta, y_0 + tb)$ est dérivable en 0, on dit alors que f est dérivable en (x_0, y_0) selon le vecteur u et on pose :

$$D_u(f)(x_0, y_0) = \lim_{t \to 0} \frac{f(x_0 + ta, y_0 + tb) - f(x_0, y_0)}{t}$$

Les dérivées partielles sont ainsi les dérivées directionnelles de f dans les directions privilégiées $(1, 0)$ et $(0, 1)$.

Définition 18.4 : Fonction de classe $\mathscr{C}^1$

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 $\mathscr{U}$.

Exemples

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

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

- L'application $\|\cdot\|$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

On admet provisoirement le résultat suivant.

Théorème 18.5 : 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 [1] en tout point (x_0, y_0, z_0) tel que $z_0 = f(x_0, y_0)$. Celui-ci a pour équation :

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

Cette équation est de la forme $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ avec $a = \dfrac{\partial f}{\partial x}(x_0, y_0)$, $b = \dfrac{\partial f}{\partial y}(x_0, y_0)$ et $c = -1$.

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

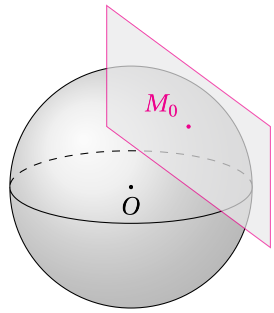
1. il sera défini dans un cadre plus général dans la section suivante

Exemple

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

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

La surface représentative de f est une demi-sphère. Déterminons une équation du plan tangent au point M_0 de coordonnées $(x_0,y_0,z_0)=(0,1/\sqrt{2},1/\sqrt{2})$.



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

f est tout d'abord de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}=\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2<1\}$ et pour tout $(x_0,y_0)\in\mathscr{U}$,

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

Ici, $x_0=0$ et $y_0=z_0=\frac{1}{\sqrt{2}}$ d'où l'équation recherchée du plan tangent en M_0 :

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

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) et on la note $\mathrm{d}f_{(x_0,y_0)}$. On observera que $\mathrm{d}f_{(x_0,y_0)}\in\mathscr{L}(\mathbb{R}^2,\mathbb{R})$.

Définition 18.6 : 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 $\nabla f(x_0,y_0)$.

Le gradient de f en (x_0,y_0) est donc l'unique vecteur vérifiant, pour tout $u=(h,k)$,

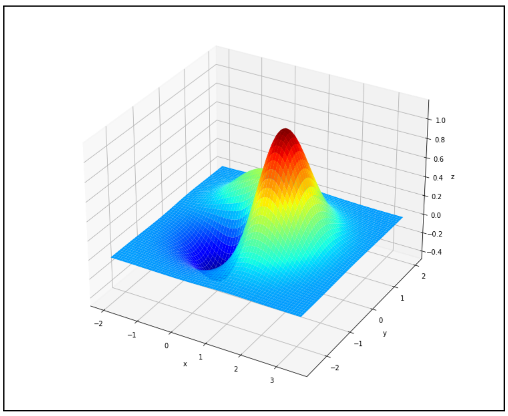
$$\mathrm{d}f_{(x_0,y_0)}(u)=\frac{\partial f}{\partial x}(x_0,y_0)\cdot h+\frac{\partial f}{\partial y}(x_0,y_0)\cdot k=\langle\nabla f(x_0,y_0),u\rangle$$

Quand ils manipulent des gradients, nos amis physiciens omettent parfois de préciser « (x_0,y_0) » mais il faut garder à l'esprit que l'on parle du gradient en un point donné.

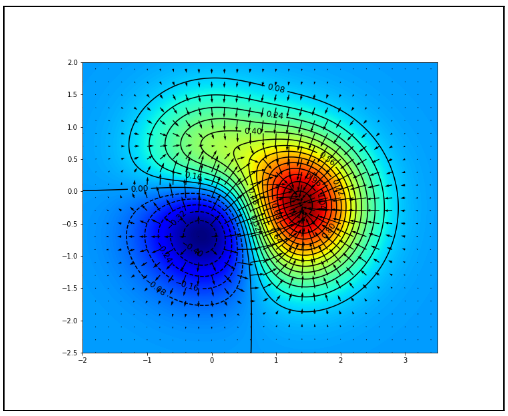
Proposition 18.7

Le gradient au point (x_0,y_0) est orthogonal à la ligne de niveau passant par (x_0,y_0).

Le gradient indique en outre la ligne de plus grande pente. Voici une illustration de ces deux propriétés pour une certaine fonction $f:\mathbb{R}^2\to\mathbb{R}$ de classe $\mathscr{C}^1$.



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



Gradient et lignes de niveau associés

3 – Notations différentielles

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 :

$$\mathrm{d}f_{(x_0,y_0)} = \frac{\partial f}{\partial x}(x_0,y_0)\,\mathrm{d}x + \frac{\partial f}{\partial y}(x_0,y_0)\,\mathrm{d}y \quad \text{(comb. lin. de deux applications linéaires)}$$

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 $\mathscr{L}(\mathbb{R}^2,\mathbb{R})$:

$$\mathrm{d}f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y \quad (\mathrm{d}f \text{ n'est en général pas linéaire})$$

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 $\mathscr{C}^1$:

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

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

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

Démonstration

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

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

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

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

■

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 18.8

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 18.9 : Théorème de Schwarz

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

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

Théorème 18.10 : 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 $(x_0, y_0) \in \mathscr{U}$:

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

Cette dernière formule rendra de précieux services dans l'étude des extrema d'une fonction numérique (cf. dernière section du chapitre).

D – 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.

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x}\cdot\frac{\partial g}{\partial u} + \frac{\partial v}{\partial x}\cdot\frac{\partial g}{\partial v} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}$$

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

De même, on trouve $\dfrac{\partial^2 f}{\partial t^2} = c^2\left(\dfrac{\partial^2 g}{\partial u^2} - 2\dfrac{\partial^2 g}{\partial u \partial v} + \dfrac{\partial^2 g}{\partial v^2}\right)$

L'équation $(\star)$ devient : $\quad \forall (u,v) \in \mathbb{R}^2, \quad \dfrac{\partial^2 g}{\partial u \partial v}(u,v) = 0$

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, applications de classe $\mathscr{C}^1$

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 :

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

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

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

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

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

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

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

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

A – Différentielle

Théorème / Définition 18.11 : 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 :

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

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 18.12

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, $f(a+h)-f(a)\underset{h\to0}{=}\mathrm{d}f_a(h)+\mathrm{o}(\|h\|)\xrightarrow[h\to0_E]{}\mathrm{d}f_a(0_E)=0_F$, par continuité de $\mathrm{d}f_a$ en tant qu'application linéaire définie sur un espace normé de dimension finie. ■

Proposition 18.13 : 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 app. 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 $f(a)$,

$$(g\circ f)(a+h)\underset{h\to0}{=}g(f(a)+\underbrace{\mathrm{d}f_a(h)+\mathrm{o}(\|h\|)}_{=k})=g(f(a))+\mathrm{d}g_{f(a)}(k)+\mathrm{o}(k)=g(f(a))+\left[\mathrm{d}g_{f(a)}\circ\mathrm{d}f_a\right](h)+\mathrm{o}(\|h\|)$$

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.

(i) Montrer que $f\times g$ est différentiable en a et déterminer la différentielle en ce point.

(ii) On suppose de plus que $f(a)\neq0$. Montrer que $\frac{1}{f}$ est différentiable en a.

Définition 18.14 : 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 :

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

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

Exemple

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

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 – Dérivées selon un vecteur et dérivées partielles

Définition 18.15 : Dérivée selon un vecteur

Soient $f : \mathscr{U} \subset E \to F$ et $a \in E$. Si la fonction de la variable réelle $t \mapsto f(a+tu)$ 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 18.16

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 18.17 : 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 :

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

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 (abusivement) $\mathrm{d}x_j$ les applications $h \mapsto h_j$,

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

Ainsi, lorsque f est différentiable sur $\mathscr{U}$ et si $a \in \mathscr{U}$,

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

Théorème 18.18 : Caractérisation des applications de classe $\mathscr{C}^1$

Soient $f : \mathscr{U} \subset E \to F$ et $(e_1,\dots,e_p)$ une base de E. 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 est essentiel : 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).

Démonstration *(à passer en première lecture)*

On note $\|\cdot\|_F$ une norme sur F. Soit $(e_1,\dots,e_p)$ une base de E. L'équivalence des normes en dimension finie nous permet de choisir une norme particulière sur E, celle définie par $\|h\| = \sum_{j=1}^p |h_j|$ pour $h = \sum_{j=1}^p h_j e_j$.

On munit $\mathscr{L}(E,F)$ de la norme subordonnée $\|\cdot\|_{\text{op}}$ relative à $\|\cdot\|$ et $\|\cdot\|_F$.

$\boxed{\Longrightarrow}$ Supposons la différentielle de f continue en tout point de l'ouvert $\mathscr{U}$. L'existence des dérivées partielles ayant été établie, il reste à justifier leur continuité. Soit $a \in \mathscr{U}$. Pour tous $x \in \mathscr{U}$ et $j \in [\![1,p]\!]$,

$$\left\|\frac{\partial f}{\partial x_j}(x)-\frac{\partial f}{\partial x_j}(a)\right\|_F = \left\|\mathrm{d}f_x(e_j)-\mathrm{d}f_a(e_j)\right\|_F \leqslant \left\|\mathrm{d}f_x-\mathrm{d}f_a\right\|_{\text{op}}\cdot\|e_j\| \xrightarrow[x\to a]{} 0$$

$\boxed{\Longleftarrow}$ Supposons maintenant l'existence et la continuité des dérivées partielles. Il s'agit de prouver que f est différentiable en tout point $a \in \mathscr{U}$ et de justifier la continuité de la différentielle.

Soient $a \in \mathscr{U}$ et $h = \sum_{j=1}^{p} h_j e_j$ tel que $a+h \in \mathscr{U}$. Par dérivabilité de f suivant e_p,

$$f(a+h) = f\left(a+\sum_{j=1}^{p} h_j e_j\right) = f\left(a+\sum_{j=1}^{p-1} h_j e_j\right) + h_p \frac{\partial f}{\partial x_p}\left(a+\sum_{j=1}^{p-1} h_j e_j\right) + \mathrm{o}(\|h\|)$$

Par continuité de $\dfrac{\partial f}{\partial e_p}$ en a, on obtient $f(a+h) = f\left(a+\sum_{j=1}^{p-1} h_j e_j\right) + h_p\dfrac{\partial f}{\partial x_p}(a)+\mathrm{o}(\|h\|)$.

De proche en proche,

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

f est bien différentiable en a et $\mathrm{d}f_a(h) = \sum_{j=1}^{p} h_j \dfrac{\partial f}{\partial x_j}(a)$.

Justifions désormais la continuité de la différentielle. Soit $x \in \mathscr{U}$.

$$\forall h \in E, \quad \|\mathrm{d}f_x(h)-\mathrm{d}f_a(h)\|_F \leqslant \sum_{j=1}^{p} |h_j| \cdot \left\|\frac{\partial f}{\partial x_j}(x)-\frac{\partial f}{\partial x_j}(a)\right\|_F \leqslant \|h\| \cdot \max_{1\leqslant j\leqslant p} \left\|\frac{\partial f}{\partial x_j}(x)-\frac{\partial f}{\partial x_j}(a)\right\|_F$$

Ainsi, $\|\mathrm{d}f_x-\mathrm{d}f_a\|_{\text{op}} \leqslant \max_{1\leqslant j\leqslant p}\left\|\dfrac{\partial f}{\partial x_j}(x)-\dfrac{\partial f}{\partial x_j}(a)\right\|_F \xrightarrow[x\to a]{} 0$ par continuité des $\dfrac{\partial f}{\partial x_j}$. ■

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 en un point

On munit à 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 certaine 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 18.19 : 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 $\left(\dfrac{\partial f_i}{\partial x_j}(x)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} = \big(\partial_j f_i(x)\big)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$.

Elle est appelée matrice jacobienne[2] 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 :

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

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 :

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

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

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 d'une fonction numérique

On considère dans ce paragraphe uniquement des fonctions à valeurs dans $\mathbb{R}$ et on suppose E, espace vectoriel de dimension p, muni d'une structure euclidienne.

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 $\mathbb{R}^p$,

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

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

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

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

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

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

$$f(a+h) \underset{h\to 0}{=} f(a) + \langle \nabla f(a), h \rangle + \mathrm{o}(\|h\|)$$

Il est également possible de définir le gradient de manière intrinsèque, sans l'usage des coordonnées et donc d'une base de E, à l'aide du théorème de représentation de Riesz.

2. *Charles Gustave Jacob Jacobi* (1804-1851), mathématicien allemand dont les travaux ont notamment porté sur les équations aux dérivées partielles et la théorie des nombres.

Définition 18.20

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 $h \in E$,

$$\mathrm{d}f_a(h) = \langle \nabla f(a), h \rangle$$

Si $f : \mathscr{U} \subset E \to \mathbb{R}$ est différentiable en a, pour h proche de 0, l'inégalité de Cauchy-Schwarz donne :

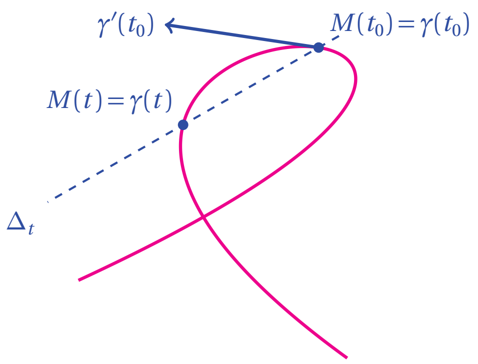
$$|f(a+h) - f(a)| \approx \left| \langle \nabla f(a), h \rangle \right| \leqslant \|\nabla f(a)\| \times \|h\|$$

L'écart absolu $f(a+h) - f(a)$ est maximal dans le cas d'égalité $\left| \langle \nabla f(a), h \rangle \right| = \|\nabla f(a)\| \times \|h\|$, 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, la direction où les variations sont les plus importantes.

E – Arcs paramétrés et dérivées le long d'un arc

On appelle arc paramétré de classe $\mathscr{C}^k$ toute fonction vectorielle de classe $\mathscr{C}^k$, définie sur un intervalle I de $\mathbb{R}$ et à valeurs dans E. Soit désormais un arc paramétré $\gamma : I \subset \mathbb{R} \to E$ supposé de classe $\mathscr{C}^1$.

Soit $t_0 \in I$. On suppose que $\gamma'(t_0) \neq 0$. Alors, la courbe représentative de γ admet une tangente au point $\gamma(t_0)$, dirigée par le vecteur $\gamma'(t_0)$.



Pour prouver ce résultat, considérons la droite Δ_t issue de $M(t_0)$, dirigée par le vecteur $\overrightarrow{M(t_0)M(t)}$, c'est-à-dire $\gamma(t) - \gamma(t_0)$, et faisons tendre t vers t_0 : intuitivement, la droite limite sera tangente à la courbe au point $M(t_0)$. Plus formellement, on définit la tangente à la courbe en $M(t_0)$ comme la droite passant par $M(t_0)$ et dirigée par $\displaystyle\lim_{t \to t_0} \frac{\overrightarrow{M(t_0)M(t)}}{\|\overrightarrow{M(t_0)M(t)}\|}$ lorsque cette dernière existe.

Sous les conditions précédentes, c'est le cas, car pour tout $t \neq t_0$,

$$\frac{\gamma(t) - \gamma(t_0)}{\|\gamma(t) - \gamma(t_0)\|} = \frac{\gamma(t) - \gamma(t_0)}{t - t_0} \cdot \frac{t - t_0}{\|\gamma(t) - \gamma(t_0)\|} \xrightarrow[t \to t_0]{} \frac{\pm\gamma'(t_0)}{\|\gamma'(t_0)\|}$$

On retrouve un résultat couramment utilisé en mécanique du point : le vecteur vitesse $\dfrac{\mathrm{d}\overrightarrow{OM}}{\mathrm{d}t}(t_0)$, s'il est non nul, dirige la tangente à la courbe en $M(t_0)$.

Soient désormais $f : \mathscr{U} \subset E \to F$ 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 18.21 : Dérivation le long d'un arc

Si $f : \mathscr{U} \subset E \to F$ et $\gamma : I \subset \mathbb{R} \to \mathscr{U}$ 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 :

$$\forall t \in I, \quad (f \circ \gamma)'(t) = \mathrm{d}(f \circ \gamma)_t(1) = (\mathrm{d}f_{\gamma(t)} \circ \mathrm{d}\gamma_t)(1) = \mathrm{d}f_{\gamma(t)}(\gamma'(t))$$

puisque que pour toute fonction $g : I \subset \mathbb{R} \to F$ dérivable, $g'(t) = \mathrm{d}g_t(1)$. ■

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 $\mathrm{d}f_{\gamma(t_0)}$.

Ce résultat appelle plusieurs remarques :

- Si $\gamma(t) = x + t\,u$ avec $x, u \in E$, alors γ est un paramétrage de la droite affine passant par x et dirigée par u. $\gamma'(t) = u$ est comme attendu constante et alors, $(f \circ \gamma)'(t) = \mathrm{d}f_{\gamma(t)}(u)$.
- Si $\gamma(t) = (x_1(t), \dots, x_p(t))$ est de classe $\mathscr{C}^1$, on généralise la règle de la chaîne :
$$(f \circ \gamma)'(t) = \mathrm{d}f_{\gamma(t)}(\gamma'(t)) = \sum_{j=1}^{p} x_j'(t) \frac{\partial f}{\partial x_j}(x_1(t), \dots, x_p(t))$$
- Enfin, si f est une application numérique, $(f \circ \gamma)'(t) = \mathrm{d}f_{\gamma(t)}(\gamma'(t)) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$.
  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 $\langle \nabla f(\gamma(t)), \gamma'(t) \rangle = 0$, ce qui assure l'orthogonalité du gradient aux vecteurs qui dirigent les tangentes aux lignes de niveau (cf. l'illustration page 4).

Proposition 18.22 : Intégration le long d'un arc

Si $f : \mathscr{U} \subset E \to F$ et $\gamma : [0,1] \subset \mathbb{R} \to \mathscr{U}$ 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 18.23 : Caractérisation des fonctions constantes

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

f est constante sur $\mathscr{U}$ 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 $\gamma(t) = (1-t)a + t\,b$. 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 : une 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.

F – Vecteurs tangents à une partie

On étend maintenant la notion de vecteur tangent aux parties d'un espace vectoriel de dimension finie.

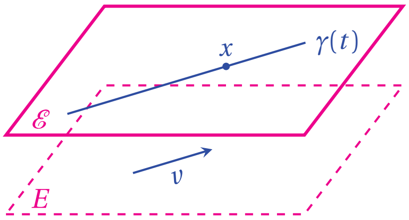
Définition 18.24 : Vecteur tangent

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 paramétré γ défini sur $]-\varepsilon, \varepsilon[$, dérivable en 0 et à valeurs dans X, tels que $\gamma(0) = x, \gamma'(0) = v$.
On note $\mathrm{T}_x X$ l'ensemble des vecteurs tangents à X en x.

Un tel ensemble $\mathrm{T}_x X$ n'est pas, en général, un espace vectoriel bien que les exemples suivants puissent suggérer le contraire. Pour déterminer les vecteurs tangents à une partie, on procède souvent par double inclusion.

Exemple

Si $\mathscr{E}$ est un sous-espace affine de dimension finie et de direction E, alors, en tout point $x \in \mathscr{E}$, $\mathrm{T}_x\mathscr{E} = E$.



- Soit $v \in E$. Il suffit de considérer la droite de $\mathscr{E}$ paramétrée par $\gamma(t) = x + t\,v$ pour vérifier que $v \in \mathrm{T}_x\mathscr{E}$.
- Réciproquement, soient $v \in \mathrm{T}_x\mathscr{E}$ et γ un arc de classe $\mathscr{C}^1$ défini sur $]-\varepsilon, \varepsilon[$, à valeurs dans $\mathscr{E}$ et tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

$$\forall t \in]-\varepsilon, \varepsilon[\setminus\{0\}, \quad \frac{\gamma(t) - \gamma(0)}{t} \in E$$

Par passage à la limite, E étant fermé en tant qu'espace vectoriel de dimension finie, $\gamma'(0) = v \in E$.

Exercice 6

Soit E un espace euclidien. Montrer que :

(i) pour tout point $x \in \mathscr{B}(0,1)$, $\mathrm{T}_x\mathscr{B}(0,1) = E$.

(ii) pour tout point $x \in \mathscr{S}(0,1)$, $\mathrm{T}_x\mathscr{S}(0,1) = \mathrm{Vect}(x)^{\perp}$.

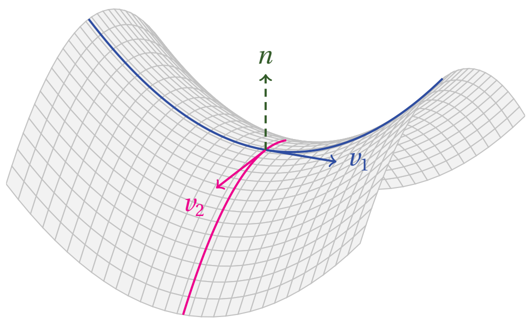
Concluons cette série d'exemples par l'étude du plan tangent (affine) à une surface de $\mathbb{R}^3$ d'équation $z = f(x,y)$ en un point donné. Soient $\mathscr{S}$ la surface d'équation $z = f(x,y)$ où $f : \mathbb{R}^2 \to \mathbb{R}$ est supposée de classe $\mathscr{C}^1$ et $M_0(x_0, y_0, z_0)$ un point de $\mathscr{S}$. Le plan tangent en M_0 peut être défini comme la réunion des tangentes aux courbes tracées le long de $\mathscr{S}$ passant par le point M_0, c'est-à-dire comme le plan affine passant par M_0 et de direction le plan vectoriel $\mathrm{T}_{M_0}\mathscr{S}$. Il reste à vérifier que $\mathrm{T}_{M_0}\mathscr{S}$ est, comme attendu, un plan vectoriel !

- On construit un arc sur $\mathscr{S}$ en considérant deux fonctions $x, y :]-\varepsilon, \varepsilon[\to \mathbb{R}$ de classe $\mathscr{C}^1$ vérifiant $x(0) = x_0$ et $y(0) = y_0$ et γ définie par :

$$\forall t \in]-\varepsilon, \varepsilon[, \quad \gamma(t) = (x(t), y(t), f(x(t), y(t)))$$

Cette application γ est de classe $\mathscr{C}^1$ sur $]-\varepsilon, \varepsilon[$, à valeurs dans $\mathscr{S}$ et vérifie $\gamma(0) = M_0$. En dérivant γ et en évaluant en 0, on obtient :

$$\gamma'(0) = \left(x'(0), y'(0), x'(0)\frac{\partial f}{\partial x}(x_0, y_0) + y'(0)\frac{\partial f}{\partial y}(x_0, y_0)\right)$$



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

Ce vecteur tangent à la courbe en M_0 est, quel que soit l'arc γ considéré, orthogonal au vecteur :

$$n = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1\right)$$

- Réciproquement, soit $v = (v_1, v_2, v_3)$ orthogonal à n. Il suffit de considérer l'arc γ défini par :

$$\forall t \in]-\varepsilon, \varepsilon[, \quad \gamma(t) = (x_0 + t\,v_1, y_0 + t\,v_2, f(x_0 + t\,v_1, y_0 + t\,v_2))$$

pour constater que $v \in \mathrm{T}_{M_0}\mathscr{S}$. En effet, γ est de classe $\mathscr{C}^1$ et à valeurs dans $\mathscr{S}$, $\gamma(0) = M_0$ et $\gamma'(0) = v$.

Nous venons ainsi de prouver que $\mathrm{T}_{M_0}\mathscr{S} = \mathrm{Vect}(n)^{\perp}$.

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) :

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

La description d'une surface de $\mathbb{R}^3$ par une équation de la forme $z = f(x, y)$ n'englobe cependant qu'un nombre restreint de surfaces de l'espace. On définit plus généralement une surface de $\mathbb{R}^3$ par la donnée d'une équation implicite $g(x, y, z) = 0$, comme par exemple $(x-x_0)^2+(y-y_0)^2+(z-z_0)^2 = R^2$. Il n'est pas toujours possible d'exprimer z en fonction de x et y pour se ramener au cas particulier précédent. Le théorème suivant permet de contourner cette objection pour déterminer le plan tangent à une telle surface. Sa démonstration (hors programme) fait appel au théorème des fonctions implicites, ce dernier précisant les conditions pour exprimer localement z comme une fonction de x et de y.

Théorème 18.25 : Hyperplan tangent

Soient g est une fonction numérique de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$, X l'ensemble des zéros de g et $x \in X$. Si $\mathrm{d}g_x$ est non nulle, $\mathrm{T}_x X = \mathrm{Ker}(\mathrm{d}g_x) = \nabla g(x)^\perp$.

Démonstration

Soient $g : \mathscr{U} \subset \mathbb{R}^p \to \mathbb{R}$ de classe $\mathscr{C}^1$, X l'hypersurface d'équation $g(x_1, \dots, x_p) = 0$ et $x \in X$ tel que $\mathrm{d}g_x \neq 0$.
Montrons que $\mathrm{T}_x X = \mathrm{Ker}(\mathrm{d}g_x) = \nabla g(x)^\perp$.
Remarquons pour commencer que $\mathrm{d}g_x \neq 0_{\mathscr{L}(\mathbb{R}^p,\mathbb{R})}$, $\mathrm{Ker}(\mathrm{d}g_x)$ est un hyperplan de $\mathbb{R}^p$. De plus,

$$h \in \mathrm{Ker}(\mathrm{d}g_x) \iff \mathrm{d}g_x(h) = 0 \iff \langle \nabla g(x), h\rangle = 0 \iff h \in \nabla g(x)^\perp$$

$\boxed{\subset}$ Soit $v \in \mathrm{T}_x X$. Il existe donc $\gamma :]-\varepsilon, \varepsilon[\to X$ de classe $\mathscr{C}^1$ tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

$$\forall t \in]-\varepsilon, \varepsilon[, \quad g(\gamma(t)) = 0 \quad \text{donc} \quad \mathrm{d}g_{\gamma(t)}(\gamma'(t)) = 0$$

En évaluant en 0, il vient $\mathrm{d}g_x(v) = 0$. Ainsi, $\mathrm{T}_x X \subset \mathrm{Ker}(\mathrm{d}g_x)$.

$\boxed{\supset}$ Admis. ■

Le plan tangent à une surface $\mathscr{S}$ d'équation $g(x, y, z) = 0$ en $M_0(x_0, y_0, z_0)$ est donc le plan affine passant par M_0 et de direction le plan vectoriel $\mathrm{T}_{M_0}\mathscr{S} = \nabla g(x_0, y_0, z_0)^\perp$. Il a donc pour équation :

$$\frac{\partial g}{\partial x}(x_0, y_0, z_0)\cdot(x-x_0) + \frac{\partial g}{\partial y}(x_0, y_0, z_0)\cdot(y-y_0) + \frac{\partial g}{\partial z}(x_0, y_0, z_0)\cdot(z-z_0) = 0 \ \text{ où } \ g(x_0, y_0, z_0) = 0$$

Exercice 7

Déterminer une équation du plan tangent à la sphère unité de $\mathbb{R}^3$, muni de sa structure euclidienne, en un point $M_0(x_0, y_0, z_0)$.

III | Applications de classe $\mathscr{C}^k$

Ce paragraphe a pour simple objectif 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 $\partial_{i,j} f$ 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 $k \in \mathbb{N}^*$:

$$\frac{\partial^k f}{\partial x_{j_k} \cdots \partial x_{j_1}} = \frac{\partial}{\partial x_{j_k}}\left(\frac{\partial^{k-1} f}{\partial x_{j_{k-1}} \cdots \partial x_{j_1}}\right)$$

Définition 18.26 : 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 18.27 : Théorème de Schwarz

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

$$\forall i, j \in [\![1, p]\!], \quad \forall x \in \mathscr{U}, \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

Ce théorème, que l'on admet et qui s'étend aux dérivées d'ordres supérieurs, montre que l'ordre de dérivation importe peu, du moment que la fonction est suffisamment régulière.

Exercice 8

(i) Justifier l'existence et calculer les dérivées partielles secondes sur $\mathbb{R}^2$ de $f : (x, y) \mapsto x\sin(y) + x^2$.

(ii) Comparer $\dfrac{\partial^2 g}{\partial x \partial y}(0,0)$ et $\dfrac{\partial^2 g}{\partial y \partial x}(0,0)$ puis conclure, lorsque $g(x,y) = \begin{cases} xy \cdot \dfrac{x^2 - y^2}{x^2 + y^2} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) = (0,0) \end{cases}$

IV | Optimisation avec et sans contrainte

Dans cette section, $f : \mathscr{U} \subset E \to \mathbb{R}$ désignera une application de classe $\mathscr{C}^2$ définie sur un ouvert $\mathscr{U}$ d'un espace vectoriel de dimension n, souvent $\mathbb{R}^n$, et à valeurs dans $\mathbb{R}$.

A – Formule de Taylor-Young à l'ordre 2 pour une fonction numérique

Pour les besoins de notre étude, commençons par énoncer la formule de Taylor-Young à l'ordre 2. Cette dernière nous permet d'améliorer notre grand principe introductif de la façon suivante :

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

Théorème 18.28 : Formule de Taylor-Young à l'ordre 2

Si f est de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de E alors, pour tout $a \in \mathscr{U}$:

$$f(a+h) \underset{h\to 0}{=} f(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2} \sum_{1 \leqslant i,j \leqslant n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \mathrm{o}(\|h\|^2)$$

Démonstration *(à passer en première lecture)*

Soit $a \in \mathscr{U}$. $\mathscr{U}$ étant ouvert, il existe $r > 0$ tel que pour tous $h \in \mathscr{B}(a, r)$ et $t \in [0,1]$, $a + th \in \mathscr{U}$.

- On introduit la fonction de la variable réelle φ définie sur $[0,1]$ par $\varphi(t) = f(a + th)$. φ est de classe $\mathscr{C}^2$ comme composée de fonctions de classe $\mathscr{C}^2$ et :

$$\forall t \in [0,1], \quad \varphi'(t) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a+th) \quad \text{et} \quad \varphi''(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a+th)$$

  La formule de Taylor avec reste intégral à l'ordre 2 pour φ, i.e. $\varphi(1) = \varphi(0) + \varphi'(0) + \displaystyle\int_0^1 (1-t)\varphi''(t)\,\mathrm{d}t$, nous donne celle relative à f :

$$f(a+h) = f(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \int_0^1 (1-t) \left[\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a+th)\right] \mathrm{d}t$$

- Exploitons maintenant la continuité des dérivées partielles secondes :

$$\forall i,j\in[\![1,n]\!],\quad \frac{\partial^2 f}{\partial x_i\partial x_j}(a+th)=\frac{\partial^2 f}{\partial x_i\partial x_j}(a)+\varepsilon_{i,j}(th)\quad\text{où}\quad \varepsilon_{i,j}(u)\xrightarrow[u\to 0_E]{}0$$

La formule de Taylor avec reste intégral devient :

$$f(a+h)=f(a)+\sum_{j=1}^n h_j\frac{\partial f}{\partial x_j}(a)+\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(a)+\sum_{1\leqslant i,j\leqslant n}h_ih_j\int_0^1(1-t)\varepsilon_{i,j}(th)\,\mathrm{d}t$$

Il reste à prouver que $R(h)=\displaystyle\sum_{1\leqslant i,j\leqslant n}h_ih_j\int_0^1(1-t)\varepsilon_{i,j}(th)\,\mathrm{d}t\underset{h\to 0}{=}\mathrm{o}(\|h\|^2)$.

Pour cela, fixons $\varepsilon>0$. Il existe $\delta>0$ tel que pour tous $i,j\in[\![1,p]\!]$, si $\|u\|\leqslant\delta$ alors $|\varepsilon_{i,j}(u)|<\varepsilon$.
Soit $h\in E$ tel que $\|h\|\leqslant\delta$. Alors, pour tout $t\in[0,1]$, $\|th\|\leqslant\delta$ et donc :

$$\forall i,j\in[\![1,p]\!],\quad \left|\int_0^1(1-t)\varepsilon_{i,j}(th)\right|\leqslant\varepsilon\int_0^1(1-t)\,\mathrm{d}t=\frac{\varepsilon}{2}$$

Ainsi, chaque terme de la somme est un $\mathrm{o}(h_ih_j)$. De plus, $|h_ih_j|\leqslant\|h\|_\infty^2$ donc $R(h)\underset{h\to 0}{=}\mathrm{o}(\|h\|^2)$. ■

La formule de Taylor-Young à l'ordre 2 se réécrit sous la forme plus concise suivante :

$$f(a+h)\underset{h\to 0}{=}f(a)+\mathrm{d}f_a(h)+\frac{1}{2}\mathrm{d}^2f_a(h,h)+\mathrm{o}(\|h\|^2)$$

où l'application bilinéaire $\mathrm{d}^2f_a:(h,k)\mapsto\displaystyle\sum_{1\leqslant i,j\leqslant n}h_ik_j\frac{\partial^2 f}{\partial x_i\partial x_j}(a)$ est appelée différentielle seconde[3] de f en a.

Il importe de savoir réécrire la formule pour une fonction de deux variables (usage le plus fréquent) :

$$\begin{aligned}f(x_0+h_1,y_0+h_2)\underset{(h_1,h_2)\to(0,0)}{=}\;&f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)\cdot h_1+\frac{\partial f}{\partial y}(x_0,y_0)\cdot h_2\\&+\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_0,y_0)\cdot h_1^2+2\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)\cdot h_1h_2+\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\cdot h_2^2\right]+\mathrm{o}(h_1^2+h_2^2)\end{aligned}$$

f étant à valeurs dans $\mathbb{R}$, nous disposons d'une écriture matricielle de la formule de Taylor-Young à l'ordre 2 assez commode. Elle s'appuie sur les relations suivantes :

$$\mathrm{d}f_a(h)=\sum_{j=1}^n h_j\frac{\partial f}{\partial x_j}(a)=\langle\nabla f(a),h\rangle\quad\text{et}\quad \mathrm{d}^2f_a(h)=\sum_{1\leqslant i,j\leqslant n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(a)=\langle H_f(a)h,h\rangle$$

où l'on a noté $h=(h_j)_{1\leqslant j\leqslant n}$ et $H_f(a)=\left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}(a)\right)_{1\leqslant i,j\leqslant n}$

Définition 18.29 : Hessienne

La hessienne au point $a\in\mathscr{U}$ de la fonction numérique $f:\mathscr{U}\subset\mathbb{R}^n\to\mathbb{R}$ de classe $\mathscr{C}^2$ est la matrice :

$$H_f(a)=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(a)\right)=\begin{bmatrix}\dfrac{\partial^2 f}{\partial {x_1}^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n}(a)\\ \vdots & \ddots & \vdots\\ \dfrac{\partial^2 f}{\partial x_1\partial x_n}(a) & \cdots & \dfrac{\partial^2 f}{\partial {x_n}^2}(a)\end{bmatrix}\in\mathrm{S}_n(\mathbb{R})$$

3. l'étude de la différentielle seconde ne figure pas au programme

Notons que la hessienne est bien symétrique, dès lors que f est de classe $\mathscr{C}^2$, en vertu du théorème de Schwarz.

On retiendra la version matricielle suivante de la formule de Taylor-Young à l'ordre 2 pour une fonction de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de E. Pour tout $a \in \mathscr{U}$,

$$\boxed{f(a+h) \underset{h\to 0}{=} f(a) + \langle \nabla f(a), h\rangle + \frac{1}{2}\langle H_f(a)h, h\rangle + \mathrm{o}(\|h\|^2) \underset{h\to 0}{=} f(a) + \nabla f(a)^\top h + \frac{1}{2} h^\top H_f(a) h + \mathrm{o}(\|h\|^2)}$$

B – Étude des extrema libres

Par la suite, f désignera une application définie sur une partie A de E, espace vectoriel de dimension n, et à valeurs dans $\mathbb{R}$.

Définition 18.30 : Extremum

On dit que f admet un minimum local (resp. un maximum local) en $a \in A$ s'il existe un voisinage $\mathscr{U}$ de a tel que :

$$\forall x \in \mathscr{U}, \quad f(x) \geqslant f(a) \quad (\text{respectivement } f(x) \leqslant f(a))$$

Le minimum (resp. le maximum) est qualifié de global lorsque l'inégalité est valable sur A.

Rappelons que si A est une partie compacte de E, f admet nécessairement un minimum et un maximum. De plus, pour une fonction de la variable réelle f de classe $\mathscr{C}^1$ sur un intervalle I :

- si f atteint un extremum local en $a \in \mathring{I}$, $f'(a) = 0$;
- f' peut s'annuler sans que f atteigne un extremum (cf. $x \mapsto x^3$ sur $]-1,1[$) ;
- si I n'est pas ouvert, f peut atteindre un extremum sans que sa dérivée s'annule (cf. $x \mapsto x$ sur $[0,1]$).

Ajoutons que lorsque f est de classe $\mathscr{C}^2$:

- si f atteint un minimum local en $a \in \mathring{I}$, $f''(a) \geqslant 0$;
- si $f'(a) = 0$ et $f''(a) > 0$ pour $a \in \mathring{I}$, f atteint un minimum local.

La recherche des extrema ne diffère guère de celle que les lecteurs connaissent pour les fonctions de la variable réelle. On prendra garde au fait que la plupart des résultats énoncés ci-dessous sont valables uniquement sur des parties ouvertes. Il sera donc parfois nécessaire de dissocier la recherche des extrema de f à l'intérieur de A et le long de sa frontière.

1 – Condition du premier ordre

Définition 18.31 : Point critique

Soit f une fonction de classe $\mathscr{C}^1$ sur un **ouvert** $\mathscr{U}$ de E. On dit que $a \in \mathscr{U}$ est un point critique de f si :

$$\mathrm{d}f_a = 0_{\mathscr{L}(E,\mathbb{R})} \quad \text{soit} \quad \nabla f(a) = \overrightarrow{0}$$

Cela revient à dire que pour tout $j \in [\![1,n]\!]$, $\dfrac{\partial f}{\partial x_j}(a) = 0$.

Théorème 18.32 : Condition nécessaire d'existence d'un extremum

Si $f : \mathscr{U} \subset E \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ admet un extremum en a, alors a est un point critique.

Démonstration

Soit $j \in [\![1,n]\!]$. Supposons que f atteigne un extremum en $a \in \mathscr{U}$. Alors, l'application de la variable réelle $\varphi : t \mapsto f(a + t e_j)$ est de classe $\mathscr{C}^1$ au voisinage de 0 et admet un extremum en 0. Ainsi, $\varphi'(0) = \dfrac{\partial f}{\partial x_j}(a) = 0$. ■

Les extremums locaux sont donc à rechercher parmi les points critiques. Cependant, comme en dimension 1, la réciproque est fausse ! Tout point critique ne correspond pas nécessairement à un extremum local.

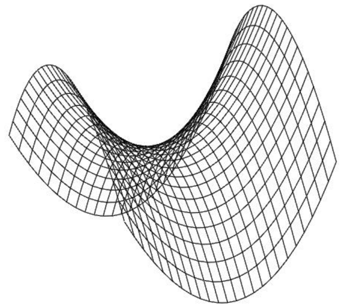
Exemple (Point selle)

Soit la fonction $f:(x,y)\mapsto x^2-y^2$; $f\in\mathscr{C}^1(\mathbb{R}^2,\mathbb{R})$ et :

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

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. Pour justifier que l'on dispose d'un extrema, on peut revenir à la définition d'un extremum.

Exemple

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

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

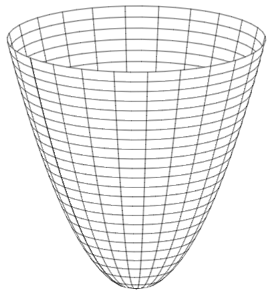
- *Recherche des points critiques*
  f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ car polynomiale.

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

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

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

  Ainsi, f atteint en M_0 un minimum (global) que l'on visualise bien sur le paraboloïde ci-dessus.



Présence d'un minimum

Exercice 9

Soient $(E,\langle\cdot,\cdot\rangle)$ un espace préhilbertien, F un sous-espace vectoriel de dimension finie de E et $a\in E$.

(i) Montrer que l'application $f:u\mapsto\|u-a\|$ admet un minimum sur F en calculant $\lim\limits_{\|u\|\to+\infty} f(u)$.

(ii) Justifier que ce minimum est atteint en un seul $x\in F$. Que dire de x ?

2 – Condition du second ordre

L'étude globale du signe de $f(x,y)-f(x_0,y_0)$ menée dans l'exemple précédent pour justifier la présence d'un minimum n'est pas satisfaisante car difficilement reproductible pour une fonction quelconque. Pour pallier cette difficulté, il suffit de pousser l'étude locale au voisinage d'un point critique au moyen de la formule de Taylor-Young. En effet, si f est de classe $\mathscr{C}^2$ sur l'ouvert $\mathscr{U}$ et a un point critique de f, alors :

$$f(a+h)\underset{h\to 0}{=}f(a)+\frac{1}{2}h^\top H_f(a)h+\mathrm{o}(\|h\|^2)$$

Le signe de $f(a+h)-f(a)$ est localement celui de $\langle H_f(a)h,h\rangle=h^\top H_f(a)h$, ce qui conduit au résultat suivant.

Théorème 18.33 : Condition suffisante d'existence d'un extremum

Soient $f:\mathscr{U}\subset E\to\mathbb{R}$ de classe $\mathscr{C}^2$ sur l'ouvert $\mathscr{U}$ et $a\in\mathscr{U}$.

- si f atteint en a un minimum local, $H_f(a)\in\mathrm{S}_n^+(\mathbb{R})$.
- si a est un point critique de f et si $H_f(a)\in\mathrm{S}_n^{++}(\mathbb{R})$, alors f atteint en a un minimum local.

Bien entendu, appliqué à $-f$, le théorème précédent affirme que lorsque a est un point critique de f et si $H_f(a)\in\mathrm{S}_n^{--}(\mathbb{R})$, alors f atteint en a un maximum local.

Démonstration

Soient $f : \mathscr{U} \subset E \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur l'ouvert $\mathscr{U}$ et $a \in \mathscr{U}$.

- Supposons que $f(a)$ est un minimum local de f. Alors a est un point critique de f. De plus, si $h \in E$, pour tout $t \in \mathbb{R}$ suffisamment petit, $a + th \in \mathscr{U}$ et $f(a+th) - f(a) \geqslant 0$.
  Or, $f(a+th) - f(a) \underset{t\to 0}{=} \frac{t^2}{2}\langle H_f(a)h, h\rangle + \mathrm{o}(t^2)$ c'est-à-dire, $\langle H_f(a)h, h\rangle \underset{t\to 0^*}{=} \frac{2}{t^2}\big(f(a+th) - f(a)\big) + \mathrm{o}(1)$.
  Un passage à la limite permet de conclure : $\langle H_f(a)h, h\rangle \geqslant 0$ pour tout $h \in E$.
- Supposons maintenant que a est un point critique de f et que $H_f(a) \in \mathrm{S}_n^{++}(\mathbb{R})$. Alors, en notant $\lambda_{\min}$ la plus petite des valeurs propres (nécessairement réelle) de $H_f(a)$, d'après le théorème spectral,

$$\forall h \in E \setminus \{0_E\}, \quad \langle H_f(a)h, h\rangle \geqslant \lambda_{\min}\|h\|^2 > 0$$

  $f(a+h) - f(a) = \frac{1}{2}h^\top H_f(a)h + \|h\|^2\varepsilon(h)$ avec $\varepsilon(u) \xrightarrow[u\to 0]{} 0$. Il existe donc $r > 0$ tel que pour tout $h \in \mathscr{B}(0, r)$, $\|h\|^2\varepsilon(h) \leqslant \lambda_{\min}\frac{\|h\|^2}{4}$. Ainsi, pour tout $h \in \mathscr{B}(0, r)$,

$$f(a+h) - f(a) = \frac{1}{2}h^\top H_f(a)h + \|h\|^2\varepsilon(h) \geqslant \lambda_{\min}\frac{\|h\|^2}{2} - \lambda_{\min}\frac{\|h\|^2}{4} = \lambda_{\min}\frac{\|h\|^2}{4} \geqslant 0$$

  $f(a)$ est bien un minimum local de f. ■

Corollaire 18.34

Soient $f : \mathscr{U} \subset E \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur l'ouvert $\mathscr{U}$ et $a \in \mathscr{U}$ un point critique de f.

- Si $\mathrm{Sp}(H_f(a)) \subset \mathbb{R}_+^*$, f atteint en a un minimum.
- Si $\mathrm{Sp}(H_f(a)) \subset \mathbb{R}_-^*$, f atteint en a un maximum.
- Si $H_f(a)$ possède deux valeurs propres de signes distincts, a est un point selle.

Lorsque les valeurs propres $H_f(a)$ sont seulement supposées positives (ou négatives), on ne peut pas conclure.

On notera que pour $n = 2$, on dispose d'une caractérisation simple au moyen du déterminant et de la trace. En effet, en notant λ et μ les deux valeurs propres de $H_f(a)$, $\det(H_f(a)) = \lambda\mu$ et $\mathrm{Tr}(H_f(a)) = \lambda + \mu$.

Corollaire 18.35

Soient $f : \mathscr{U} \subset \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur l'ouvert $\mathscr{U}$ et $a \in \mathscr{U}$ un point critique de f.

- Si $\det(H_f(a)) > 0$, f admet un extremum en a.
  Il s'agit d'un minimum lorsque $\mathrm{Tr}(H_f(a)) > 0$, d'un maximum sinon.
- Si $\det(H_f(a)) < 0$, f admet un point selle.
- Si $\det(H_f(a)) = 0$, on ne peut pas conclure.

Pour le dernier cas, on pourra penser aux fonctions $x \mapsto x^3$ et $x \mapsto x^4$ de dérivée et dérivée seconde nulles en 0.

Exemple

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

$$f(x, y) = x^4 + y^4 - (x-y)^2$$

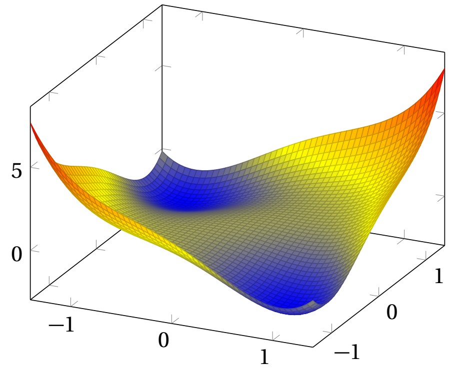
- La fonction f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^2$ en tant que fonction polynomiale.
- *Recherche des points critiques*

$$\nabla f(x, y) = \overrightarrow{0} \iff \begin{cases} 4x^3 - 2(x-y) = 0 \\ 4y^3 + 2(x-y) = 0 \end{cases} \iff \begin{cases} x^3 = -y^3 \\ 4x^3 - 2(x-y) = 0 \end{cases} \iff \begin{cases} x = -y \\ x^3 - x = 0 \end{cases}$$

  Nous avons donc trois points criques : $A(1, -1)$, $B(-1, 1)$ et $C(0, 0)$.

- *Détermination des extrema*

La hessienne en un point critique (x, y) est $H_f(x, y) = \begin{bmatrix} 12x^2 - 2 & 2 \\ 2 & 12y^2 - 2 \end{bmatrix}$.



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

⋆ Extremum au point $A(1, -1)$?
$\det(H_f(A)) = 96 > 0$ et $\mathrm{Tr}(H_f(A)) = 20 > 0$ donc f présente un minimum local en $(1, -1)$ qui vaut $f(1, -1) = -2$.

⋆ Extremum au point $B(-1, 1)$?
Par symétrie, f présente un minimum local en $(-1, 1)$.

⋆ Extremum au point $C(0, 0)$?
$\det(H_f(C)) = 0$, on ne peut pas conclure directement.
On remarquera cependant que $f(0, 0) = 0$ et que :

$$\begin{aligned} & f(x, x) = 2x^4 > 0 \text{ pour } x \neq 0 \\ \text{et} \quad & f(x, 0) = x^2(x^2 - 1) < 0 \text{ pour } x \in]-1, 1[\end{aligned}$$

On trouve ici un point selle.

C – Étude des extrema liés

On cherche dans cette dernière partie à déterminer les extrema d'une fonction $f : \mathscr{U} \subset E \to \mathbb{R}$ de classe $\mathscr{C}^1$ sous la contrainte $g(x) = 0$, où g est elle-même une fonction numérique de classe $\mathscr{C}^1$ sur $\mathscr{U}$.

Pour déterminer $\min\limits_{g(x)=0} f(x)$, une approche naïve consiste à tirer de la contrainte $g(x_1, \ldots, x_n) = 0$ une relation de la forme $x_n = \varphi(x_1, \ldots, x_{n-1})$. On se ramène de la sorte aux techniques d'optimisation du paragraphe précédent :

$$\min_{g(x)=0} f(x) = \min_{(x_1,\ldots,x_{n-1}) \in \mathscr{U}'} f(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_{n-1}))$$

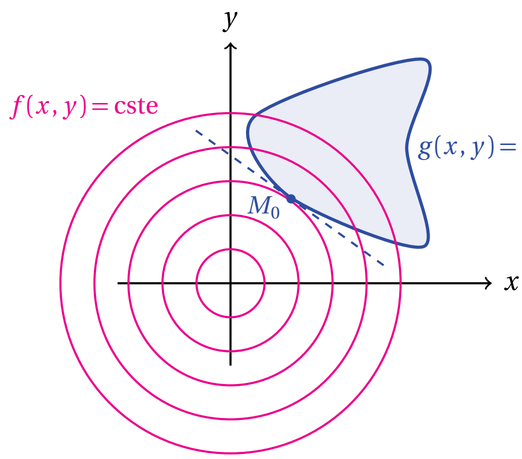
où $\mathscr{U}'$ serait une partie de E à préciser.

Exercice 10

Déterminer de deux manières $\min\limits_{x^2+y^2=1} xy$.

Néanmoins, même si des résultats spécifiques justifient, sous la condition $\mathrm{d}g_x \neq 0$, la possibilité d'exprimer une des variables en fonction des autres, l'approche proposée s'avère en pratique rapidement inopérante : on ne dispose pas en général d'une expression explicite et cette dernière n'est souvent valable que localement.

La mathématicien Joseph-Louis Lagrange a proposé une autre démarche, fondée sur le point de vue géométrique suivant. Considérons la fonction f définie sur $\mathbb{R}^2$ par $f(x, y) = x^2 + y^2$ et une contrainte de la forme $g(x, y) = 0$ où $g : \mathbb{R}^2 \to \mathbb{R}$ est supposée de classe $\mathscr{C}^1$. f atteint un minimum (global) en $(0, 0)$ mais celui-ci différera de $\min\limits_{g(x,y)=0} f(x, y)$ si $g(0, 0) \neq 0$.



Comme l'illustre le graphe ci-contre, lorsque f admet en $M_0 = (x_0, y_0)$ un minimum sous la contrainte $g(x_0, y_0) = 0$, la ligne de niveau relative à f en M_0 est tangente à la courbe d'équation $g(x, y) = 0$. En d'autres termes, les deux gradients $\nabla f(M_0)$ et $\nabla g(M_0)$ sont colinéaires : $\nabla f(M_0) = \lambda \nabla g(M_0)$.

Le rapport λ est appelé *multiplicateur de Lagrange*.

Théorème 18.36 : Multiplicateur de Lagrange

Soient f et g deux fonctions numériques de classe $\mathscr{C}^1$ sur l'**ouvert** $\mathscr{U}$ de E et $X = \{x \in \mathscr{U} \mid g(x) = 0\}$.
Si la restriction de f à X admet un extremum local en $a \in X$ et $\mathrm{d}g_a \neq 0$, alors $\mathrm{d}f_a$ est colinéaire à $\mathrm{d}g_a$.

Démonstration

- Si $\mathrm{d}f_a$ est nulle, alors $\mathrm{d}f_a = \lambda \mathrm{d}g_a$ avec $\lambda = 0$. Supposons désormais $\mathrm{d}f_a$ non nulle.
- Rappelons que deux formes linéaires non nulles φ et ψ ont même noyau si et seulement s'il existe $\lambda \in \mathbb{R}^*$ tel que $\varphi = \lambda\psi$ (cf. « Compléments d'algèbre linéaire »). Prouvons donc que $\mathrm{Ker}(\mathrm{d}g_a) = \mathrm{Ker}(\mathrm{d}f_a)$.
L'application $\mathrm{d}g_a$ étant non nulle, $\mathrm{Ker}(\mathrm{d}g_a) = \mathrm{T}_a X$. Soit $x \in \mathrm{Ker}(\mathrm{d}g_a)$. On sait qu'il existe $\gamma :]-\varepsilon, \varepsilon[\to X$ de classe $\mathscr{C}^1$ tel que $\gamma(0) = a$ et $\gamma'(0) = x$. De plus, $f_{|X}$ admet un extremum local en a donc $f \circ \gamma$ admet un extremum local en 0, point intérieur de $]-\varepsilon, \varepsilon[$. Ainsi,

$$(f \circ \gamma)'(0) = \mathrm{d}f_{\gamma(0)}(\gamma'(0)) = \mathrm{d}f_a(x) = 0$$

Donc $x \in \mathrm{Ker}(\mathrm{d}f_a)$. Par égalité des dimensions, $\mathrm{Ker}(\mathrm{d}f_a) = \mathrm{Ker}(\mathrm{d}g_a)$.
D'où l'existence de $\lambda \in \mathbb{R}^*$ tel que $\mathrm{d}f_a = \lambda \mathrm{d}g_a$. ■

Précisons que cette condition n'est que nécessaire et qu'il sera en général indispensable d'accompagner cette recherche de points critiques (toujours sur un ouvert!) d'une étude locale pour distinguer les extrema des points selles ou d'un argument de compacité.

Exemple

Déterminons à nouveau $\min\limits_{x^2+y^2=1} xy$ en posant cette fois-ci $f(x,y) = xy$ et $g(x,y) = x^2 + y^2 - 1$.

- Les deux fonctions numériques f et g sont polynomiales donc de classe $\mathscr{C}^\infty$ sur l'ouvert $\mathbb{R}^2$.
- Le cercle unité $X = \{(x,y) \in \mathbb{R}^2 \mid g(x,y) = 0\}$ étant compact, $f_{|X}$ admet un minimum en $(x_0, y_0) \in X$.
- $\nabla f(x,y) = \begin{bmatrix} y \\ x \end{bmatrix}$ et $\nabla g(x,y) = 2\begin{bmatrix} x \\ y \end{bmatrix} \neq 0$ pour tout $(x,y) \in X$. D'après ce qui précède, il existe $\lambda \in \mathbb{R}^*$ tel que :

$$\begin{cases} y_0 = 2\lambda x_0 \\ x_0 = 2\lambda y_0 \\ x_0^2 + y_0^2 = 1 \end{cases}$$

On trouve $(x_0, y_0, \lambda) = \left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$ ou bien $(x_0, y_0, \lambda) = \left(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}}, -\frac{1}{2}\right)$. D'où $\min\limits_{x^2+y^2=1} xy = -\frac{1}{2}$.

Exercice 11 (CCP 2023)

Soit f l'application définie sur $\mathbb{R}^2$ par $f(x,y) = 4x^2 + 12xy - y^2$. On note $\mathscr{C} = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 13\}$.

(i) Montrer que f admet un minimum et un maximum sur $\mathscr{C}$.

(ii) Les déterminer.

Exercice 12 (inégalité arithmético-géométrique)

En considérant l'application $f : (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto x_1 \times \cdots \times x_n$ et la contrainte $C_s = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n \mid x_1 + \cdots + x_n = s\}$, retrouver l'inégalité arithmético-géométrique.

Rajoutons pour finir le résultat plus général suivant, hors programme.

Théorème 18.37 : Multiplicateurs de Lagrange (HP, pour votre culture)

Soient f et $g_1, \dots, g_p$ des fonctions numériques de classe $\mathscr{C}^1$ sur l'**ouvert** $\mathscr{U}$ de E et :

$$X = \{x \in \mathscr{U} \mid g_1(x) = \cdots = g_p(x) = 0\}$$

Si la restriction de f à X admet un extremum local en $a \in X$ et si les formes linéaires $\mathrm{d}g_{1,a}, \dots, \mathrm{d}g_{p,a}$ sont indépendantes, alors il existe $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ tels que $\mathrm{d}f_a = \lambda_1 \mathrm{d}g_{1,a} + \ldots + \lambda_p \mathrm{d}g_{p,a}$.
Les réels $\lambda_1, \dots, \lambda_p$ sont appelés *multiplicateurs de Lagrange*.