

FONCTIONS DE PLUSIEURS VARIABLES

On étudie ici des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$ ($p \geq 2$, dans la pratique, $p = 2$ ou $p = 3$). $\mathbb{R}^p$ étant un espace vectoriel de dimension finie, toutes les normes y sont équivalentes. On notera donc $\| \cdot \|$ une norme quelconque sur $\mathbb{R}^p$ (dans la pratique, on utilisera le plus souvent la norme $\| \cdot \|_\infty$).

Les applications considérées seront toujours définies sur un ouvert Ω de $\mathbb{R}^p$; cela assure que, si f est définie en $a \in \Omega$, alors elle reste définie dans tout un voisinage de a .

L'ensemble $\mathbb{R}^p$ peut être considéré comme un espace vectoriel (un p -uplet $(x_1, \dots, x_n)$ étant alors un vecteur $\vec{V}$) ou comme un espace affine (une origine O étant choisie, le p -uplet $(x_1, \dots, x_p)$ représente les coordonnées d'un point M tel que $\overrightarrow{OM} = \vec{V}$).

I. Rappels

Déf 1: Applications partielles

Soit $f: \Omega \subset \mathbb{R}^p \rightarrow \mathbb{R}$, et $a = (a_1, \dots, a_p) \in \Omega$. Pour tout $i \in \llbracket 1; p \rrbracket$, la i -ème **application partielle** de f en a est l'application :

$$f_{a,i}: x \mapsto f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p).$$

Elle est définie sur un ouvert de $\mathbb{R}$ contenant a_i et à valeurs dans $\mathbb{R}$.

Déf 2: Continuité

Une application $f: \Omega \subset \mathbb{R}^p \rightarrow \mathbb{R}$ d'un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$ est dite **continue** en $a \in \Omega$ si :

$$\forall \varepsilon > 0, \exists \alpha > 0 \text{ tel que } \forall x \in \Omega, \|x - a\| < \alpha \implies |f(x) - f(a)| < \varepsilon$$

ou, plus simplement :

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Propriétés:

- a) Les applications coordonnées : $(x_1, \dots, x_p) \mapsto x_i$ sont continues (linéaires sur un espace vectoriel de dimension finie).
- b) Si $f: \mathbb{R}^p \rightarrow \mathbb{R}$ est une fonction polynomiale en les coordonnées, elle est continue.
- c) Une combinaison linéaire, un produit ou un quotient d'applications continues est une application continue là où elle est définie.
- d) Si $f: \Omega \subset \mathbb{R}^p \rightarrow \mathbb{R}$ est continue en $a \in \Omega$, alors toutes ses applications partielles $f_{a,j}$ en ce point sont continues.

Remarque : La réciproque de cette propriété est fausse : la continuité des seules applications partielles ne suffit pas à assurer la continuité en un point.

Exemple

Étudier la continuité de l'application $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$f(x, y) = \frac{x^3 y^3}{x^2 + y^4} \text{ pour } (x, y) \neq (0, 0) \text{ et } f(0, 0) = 0.$$

Solution:

– Déjà, les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto y$ étant continues, f est continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ comme quotient de fonctions continues dont le dénominateur ne s'annule pas.

– Étude en $(0, 0)$:

Plusieurs solutions :

• Puisque $x^2 + y^4 \geq x^2$ on a : $\forall (x, y) \neq (0, 0), |f(x, y)| \leq |xy^3| \leq \|(x, y)\|_\infty^4$. Ainsi,

$$|f(x, y) - f(0, 0)| \leq \|(x, y)\|_\infty^4.$$

Par suite, f est continue en $(0, 0)$.

- Utilisation des coordonnées polaires :
pour $r \neq 0$:

$$|f(r \cos \theta, r \sin \theta)| = \frac{r^6 |\cos^3 \theta \sin^3 \theta|}{r^2 \cos^2 \theta + r^4 \sin^4 \theta} = \frac{r^4 |\cos^3 \theta \sin^3 \theta|}{\cos^2 \theta + r^2 \sin^2 \theta} \leq r^4 |\cos \theta \sin^3 \theta| \leq r^4$$

donc : $\forall (x, y) \in \mathbb{R}^2, |(x, y) - f(0, 0)| \leq \|(x, y)\|_2^4$, d'où la continuité en $(0, 0)$.

Rem: Pour prouver qu'une fonction n'admet pas de limite ou n'est pas continue en $(0, 0)$ il est souvent utile de se ramener au cas des fonctions d'une variable en considérant des « chemins » de la forme (x^α, x^β) ou du type $(x, 0)$ (ou $(0, x)$) et de faire tendre x vers 0.

Exemples

- Soit f définie par $f(x, y) = \frac{|x| \sqrt{|y|}}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$. Étudier la limite de f en $(0, 0)$.

 **Solution:**

Pour $x \neq 0$, $f(x, 0) = 0 \xrightarrow{x \rightarrow 0} 0$. Si f possède une limite en $(0, 0)$, c'est donc 0 (théorème sur la limite d'une composée).

Mais $f(x, x^2) = \frac{1}{1 + x^2} \xrightarrow{x \rightarrow 0} 1 \neq 0$ donc f n'est pas continue en $(0, 0)$.

- Soit f définie par $f(x, y) = \begin{cases} \frac{x^2 y^3}{x^4 + y^6} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases}$ Étudier la continuité de f en $(0, 0)$.

 **Solution:**

Pour $x > 0$, on a $f(\sqrt{x}, \sqrt[3]{x}) = \frac{x^2}{2x^2} \xrightarrow{x \rightarrow 0+} \frac{1}{2} \neq 0$ donc f n'est pas continue en $(0, 0)$.

- Soit f définie par $f(x, y) = \frac{((x-1)^2 + y^2) \ln((x-1)^2 + y^2)}{|x| + |y|}$. Étudier la continuité de f .

 **Solution:**

f est continue en tout point différent de $(0, 0)$ et $(1, 0)$ comme produit, différence et quotient de fonctions continues dont le dénominateur ne s'annule pas.

Mais $f(x, 0) = \frac{(x-1)^2 \ln((x-1)^2)}{|x|} = \frac{2(x-1)^2 \ln|1-x|}{|x|} \sim \frac{2x}{|x|}$ qui n'a pas de limite en 0. Donc f n'a pas de limite en 0.

En $(1, 0)$, on montre facilement par croissances comparées qu'elle tend vers 0, donc peut se prolonger par continuité en ce point.

II. Dérivées partielles

Déf 3: Dérivée selon un vecteur

Soit f une application définie sur un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$, et a un point de Ω . Soit $\vec{v}$ un vecteur non nul de $\mathbb{R}^p$, et I un intervalle ouvert autour de 0 tel que, pour tout $t \in I$, $a + t\vec{v} \in \Omega$. On dit que f admet une **dérivée partielle selon le vecteur $\vec{v}$** au point a si la fonction $t \in I \mapsto f(a + t\vec{v})$ est dérivable en 0.

La dérivée correspondante est appelée dérivée de f suivant le vecteur $\vec{v}$, et est noté $\frac{\partial f}{\partial \vec{v}}(a)$ ou $D_{\vec{v}} f(a)$. Ainsi :

$$\frac{\partial f}{\partial \vec{v}}(a) = D_{\vec{v}} f(a) = \lim_{t \rightarrow 0} \frac{f(a + t\vec{v}) - f(a)}{t}.$$

Déf 4: Dérivées partielles

Soit f une application définie sur un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$, et a un point de Ω .

Soit $(\vec{e}_1, \dots, \vec{e}_p)$ la base canonique de $\mathbb{R}^p$.

On appelle **dérivées partielles** (premières) de f ses dérivées suivant les vecteurs $\vec{e}_1, \dots, \vec{e}_p$ (si elles existent).

La i -ème dérivée partielle en a est notée $\frac{\partial f}{\partial x_i}(a)$. Ainsi :

$$\forall j \in \llbracket 1; p \rrbracket, \frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i^j}(a) = \lim_{t \rightarrow a} \frac{f(a + t \vec{e}_i^j) - f(a)}{t}.$$

C'est donc aussi la dérivée en a_i de la i -ème application partielle $f_{a,i}$:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_p) - f(a_1, \dots, a_p)}{t}.$$



Rem : Dans l'écriture $\frac{\partial f}{\partial x_i}$, le « x_i » ne désigne pas véritablement une variable mais une direction.

Par exemple, soit $f : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto x^2 y^3 \end{cases}$. Que vaut $\frac{\partial f}{\partial x}(y, x)$?

III. Fonctions de classe $\mathcal{C}^1$ **Déf 5:**

Soient Ω un ouvert de $\mathbb{R}^p$ et f une application de Ω dans $\mathbb{R}$.

On dit que f est **de classe $\mathcal{C}^1$ sur Ω** si f admet des dérivées partielles par rapport à chacune des p variables en tout point de Ω , et si ces p dérivées partielles sont continues sur Ω .

Théorème 1: Développement limité d'une fonction de classe $\mathcal{C}^1$

Si f est de classe $\mathcal{C}^1$ sur Ω , alors f admet en tout point $a \in \Omega$ le développement limité d'ordre 1 :

$$f(a + h) = f(a) + \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i + o(\|h\|),$$

avec $h = (h_1, \dots, h_p)$.

Démonstration:

Pour simplifier les notations, nous nous placerons dans le cas $p = 2$.

Soit $a = (a_1, a_2) \in \Omega$ et $h = (h_1, h_2) \in \mathbb{R}^2$ suffisamment petit pour que $a + h$ soit dans Ω (cela est possible car Ω est un ouvert). On écrit alors :

$$\begin{aligned} f(a + h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) \\ &= f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1 + h_1, a_2) - f(a_1, a_2). \end{aligned}$$

En appliquant alors le théorème des accroissements finis aux applications partielles (c'est possible compte tenu des hypothèses) on obtient l'existence de $c_j \in]a_j; a_j + h_j[$ ($j = 1, 2$) tels que :

$$f(a + h) - f(a) = h_2 \frac{\partial f}{\partial x_2}(a_1 + h_1, c_2) + h_1 \frac{\partial f}{\partial x_1}(c_1, a_2).$$

D'où :

$$f(a + h) - f(a) - \sum_{i=1}^2 h_i \frac{\partial f}{\partial x_i}(a) = h_2 \left[\frac{\partial f}{\partial x_2}(a_1 + h_1, c_2) - \frac{\partial f}{\partial x_2}(a) \right] + h_1 \left[\frac{\partial f}{\partial x_1}(c_1, a_2) - \frac{\partial f}{\partial x_1}(a) \right].$$

Puisque les $\frac{\partial f}{\partial x_i}$ sont continues en a , les termes entre crochets tendent vers 0 quand $h \rightarrow (0, 0)$, d'où le résultat.

Corollaire 1.1:

| Si f est de classe $\mathcal{C}^1$ sur Ω , elle y est continue.

 Démonstration:

En effet, le résultat précédent implique, pour tout $a \in \Omega$, $\lim_{h \rightarrow 0} f(a+h) = f(a)$.

Déf 6: Différentielle

Soit f une application de classe $\mathcal{C}^1$ sur un ouvert Ω de $\mathbb{R}^p$, à valeurs dans $\mathbb{R}$. Pour tout $a \in \Omega$, on appelle **différentielle de f en a** la forme linéaire notée df_a , de $\mathbb{R}^p$ dans $\mathbb{R}$, définie par :

$$\forall h = (h_1, \dots, h_p) \in \mathbb{R}^p, \quad df_a(h) = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i \quad (*)$$

Avec cette notation, le développement limité de f en a s'écrit donc :

$$f(a+h) = f(a) + df_a(h) + o(\|h\|).$$

Rem : on peut aussi noter $df_a(h) = df_a \cdot h$

Prop 1:

Si f est de classe $\mathcal{C}^1$ sur Ω , alors pour tout $a \in \Omega$ et pour tout vecteur $\vec{v}$ de $\mathbb{R}^p$, f admet une dérivée en a selon $\vec{v}$ et :

$$\frac{\partial f}{\partial \vec{v}}(a) = df_a(\vec{v}).$$

 Démonstration:

En effet, $f(a+t\vec{v}) = f(a) + df_a(t\vec{v}) + o(\|t\vec{v}\|) = f(a) + t df_a(\vec{v}) + o(t)$ donc $\lim_{t \rightarrow 0} \frac{f(a+t\vec{v}) - f(a)}{t} = df_a(\vec{v})$.

Remarque : La réciproque est fausse : une fonction peut admettre une dérivée en un point selon n'importe quel vecteur sans être même continue en ce point.

Exemple : considérer $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

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

Notation différentielle

On suppose toujours f de classe $\mathcal{C}^1$ sur Ω .

Soit $(\vec{e}_1, \dots, \vec{e}_p)$ la base canonique de $\mathbb{R}^p$, et $(\vec{e}_1^*, \dots, \vec{e}_p^*)$ sa base duale, formée des formes linéaires coordonnées.

Ainsi, pour tout $h = (h_1, \dots, h_p) \in \mathbb{R}^p$, $\vec{e}_i^*(h) = h_i$.

Donc si f est de classe $\mathcal{C}^1$ on a :

$$\forall a \in \Omega, \forall h = (h_1, \dots, h_p) \in \mathbb{R}^p, \quad df_a(h) = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) \vec{e}_i^*(h),$$

ou encore :

$$\forall a \in \Omega, \quad df_a = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) \vec{e}_i^*.$$

Or, il est facile de vérifier, que, pour tout $a \in \mathbb{R}^p$, $(d\vec{e}_i^*)_a = \vec{e}_i^*$, puisque $\frac{\partial \vec{e}_i^*}{\partial x_j} = \delta_{i,j}$. Donc

$$\forall a \in \Omega, \quad df_a = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) \cdot (d\vec{e}_i^*)_a,$$

ou encore, de façon un peu abusive :

$$df = \sum_{i=1}^p \frac{\partial f}{\partial x_i} \cdot d\vec{e}_i^* = \sum_{i=1}^p \frac{\partial f}{\partial \vec{e}_i^*} \cdot d\vec{e}_i^*.$$

En notant plus simplement $d\vec{e}_i^* = dx_i$ (de sorte que $dx_i(h) = h_i$), la relation précédente s'écrit alors :

$$df_a = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) dx_i,$$

ou encore, de façon un peu abusive :

$$df = \sum_{i=1}^p \frac{\partial f}{\partial x_i} dx_i.$$

Exemple

Soit $f : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (r, \theta) & \longmapsto r \cos \theta \end{cases}$. Alors, de façon abrégée :

$$df = \cos \theta dr - r \sin \theta d\theta.$$

Théorème 2: Opérations sur les fonctions de classe $\mathcal{C}^1$

Soit Ω un ouvert de $\mathbb{R}^p$, et f, g des fonctions de classe $\mathcal{C}^1$ de Ω dans $\mathbb{R}$. Alors :

- Pour tout λ réel, $\lambda f + g$ est de classe $\mathcal{C}^1$ sur Ω et

$$\forall a \in \Omega, \forall i \in \llbracket 1; p \rrbracket, \frac{\partial(\lambda f + g)}{\partial x_i}(a) = \lambda \frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a)$$

ou encore, de façon abrégée : $d(\lambda f + g) = \lambda df + dg$.

- La fonction produit fg est de classe $\mathcal{C}^1$ sur Ω et :

$$\forall a \in \Omega, \forall i \in \llbracket 1; p \rrbracket, \frac{\partial(fg)}{\partial x_i}(a) = g(a) \frac{\partial f}{\partial x_i}(a) + f(a) \frac{\partial g}{\partial x_i}(a)$$

ou encore, de façon abrégée : $d(fg) = g df + f dg$.

- Si g ne s'annule pas sur Ω , la fonction $\frac{f}{g}$ est de classe $\mathcal{C}^1$ sur Ω et :

$$\forall a \in \Omega, \forall i \in \llbracket 1; p \rrbracket, \frac{\partial\left(\frac{f}{g}\right)}{\partial x_i}(a) = -\frac{1}{g(a)^2} \left(g(a) \frac{\partial f}{\partial x_i}(a) - f(a) \frac{\partial g}{\partial x_i}(a) \right)$$

ou encore, de façon abrégée : $d\left(\frac{f}{g}\right) = \frac{g df - f dg}{g^2}$.

- Si φ est une application de classe $\mathcal{C}^1$ de $f(\Omega)$ dans $\mathbb{R}$, alors $\varphi \circ f$ est de classe $\mathcal{C}^1$ sur Ω et :

$$\forall a \in \Omega, \forall i \in \llbracket 1; p \rrbracket, \frac{\partial(\varphi \circ f)}{\partial x_i}(a) = \varphi'(f(a)) \frac{\partial f}{\partial x_i}(a).$$

Démonstration:

Tous ces résultats découlent des résultats similaires sur les fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Démontrons par exemple le résultat sur le produit.

Pour tout $a \in \Omega$, et tout $i \in \llbracket 1; p \rrbracket$, la fonction $\varphi_i : t \mapsto f(a + t\vec{e}_i) \cdot g(a + t\vec{e}_i)$, définie sur un voisinage de 0, admet une dérivée en 0 comme produit de deux telles fonctions.

Et, d'après la formule donnant la dérivée d'un produit :

$$\lim_{t \rightarrow 0} \frac{(fg)(a + t\vec{e}_i) - (fg)(a)}{t} = \varphi_i'(0) = g(a) \frac{\partial f}{\partial x_i}(a) + f(a) \frac{\partial g}{\partial x_i}(a).$$

Cela prouve que fg admet en a des dérivées partielles, et puisque f et g sont de classe $\mathcal{C}^1$, ces dérivées partielles sont continues sur Ω , c'est-à-dire fg est de classe $\mathcal{C}^1$.

Elle admet donc en tout point a une différentielle, donnée par :

$$d(fg)_a = \sum_{i=1}^p \left(g(a) \frac{\partial f}{\partial x_i}(a) + f(a) \frac{\partial g}{\partial x_i}(a) \right) dx_i = g(a) \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) dx_i + f(a) \sum_{i=1}^p \frac{\partial g}{\partial x_i}(a) dx_i = g(a) df_a + f(a) dg_a.$$

Déf 7: Gradient

Soit f une application de classe $\mathcal{C}^1$ d'un ouvert Ω de $\mathbb{R}^p$, à valeurs dans $\mathbb{R}$. Soit $a \in \Omega$.

On appelle **gradient de f en a** , noté $\overrightarrow{\text{grad}} f(a)$ ou $\nabla f(a)$, le vecteur de $\mathbb{R}^p$ dont les coordonnées dans la base canonique sont :

$$\left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a) \right).$$

Lorsqu'on munit $\mathbb{R}^p$ de son produit scalaire canonique, le gradient est donc caractérisé par la relation :

$$\forall \vec{v} \in \mathbb{R}^p, \text{d}f_a(\vec{v}) = \langle \nabla f(a) | \vec{v} \rangle.$$

Compte tenu de l'inégalité de Cauchy-Schwarz, cette relation montre que le vecteur gradient (lorsqu'il est non nul) donne la direction selon laquelle la dérivée de f en a , $\frac{\partial f}{\partial \vec{v}}(a)$, est maximale : en effet, $|\langle \nabla f(a) | \vec{v} \rangle| \leq \|\nabla f(a)\| \|\vec{v}\|$, avec égalité si et seulement si $\vec{v}$ est colinéaire à $\nabla f(a)$.

IV. Dérivée le long d'un arc paramétré**IV.1. Arcs paramétrés****IV.1.1. Définitions****Déf 8:**

On appelle arc paramétré ou courbe paramétrée de $\mathbb{R}^p$ un couple $(I, \vec{F})$ où I est un intervalle de $\mathbb{R}$ et $\vec{F}$ une application de I dans $\mathbb{R}^p$. L'arc est dit de classe $\mathcal{C}^k$ si $\vec{F}$ est de classe $\mathcal{C}^k$.

Pour tout t on notera $M(t)$ le point de $\mathbb{R}^p$ tel que $\overrightarrow{OM}(t) = \vec{F}(t)$.

L'ensemble

$$\Gamma = \{M(t) \in \mathbb{R}^p \mid t \in I\}$$

des points de $\mathbb{R}^p$ atteints est appelé support de l'arc.

Rem: En interprétant le paramètre t comme le temps, un arc paramétré représente le mouvement d'un mobile $M(t)$; le support de l'arc est la trajectoire du mobile. Si $\vec{F}$ est de classe $\mathcal{C}^k$ avec $k \geq 2$,

$$\vec{F}'(t) = \frac{d\overrightarrow{OM}}{dt} \text{ est le vecteur vitesse et } \vec{F}''(t) = \frac{d^2\overrightarrow{OM}}{dt^2} \text{ le vecteur accélération.}$$

IV.1.2. Tangente

Soit $(I, \vec{F})$ un arc paramétré de classe $\mathcal{C}^n$, et $t_0 \in I$. On suppose qu'il existe un entier $i \in \llbracket 1; n \rrbracket$ tel que $\vec{F}^{(i)}(t_0) \neq 0$; on note k le plus petit de ces entiers.

En appliquant la formule de Taylor-Young aux fonctions coordonnées de $\vec{F}$, on obtient le développement limité :

$$\vec{F}(t_0 + h) = \vec{F}(t_0) + h^k \frac{\vec{F}^{(k)}(t_0)}{k!} + \vec{o}(h^k),$$

de sorte que $\lim_{h \rightarrow 0} \frac{\overrightarrow{M(t_0)M(t_0+h)}}{h^k} = \frac{\vec{F}^{(k)}(t_0)}{k!}.$

Or le vecteur $\frac{\overrightarrow{M(t_0)M(t_0+h)}}{h^k}$ est un vecteur directeur de la droite passant par les points $M(t_0)$ et $M(t_0 + h)$, ce qui justifie la définition suivante :

Déf 9:

On appelle tangente à l'arc au point $M(t_0)$ la droite passant par $M(t_0)$ et dirigée par le vecteur $\vec{F}^{(k)}(t_0)$.

Déf 10:

Le point $M(t_0)$ est dit régulier si $k = 1$, c'est-à-dire $\vec{F}'(t_0) \neq 0$. Il est dit stationnaire si $k > 1$, c'est-à-dire $\vec{F}'(t_0) = \vec{0}$. Un arc est dit régulier si tous ses points sont réguliers.

IV.2. Règle de la chaîne**Théorème 3: dérivée le long d'un arc paramétré, ou « règle de la chaîne »**

Soit f une application de classe $\mathcal{C}^1$ d'un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$.

Soient $(\varphi_1, \dots, \varphi_p)$ p fonctions de classe $\mathcal{C}^1$ définies sur un intervalle I de $\mathbb{R}$ telles que :

$$\forall t \in I, \varphi(t) = (\varphi_1(t), \dots, \varphi_p(t)) \in \Omega.$$

Alors la fonction $g: t \mapsto f \circ \varphi(t) = f(\varphi_1(t), \dots, \varphi_p(t))$ est de classe $\mathcal{C}^1$ sur I et l'on a :

$$\forall t \in I, g'(t) = \sum_{i=1}^p \varphi'_i(t) \frac{\partial f}{\partial x_i}(\varphi_1(t), \dots, \varphi_p(t)).$$

Démonstration:

Chaque fonction φ_i admet en tout $t \in I$ un développement limité d'ordre 1 :

$$\varphi_i(t+h) = \varphi_i(t) + h\varphi'_i(t) + h\varepsilon_i(h)$$

où $\lim_{h \rightarrow 0} \varepsilon_i(h) = 0$, et où h appartient à un certain voisinage de 0 tel que $t+h \in I$.

On aura donc :

$$g(t+h) = f(\varphi_1(t) + h\varphi'_1(t) + h\varepsilon_1(h), \dots, \varphi_p(t) + h\varphi'_p(t) + h\varepsilon_p(h)) = f(\varphi(t) + \vec{k})$$

où $\vec{k}$ est le vecteur de $\mathbb{R}^p$ de coordonnées les $h\varphi'_i(t) + h\varepsilon_i(h)$.

On écrit alors le développement limité de f en $\varphi(t)$:

$$f(\varphi(t) + \vec{k}) = f(\varphi(t)) + \sum_{i=1}^p k_i \frac{\partial f}{\partial x_i}(\varphi(t)) + o(\|\vec{k}\|).$$

Puisque les φ'_i sont bornées au voisinage de 0, on a $\|\vec{k}\| = O(h)$ donc on peut réécrire le développement limité précédent sous la forme :

$$f(\varphi(t) + \vec{k}) = f(\varphi(t)) + \sum_{i=1}^p (h\varphi'_i(t) + h\varepsilon_i(h)) \frac{\partial f}{\partial x_i}(\varphi(t)) + o(h)$$

et finalement :

$$g(t+h) = g(t) + h \sum_{i=1}^p \varphi'_i(t) \frac{\partial f}{\partial x_i}(\varphi(t)) + o(h)$$

ce qui donne le résultat.

Prop 2: caractérisation des fonctions constantes sur un ouvert convexe

Une application f de classe $\mathcal{C}^1$ sur un ouvert *convexe* Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$ est constante sur Ω si et seulement si sa différentielle df est nulle en tout point de Ω (ce qui équivaut à dire que toutes les dérivées partielles de f sont nulles sur Ω).

Démonstration:

- Il est clair que si f est constante, sa différentielle est nulle en tout point.
- Réciproquement, supposons $df_x = 0$ pour tout $x \in \Omega$. Soient alors $a, b \in \Omega$ et $g: t \in [0;1] \mapsto f(at + (1-t)b)$; cette définition a bien un sens puisque pour tout $t \in [0;1]$, on a $at + (1-t)b \in \Omega$. g est alors de classe $\mathcal{C}^1$ et pour tout $t \in [0;1]$: $g'(t) = df_{at+(1-t)b}(a-b) = 0$. Donc g est constante et en particulier : $g(1) = g(0)$ soit $f(b) = f(a)$ pour tout $b \in \Omega$ c'est-à-dire f constante.

Généralisation

Déf 11:

On dira qu'une application φ définie sur un ouvert V de $\mathbb{R}^q$ à valeurs dans $\mathbb{R}^p$ est de classe $\mathcal{C}^1$ sur V si toutes ses applications coordonnées (définies sur V à valeurs dans $\mathbb{R}$) sont de classe $\mathcal{C}^1$.

On peut étendre la « règle de la chaîne » à la composée d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$ avec une fonction de $\mathbb{R}^q$ dans $\mathbb{R}^p$. Pour simplifier les calculs, nous nous placerons dans le cas $p = q = 2$. On a alors le théorème suivant.

Théorème 4: changement de variable dans $\mathbb{R}^2$

Soit f une fonction de classe $\mathcal{C}^1$ d'un ouvert Ω de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$, et $\varphi: (u, v) \mapsto (\varphi_1(u, v), \varphi_2(u, v))$ une fonction définie sur un ouvert V de $\mathbb{R}^2$, à valeurs dans Ω , de classe $\mathcal{C}^1$ (c'est-à-dire que φ_1 et φ_2 sont de classe $\mathcal{C}^1$).

Posons alors $F = f \circ \varphi$. Alors F est de classe $\mathcal{C}^1$ sur V et, pour tout $(a, b) \in V$ on a :

$$\begin{aligned}\frac{\partial F}{\partial u}(a, b) &= \frac{\partial f}{\partial x}(\varphi_1(a, b), \varphi_2(a, b)) \cdot \frac{\partial \varphi_1}{\partial u}(a, b) + \frac{\partial f}{\partial y}(\varphi_1(a, b), \varphi_2(a, b)) \cdot \frac{\partial \varphi_2}{\partial u}(a, b) \\ \frac{\partial F}{\partial v}(a, b) &= \frac{\partial f}{\partial x}(\varphi_1(a, b), \varphi_2(a, b)) \cdot \frac{\partial \varphi_1}{\partial v}(a, b) + \frac{\partial f}{\partial y}(\varphi_1(a, b), \varphi_2(a, b)) \cdot \frac{\partial \varphi_2}{\partial v}(a, b).\end{aligned}$$

Démonstration:

On a $F(u, v) = f(\varphi_1(u, v), \varphi_2(u, v))$, et on applique la règle de la chaîne aux applications partielles $u \mapsto F(u, v)$ et $v \mapsto F(u, v)$.

On note en général de façon abusive (x, y) les applications coordonnées de φ (au lieu de φ_1 et φ_2), de sorte que le changement de variable s'écrit $F(u, v) = f(x(u, v), y(u, v))$.

Les formules précédentes s'écrivent alors, en abrégé :

$$\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u}, \quad \frac{\partial F}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v}.$$

Exemples

1. Gradient en coordonnées polaires.

Soit $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. On pose $g(r, \theta) = f(r \cos \theta, r \sin \theta)$.

$$\text{On a alors : } \begin{cases} \frac{\partial g}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} \\ \frac{\partial g}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta}. \end{cases}$$

Voir calculs en classe

Cela permet d'obtenir : $\nabla f(r \cos \theta, r \sin \theta) = \frac{\partial g}{\partial r}(r, \theta) \vec{u}_\theta + \frac{1}{r} \frac{\partial g}{\partial \theta}(r, \theta) \vec{v}_\theta$.

2. Résoudre sur $(\mathbb{R}_+^*)^2$ l'équation aux dérivées partielles :

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = \frac{y}{x}$$

en utilisant le changement de variable : $u = \frac{y}{x}$, $v = x^2 + y^2$.

3. Résoudre sur $\mathbb{R}_+^* \times \mathbb{R}_+^*$ l'équation aux dérivées partielles :

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = \sqrt{x^3 + y^3}$$

en utilisant les coordonnées polaires.

V. Dérivées partielles d'ordre deux

Déf 12:

Soit f une fonction définie sur un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$, et $(i, j) \in \llbracket 1; p \rrbracket^2$.
On suppose que f possède en tout point une j^{e} dérivée partielle $\frac{\partial f}{\partial x_j}$, qui est donc une application de Ω dans $\mathbb{R}$. Cette application peut elle-même posséder une i^{e} dérivée partielle qui sera alors notée :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right),$$

et appelée **dérivée partielle seconde** de f par rapport aux variables x_i, x_j .

On dit que $f: \Omega \rightarrow \mathbb{R}^n$ est **de classe $\mathcal{C}^2$ sur Ω** si toutes ses dérivées partielles secondes existent et sont continues sur Ω .

Théorème 5: Théorème de Schwarz (admis)

Si $f \in \mathcal{C}^2(\Omega, \mathbb{R}^n)$ où Ω est un ouvert de $\mathbb{R}^p$, on a :

$$\forall (i, j) \in \llbracket 1; p \rrbracket^2, \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Exercices

$$1. \text{ Soit } f : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto \begin{cases} \frac{x^3 y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } x = y = 0. \end{cases} \end{cases}$$

Montrer que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et que $\frac{\partial^2 f}{\partial x \partial y}(0, 0) \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0)$.

2. Résoudre sur $\mathbb{R}^2$ l'équation aux dérivées partielles :

$$\frac{\partial^2 f}{\partial x^2} - 3 \frac{\partial^2 f}{\partial x \partial y} + 2 \frac{\partial^2 f}{\partial y^2} = 0$$

en utilisant un changement de variables de la forme $\begin{cases} u = ax + by \\ v = cx + dy. \end{cases}$

3. Résoudre, sur un ouvert Ω bien choisi, l'équation aux dérivées partielles :

$$x \frac{\partial^2 f}{\partial x^2} + y \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial x} = y$$

en utilisant le changement de variables $x = uv, y = u + v$.

4. Laplacien en coordonnées polaires

Réponse : si $F(r, \theta) = f(r \cos \theta, r \sin \theta)$ alors $\Delta f = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 F}{\partial \theta^2} + \frac{1}{r} \frac{\partial F}{\partial r}$.

Matrice hessienne

Déf 13:

Soit f une application de classe $\mathcal{C}^2$ d'un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$, et soit $a \in \Omega$.

On appelle matrice hessienne de f en a la matrice carrée d'ordre p , notée $H_f(a)$, telle que, pour tout $(i, j) \in \llbracket 1; p \rrbracket^2$, le terme d'indice (i, j) de $H_f(a)$ est égal à $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$.

Autrement dit :

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_p} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_p \partial x_1} & \frac{\partial^2 f}{\partial x_p \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_p^2} \end{pmatrix}.$$

On remarquera que, compte tenu du théorème de Schwarz, la matrice $H_f(a)$ est une matrice symétrique.

Théorème 6: Formule de Taylor-Young à l'ordre 2

Soit f une application de classe $\mathcal{C}^2$ d'un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$. Alors f admet en tout point $a \in \Omega$ le développement limité d'ordre 2 :

$$f(a+h) \underset{h \rightarrow 0}{=} f(a) + \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i + \frac{1}{2} \sum_{(i,j) \in \llbracket 1;p \rrbracket^2} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(\|h\|)$$

ce qui peut aussi s'écrire, en confondant h avec le vecteur colonne correspondant de $\mathcal{M}_{p,1}(\mathbb{R})$:

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

 **Démonstration:**

Pour simplifier les notations, nous nous placerons dans le cas $p = 2$.

Soit $a = (a_1, a_2) \in \Omega$ et $h = (h_1, h_2) \in \mathbb{R}^2$ suffisamment petit pour que $a+h$ soit dans Ω (cela est possible car Ω est un ouvert.)

Pour tout t dans un voisinage de 0 convenable (pour que $a+th$ soit dans Ω), on peut poser

$$\varphi(t) = f(a+th) = f(a_1+th_1, a_2+th_2).$$

Alors φ est de classe $\mathcal{C}^2$ au voisinage de 0, et, d'après la règle de la chaîne :

$$\varphi'(t) = h_1 \frac{\partial f}{\partial x_1}(a+th) + h_2 \frac{\partial f}{\partial x_2}(a+th)$$

puis en recommençant :

$$\varphi''(t) = h_1^2 \frac{\partial^2 f}{\partial x_1^2}(a+th) + h_1 h_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(a+th) + h_2 h_1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(a+th) + h_2^2 \frac{\partial^2 f}{\partial x_2^2}(a+th).$$

La formule de Taylor avec reste intégrale appliquée à φ entre 0 et 1 s'écrit :

$$\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 (1-t) \varphi''(t) dt \underset{\text{astuce!}}{=} \varphi(0) + \varphi'(0) + \frac{1}{2} \varphi''(0) + \int_0^1 (1-t) (\varphi''(t) - \varphi''(0)) dt$$

donc

$$f(a+h) = f(a) + \sum_{i=1}^2 \frac{\partial f}{\partial x_i}(a) h_i + \frac{1}{2} \sum_{(i,j) \in \llbracket 1;2 \rrbracket^2} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + \varepsilon(h)$$

avec $\varepsilon(h) = \int_0^1 (1-t) (\varphi''(t) - \varphi''(0)) dt$, donc $|\varepsilon(h)| \leq \sup_{t \in [0;1]} |\varphi''(t) - \varphi''(0)|$. Or :

$$|\varphi''(t) - \varphi''(0)| \leq \|h\|_\infty^2 \left(\left| \frac{\partial^2 f}{\partial x_1^2}(a+th) - \frac{\partial^2 f}{\partial x_1^2}(a) \right| + 2 \left| \frac{\partial^2 f}{\partial x_1 \partial x_2}(a+th) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \right| + \left| \frac{\partial^2 f}{\partial x_2^2}(a+th) - \frac{\partial^2 f}{\partial x_2^2}(a) \right| \right)$$

et, puisque les dérivées partielles secondes de f sont continues en a , on obtient (rigoureusement, il faudrait détailler en revenant à la définition de la continuité avec des ε ...) $\varepsilon(h) = o(\|h\|_\infty^2)$, cqfd.

VI. Extrema d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$

Déf 14: Extrema

Soit f une application de classe $\mathcal{C}^1$ d'un ouvert Ω de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$. Soit $a \in \Omega$.

On dit que f admet un minimum (respectivement un maximum) global en a si :

$$\forall x \in \Omega, f(x) \geq f(a) \quad (\text{respectivement } f(x) \leq f(a)).$$

On dit que f admet un minimum (respectivement un maximum) local en a si :

$$\exists V \in \mathcal{V}(a) \text{ tel que } \forall x \in V, f(x) \geq f(a) \quad (\text{respectivement } f(x) \leq f(a)).$$

Théorème 7:

Si f admet un extremum local en a , alors $\nabla(f)(a) = 0$ (c'est-à-dire que pour tout $i \in \llbracket 1; p \rrbracket$, $\frac{\partial f}{\partial x_i}(a) = 0$).

 **Démonstration:**

Supposons que f admette un extremum local en a . Alors pour tout $\vec{v} \in \mathbb{R}^p$ la fonction $\varphi: t \mapsto f(a + t\vec{v})$ (définie au voisinage de 0), admet un extremum local en 0. Sa dérivée y est donc nulle. Or $\varphi'(t) = df_a(\vec{v}) = \langle \nabla f(a) | \vec{v} \rangle$ d'où le résultat.



Rem : La réciproque de ce théorème est fautive : le gradient de f peut s'annuler en a sans qu'il s'agisse d'un extremum.

Par exemple, soit $f: \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) \longmapsto xy \end{cases}$. On a $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0)$, mais f ne peut admettre d'extremum local en $(0,0)$ puisque, au voisinage de ce point, elle prend aussi bien des valeurs négatives que des valeurs positives.

Déf 15:

Étant donnée $f \in \mathcal{C}^1(\Omega, \mathbb{R})$, on dit que $a \in \Omega$ est un point critique de f si $\nabla f(a) = 0$.

Théorème 8: Condition suffisante d'extremum en un point critique

Soit $f \in \mathcal{C}^2(\Omega, \mathbb{R})$, et $a \in \Omega$ un point critique de f , telle que la matrice hessienne de f en a soit inversible. Alors :

- Si toutes les valeurs propres de $H_f(a)$ sont (strictement) positives, f admet un minimum local en a ;
- Si toutes les valeurs propres de $H_f(a)$ sont (strictement) négatives, f admet un maximum local en a ;
- Si $H_f(a)$ possède des valeurs propres positives et des valeurs propres négatives, f ne présente ni minimum ni maximum en a ; a s'appelle alors point col.

 **Démonstration:**

Si a est un point critique, la formule de Taylor-Young se résume à :

$$f(a+h) - f(a) = \frac{1}{2} q_a(h) + o(\|h\|^2)$$

avec $q_a(h) = h^T H_f(a) h$. On a donc aussi, pour $h \neq 0$:

$$f(a+h) - f(a) = \frac{\|h\|^2}{2} \left(q_a\left(\frac{h}{\|h\|}\right) + \varepsilon(h) \right) \quad \text{avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0.$$

• Supposons que toutes les valeurs propres de $H_f(a)$ sont strictement positives, ce qui équivaut à dire que la matrice $H_f(a)$ est définie positive, ou encore que, pour tout $h \neq 0$, $h^T H_f(a) h > 0$ (cf. cours sur les espaces euclidiens). L'application q_a étant une fonction polynomiale en les coordonnées de h est continue; en particulier, elle est continue sur la sphère unité de $\mathbb{R}^p$, qui est une partie fermée bornée. D'après le théorème des bornes atteintes, elle y admet un minimum m , qui est atteint.

Or si h est non nul, $h^T H_f(a) h$ est aussi non nul : en effet, la matrice $H_f(a)$ étant symétrique définie positive peut s'écrire $H_f(a) = B^T B$ avec B inversible, donc $h^T H_f(a) h = h^T B^T B h = \|Bh\|^2 \neq 0$.

Le minimum m de q_a sur la sphère unité est donc strictement positif, et on a ainsi, pour tout $h \neq 0$:

$$f(a+h) - f(a) \geq \frac{\|h\|^2}{2} (m + \varepsilon(\|h\|)).$$

Pour $\|h\|$ assez petit, on a donc $f(a+h) - f(a) > 0$: f possède en a un minimum local strict.

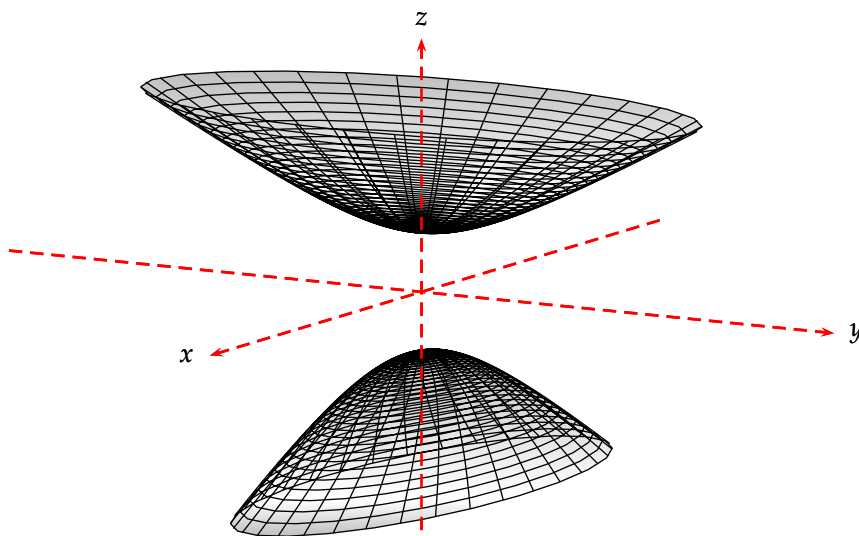
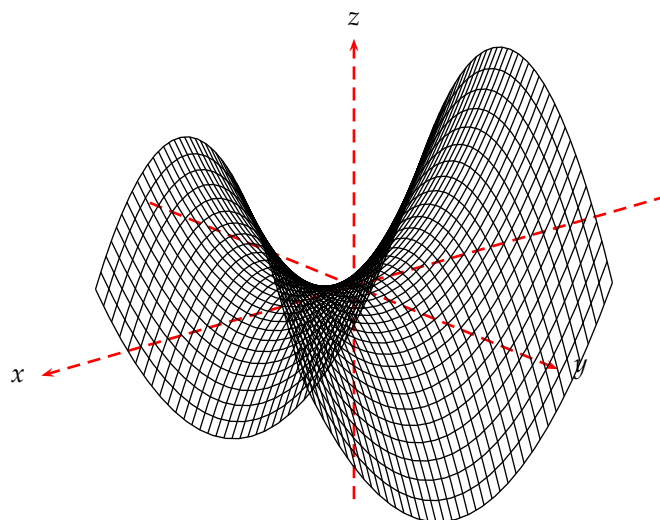
- Si toutes les valeurs propres de q_a sont strictement négatives, il suffit de remplacer f par $-f$ pour se ramener au cas précédent et conclure.
- Si q_a possède à la fois des valeurs propres positives et négatives (strictement), en considérant des vecteurs propres associés, on peut trouver des vecteurs unitaires u et v tels que $q_a(u) = \lambda > 0$ $q_a(v) = \mu < 0$ (car si u est un vecteur propre unitaire associé à λ on a $u^T H_f(a) u = u^T (\lambda u) = \lambda \|u\|^2 = \lambda$).
On pourra donc trouver au voisinage de a des points de la forme $a + tu$ où $f(a + tu) - f(a) > 0$ et des points de la forme $a + tv$ où $f(a + tv) - f(a) < 0$: f ne présente pas d'extrémum en a .

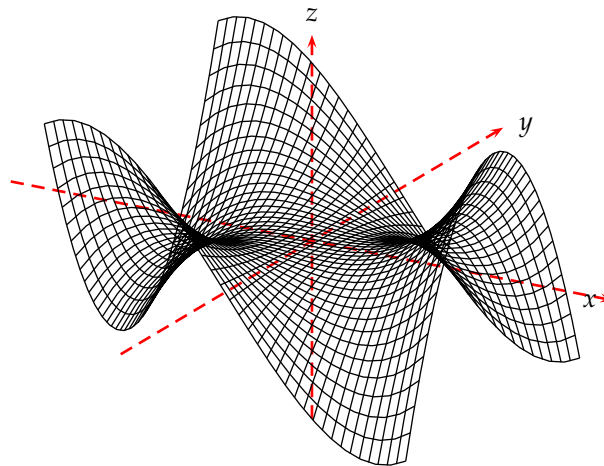
Déf 16:

Soit $f \in \mathcal{C}^2(\Omega, \mathbb{R})$, et a un point critique de f ($\nabla f(a) = 0$). On dit que $a \in \Omega$ est un point critique dégénéré si la matrice hessienne $H_f(a)$ n'est pas inversible, c'est-à-dire si elle admet 0 pour valeur propre.

Interprétation géométrique

On se place dans $\mathbb{R}^3$, et on considère la surface d'équation $z = f(x, y)$, où f est de classe $\mathcal{C}^2$. On peut avoir les dispositions suivantes :

SURFACE D'ÉQUATION $x^2 + y^2 - z^2 + 1 = 0$ – MINIMUM ET MAXIMUMSURFACE D'ÉQUATION $z = x^2 - y^2$ « SELLE DE CHEVAL » – POINT COL

SURFACE D'ÉQUATION $z = x^3 - 3xy^2$ « SELLE DE SINGE » – POINT CRITIQUE DÉGÉNÉRÉ**Cas particulier : $p=2$**

Dans le cas $p = 2$, la matrice $H_f(a)$ est symétrique réelle d'ordre 2, donc on peut directement connaître le cas dans lequel on est simplement en calculant sa trace et son déterminant.

Exemples

1. Étudier les extrema sur $\mathbb{R}^2$ de la fonction $f: (x, y) \mapsto x^4 + y^4 - x^2 + y^2$.

Solution:

- On commence par rechercher les points critiques en résolvant :

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

On trouve les couples : $(0, 0)$ et $(\pm \frac{1}{\sqrt{2}}, 0)$.

- Étude en $(0, 0)$.

On peut remarquer que f est déjà donnée sous la forme d'un développement limité d'ordre 2 en ce point :

$$f(x, y) = -x^2 + y^2 + o(\|(x, y)\|^2).$$

Ainsi, ici, $q_{(0,0)}(x, y) = -x^2 + y^2$, sans calcul. Il y a une valeur propre positive (+1) et une négative (-1), donc il s'agit d'un point col.

On pouvait aussi remarquer, sans faire appel au théorème, que $f(x, 0) = x^4 - x^2 \underset{x \rightarrow 0}{\sim} -x^2$ et $f(0, x) = x^4 + x^2 \underset{x \rightarrow 0}{\sim} x^2$, donc f prend au voisinage de $(0, 0)$ des valeurs positives et des valeurs négatives. Puisque $f(0, 0) = 0$, il ne peut s'agir d'un extremum.

- Étude en $(\frac{1}{\sqrt{2}}, 0)$ (par exemple).

On a, pour tout $(x, y) \in \mathbb{R}^2$,

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 12x^2 - 2, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 12y^2 + 2 \quad \text{et} \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 0$$

donc la matrice hessienne de f en le point critique $(\frac{1}{\sqrt{2}}, 0)$ vaut $\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$. Les deux valeurs propres sont strictement positives, il s'agit donc d'un minimum local.

On peut préciser. Posons $\delta(h, k) = f(\frac{1}{\sqrt{2}} + h, k) - f(\frac{1}{\sqrt{2}}, 0)$. Un calcul simple donne :

$$\forall (h, k) \in \mathbb{R}^2, \delta(h, k) = h^2(h^2 + 2\sqrt{2}h + 2) + k^4$$

donc $\delta(h, k) \geq 0$ pour tout (h, k) : il s'agit d'un minimum global.

2. Même fonction mais chercher cette fois-ci les extrema sur $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

Solution:

- Si f admet un extremum sur l'ouvert $\overset{\circ}{D} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$, c'est nécessairement un point critique, donc en $(\pm \frac{1}{\sqrt{2}}, 0)$. D'après l'étude précédente, il s'agit en ces points du minimum global de f , donc aussi le minimum sur D .
 - f étant continue sur le fermé borné D , elle y est bornée et atteint ses bornes. Compte tenu de ce qui précède, elle admet donc un maximum sur D qui est nécessairement atteint sur la frontière.
- On considère donc la restriction de f au cercle d'équation $x^2 + y^2 = 1$.
 Sur ce cercle, on a $f(x, y) = x^4 + (1 - x^2)^2 - x^2 + (1 - x^2) = 2(x^2 - 1)^2$, donc f atteint son maximum sur le cercle (et sur D) pour $x = 0$ et $y = \pm 1$. Ce maximum est égal à 2.

3. Déterminer les extrema locaux sur $\mathbb{R} \times \mathbb{R}_+^*$ de $f: (x, y) \mapsto y(x^2 + (\ln y)^2)$. Sont-ce des extrema globaux ?
4. Même question avec $f: (x, y) \mapsto x^4 + y^4 - 4xy$ sur $\mathbb{R}^2$.
5. Même question avec $f: (x, y) \mapsto x^3 + xy^2 - x^2y - y^3$ sur $\mathbb{R}^2$.

VII. Applications géométriques

VII.1. Application aux courbes planes

Déf 17:

Une courbe du plan définie par une équation cartésienne est un ensemble de points (x, y) du plan vérifiant une équation de la forme :

$$f(x, y) = 0$$

où f est une application de classe $\mathcal{C}^1$ sur un ouvert Ω de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.

Un point $M(x_0, y_0)$ de la courbe $\mathcal{C}$ d'équation $f(x, y) = 0$ est dit **régulier** si ce n'est pas un point critique de f , c'est-à-dire si :

$$\nabla f(x_0, y_0) \neq 0.$$

Théorème 9:

Si M_0 est un point régulier de $\mathcal{C}$, la courbe admet en ce point une tangente ; c'est la droite passant par M_0 et de vecteur normal $\nabla f(x_0, y_0)$, c'est-à-dire d'équation :

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0.$$

Démonstration:

On donne une démonstration simplifiée, en admettant qu'il est possible de paramétrer localement la courbe $\mathcal{C}$ (ce résultat est exact en un point régulier, mais difficile à démontrer).

Ainsi, on suppose qu'il existe une fonction $\vec{F}: I \rightarrow \mathbb{R}^2$ de classe $\mathcal{C}^1$ sur I intervalle de $\mathbb{R}$ telle que, au voisinage de (x_0, y_0) , $(x, y) \in \mathcal{C} \iff \exists t \in I$ tq $(x, y) = \vec{F}(t)$.

On a alors, pour tout $t \in I$, $f(x(t), y(t)) = 0$ donc en dérivant à l'aide de la règle de la chaîne :

$$x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t)) = 0.$$

Cette relation montre que le vecteur gradient est orthogonal au vecteur de coordonnées $(x'(t), y'(t))$ qui est précisément un vecteur directeur de la tangente à $\mathcal{C}$ (si le point est régulier, cf. cours sur les arcs paramétrés).

VII.2. Application aux surfaces

Déf 18:

Une **surface** définie par une équation cartésienne est un ensemble de points (x, y, z) de l'espace vérifiant une équation de la forme :

$$f(x, y, z) = 0$$

où f est une application de classe $\mathcal{C}^1$ sur un ouvert Ω de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}$.

Un point $M(x_0, y_0, z_0)$ de la surface $\mathcal{S}$ d'équation $f(x, y, z) = 0$ est dit **régulier** si ce n'est pas un point critique de f , c'est-à-dire si :

$$\nabla f(x_0, y_0, z_0) \neq 0.$$

Déf 19:

On appelle **courbe tracée sur la surface** $\mathcal{S}$ d'équation $f(x, y, z) = 0$ tout arc paramétré (I, γ) où I est un intervalle de $\mathbb{R}$ et où $\gamma: I \rightarrow \mathbb{R}^3$ est une application de classe $\mathcal{C}^1$ à valeurs dans Ω telle que :

$$\forall t \in I, f(\gamma(t)) = 0$$

c'est-à-dire que pour tout t le point $M(t)$ de l'arc γ appartient à $\mathcal{S}$.

En dérivant alors la relation précédente à l'aide de la règle de la chaîne on obtient :

$$\forall t \in I, \langle \nabla f(\gamma(t)) | \gamma'(t) \rangle = 0$$

donc en tout point régulier (de l'arc et de la surface), la tangente à l'arc est orthogonale au gradient. Cela justifie la définition suivante.

Déf 20:

Si M_0 est un point régulier de $\mathcal{S}$, la surface admet en ce point un **plan tangent**; c'est, par définition, le plan passant par M_0 et de vecteur normal $\nabla f(x_0, y_0, z_0)$, c'est-à-dire d'équation :

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0)(z - z_0) = 0.$$

La droite passant par M_0 et de vecteur directeur $\nabla f(x_0, y_0, z_0)$ s'appelle la **normale** à $\mathcal{S}$ en M_0 .

Exemples

1. On considère la surface (S) d'équation $xyz = 1$.

Montrer que pour tout point $M \in (S)$, les intersections du plan tangent à (S) en M avec les plans xOy , yOz et xOz forment un tétraèdre de volume constant.

 **Solution:**

L'équation de (S) peut s'écrire :

$$f(x, y, z) = 0$$

où f est l'application $f: \begin{cases} \mathbb{R}^3 & \longrightarrow \mathbb{R} \\ (x, y, z) & \longmapsto xyz - 1. \end{cases}$

f est évidemment de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$.

Si $M_0(x_0, y_0, z_0)$ appartient à (S) , c'est-à-dire si $x_0 y_0 z_0 = 1$, le gradient de f en ce point a pour coordonnées :

$$\frac{\partial f}{\partial x}(M_0) = y_0 z_0, \quad \frac{\partial f}{\partial y}(M_0) = x_0 z_0 \quad \text{et} \quad \frac{\partial f}{\partial z}(M_0) = x_0 y_0.$$

Puisque $x_0 y_0 z_0 = 1$, x_0 , y_0 et z_0 ne sont pas nuls donc $\forall M_0 \in (S)$, $\overrightarrow{\text{grad} f}(M_0) \neq \vec{0}$ donc la surface (S) est régulière.

D'après les calculs ci-dessus, l'équation du plan tangent en un point $M_0(x_0, y_0, z_0)$ de (S) est :

$$y_0 z_0 (x - x_0) + x_0 z_0 (y - y_0) + x_0 y_0 (z - z_0) = 0,$$

soit encore, compte tenu de $x_0 y_0 z_0 = 1$:

$$\frac{x}{x_0} + \frac{y}{y_0} + \frac{z}{z_0} = 3.$$

Le point d'intersection A de ce plan avec l'axe Ox s'obtient pour $y = z = 0$ donc A a pour coordonnées $(3x_0, 0, 0)$. De même, le point d'intersection B avec l'axe Oy a pour coordonnées $(0, 3y_0, 0)$ et le point d'intersection C avec l'axe Oz a pour coordonnées $(0, 0, 3z_0)$.

Le volume du tétraèdre $OABC$ est alors :

$$V = \frac{1}{6} |\det(\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC})| = \frac{1}{6} \begin{vmatrix} 3x_0 & 0 & 0 \\ 0 & 3y_0 & 0 \\ 0 & 0 & 3z_0 \end{vmatrix} = \frac{9}{2} x_0 y_0 z_0 = \frac{9}{2}.$$

On constate ainsi que le volume V ne dépend pas du point M_0 .

2. Trouver tous les plans tangents à la surface (S) d'équation $x^2 - 4y^2 + z^2 = 1$ passant par des points $A(a, 0, 0)$, $B(0, b, 0)$ et $C(0, 0, c)$ tels que $c = 2a = 2b$ (on supposera a, b, c non nuls).

 **Solution:**

L'équation de (S) s'écrit $f(x, y, z) = 0$ avec $f(x, y, z) = x^2 - 4y^2 + z^2 - 1$. Pour tout $(x, y, z) \in \mathbb{R}^3$, on a

$$\frac{\partial f}{\partial x}(x, y, z) = 2x, \quad \frac{\partial f}{\partial y}(x, y, z) = -8y \quad \text{et} \quad \frac{\partial f}{\partial z}(x, y, z) = 2z$$

donc tout point de (S) autre que O est régulier. L'équation du plan tangent en un tel point M_0 de coordonnées (x_0, y_0, z_0) avec $(x_0, y_0, z_0) \neq (0, 0, 0)$ est alors

$$x_0(x - x_0) - 4y_0(y - y_0) + z_0(z - z_0) = 0$$

soit, puisque M_0 appartient à (S) , donc vérifie $x_0^2 - 4y_0^2 + z_0^2 = 1$:

$$xx_0 - 4yy_0 + zz_0 = 1.$$

Ce plan passe par A si et seulement si $ax_0 = 1$, soit $x_0 = \frac{1}{a}$. De même, il passe par B si et seulement si $y_0 = -\frac{1}{4b}$ et par C si et seulement si $z_0 = \frac{1}{c}$. Cela n'est possible que si x_0, y_0 et z_0 sont tous non nuls, et dans ce cas, la condition imposée par l'énoncé s'écrit

$$\frac{1}{z_0} = \frac{2}{x_0} = -\frac{1}{2y_0} \quad \text{soit} \quad x_0y_0 = 2y_0z_0 = -\frac{1}{2}x_0z_0$$

x_0, y_0, z_0 ne pouvant être nuls, on obtient $z_0 = \frac{x_0}{2}$ et $y_0 = -\frac{x_0}{4}$, puis, en remplaçant dans l'équation de (S) : $x_0^2 - 4y_0^2 + z_0^2 = 1$, on obtient $x_0^2 = 1$. On peut donc conclure :

Les plans qui conviennent sont les plans tangents à (S) aux points de coordonnées $\left(1, -\frac{1}{4}, \frac{1}{2}\right)$ et $\left(-1, \frac{1}{4}, -\frac{1}{2}\right)$; ils ont pour équations $x + y + \frac{z}{2} = 1$ et $-x - y - \frac{z}{2} = 1$.

