C12-3 : Fonctions de plusieurs variables réelles (calcul différentiel : compléments)

I Fonctions de plusieurs variables à valeurs réelles

I.1 Opérations sur les fonctions de classe C^1

Soit $\mathscr{U}$ un ouvert d'un espace vectoriel de dimension finie E. Une combinaison linéaire, un produit de fonctions de classe C^1 sur $\mathscr{U}$ à valeurs réelles est de classe C^1 ; donc

> **Proposition :** L'ensemble $C^1(\mathscr{U}, \mathbf{R})$ des applications de classe C^1 sur $\mathscr{U}$ à valeurs réelles est une $\mathbf{R}$-algèbre.

Remarque : pour une fonction à valeurs réelles, les dérivées partielles sont des réels ; si f et g sont des applications de classe C^1 sur $\mathscr{U}$ à valeurs réelles, si les $\dfrac{\partial}{\partial x_i}$ sont les dérivations partielles relatives à une base de E, alors

$$\frac{\partial(fg)}{\partial x_i} = f\frac{\partial g}{\partial x_i} + g\frac{\partial f}{\partial x_i} .$$

I.2 Formes linéaires sur un espace euclidien

Soit $(E, (.|.))$ un espace euclidien. Pour tout élément a de E, on définit

$$\begin{array}{rccc} \phi_a : & E & \longrightarrow & \mathbf{R} \\ & x & \longmapsto & (a|x) \end{array}$$

ϕ_a est une application linéaire, c'est donc un élément de E^*.
L'application $a \longmapsto \phi_a$ est un isomorphisme de E sur E^*, appelé isomorphisme canonique de E sur E^*.
On utilise en général ce résultat sous la forme suivante :

Proposition : Soit ϕ une forme linéaire non nulle sur l'espace euclidien E.
Il existe alors un unique élément a de E tel que

$$\forall x \in E \quad (a|x) = \phi(x)$$

Ce résultat est à rapprocher des considérations sur les hyperplans, qui en général sont les noyaux des formes linéaires non nulles, et dans les espaces euclidiens se caractérisent par leurs vecteurs normaux.

I.3 Gradient

Soit f une application de classe C^1 définie sur un ouvert $\mathcal{U}$ d'un espace euclidien E, à valeurs dans $\mathbf{R}$.

a. Définition

Pour tout élément a de $\mathcal{U}$, $\mathrm{d}f(a)$ est une application linéaire de E dans $\mathbf{R}$, donc un élément de E^*. On sait qu'il existe un unique élément de E, appelé **gradient** de f en a et noté $\mathrm{grad} f(a)$ ou $\nabla f(a)$, tel que, pour tout vecteur h,

$$\mathrm{d}f(a)(h) = \big(\mathrm{grad} f(a)|h\big) = \big(\nabla f(a)|h\big)\ .$$

Autrement dit,

$$\begin{aligned} f(a+h) &= f(a) + \big(\mathrm{grad} f(a)|h\big) + \underset{h\to 0}{\mathrm{o}}(h) \\ &= f(a) + \big(\nabla f(a)|h\big) + \underset{h\to 0}{\mathrm{o}}(\|h\|) \end{aligned}$$

b. Composantes en base orthonormale

Soit $\mathscr{B} = (e_1, \dots, e_n)$ une base orthonormale de E ; on note $\dfrac{\partial}{\partial x_i}$ les dérivations partielles relatives à cette base. On a vu que, si $h = \sum_{i=1}^{n} h_i e_i$,
$\mathrm{d}f(a)(h) = \sum_{i=1}^{n} h_i \dfrac{\partial f}{\partial x_i}(a)$, ce qui montre que

$$\nabla f(a) = \operatorname{grad} f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\ e_i\ .$$

En particulier, si $E = \mathbf{R}^n$ et si $\mathscr{B}$ est la base canonique,

$$\nabla f(a) = \operatorname{grad} f(a) = \Big(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\Big).$$

c. Interprétation

Le gradient indique la direction de plus grande variation de f : le vecteur unitaire v pour lequel $D_v f(a)$ est maximale est

$$v = \frac{1}{\|\nabla f(a)\|} \nabla f(a)$$

II Dérivées partielles d'ordre ≥ 2

II.1 Définition

Soit f une fonction définie sur un ouvert $\mathscr{U}$ d'un evn de dimension finie E, à valeurs dans un evn de dimension finie F.
Munissons E d'une base $(e_1, \ldots, e_n)$, et notons $\dfrac{\partial}{\partial x_i}$ les dérivations partielles correspondantes. Les dérivées partielles définies précédemment seront dorénavant appelées dérivées partielles « premières ». On définit les dérivées partielles secondes comme les dérivées partielles des dérivées partielles premières... lorsqu'elles existent.
Plus précisément : si $\dfrac{\partial f}{\partial x_i}$ est dérivable en a suivant le j-ème vecteur de la base $(e_1, \ldots, e_n)$, on définit

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

On dira que f est de classe C^2 sur $\mathscr{U}$ lorsque toutes ses dérivées partielles secondes $\dfrac{\partial^2 f}{\partial x_j \, \partial x_i}$ $(1 \leq i, j \leq n)$ sont définies et continues sur $\mathscr{U}$.
On définit de même par récurrence les dérivées partielles d'ordre $m \geq 3$ et la classe C^m.
Une fonction est de classe C^∞ lorsqu'elle est de classe C^m pour tout m.

II.2 Théorème de Schwarz

Théorème Si f est de classe C^2 sur $\mathscr{U}$, pour tout couple (i, j) on a, sur $\mathscr{U}$

$$\frac{\partial^2 f}{\partial x_j \, \partial x_i} = \frac{\partial^2 f}{\partial x_i \, \partial x_j}$$

... et plus généralement, pour une fonction de classe C^m, les dérivées partielles m-ièmes « ne dépendent pas de l'ordre des dérivations »

Exercice : On considère la fonction p définie par

$$p(x, y) = xy\frac{x^2 - y^2}{x^2 + y^2}$$

si $(x, y) \neq (0,0)$, et $p(0,0) = 0$. Calculer $\dfrac{\partial f}{\partial y}(x,0)$ (on distinguera $x = 0$ et $x \neq 0$), $\dfrac{\partial f}{\partial x}(0, y)$ (on distinguera $y = 0$ et $y \neq 0$), puis $\dfrac{\partial^2 f}{\partial x\,\partial y}((0,0))$ et $\dfrac{\partial^2 f}{\partial y\,\partial x}((0,0))$.

II.3 Stabilité

Une composée d'applications de classe C^m est de classe C^m.
Une combinaison linéaire d'applications de classe C^m est de classe C^m. Pour des fonctions à valeurs numériques (à valeurs réelles), un produit de fonctions de classe C^m est de classe C^m. Ainsi, avec des notations évidentes, $C^m(\mathcal{U}, \mathbf{R})$ est une $\mathbf{R}$-algèbre ($1 \leq m \leq +\infty$).

III Equations aux dérivées partielles

Rien n'est à connaître au programme sur les équations aux dérivées partielles. Mais des exercices, surtout à l'oral, s'y intéressent. L'idée la plus fréquente est, par changement de variable (il faudra savoir calculer des dérivées partielles de fonctions composées), de se ramener à la résolution de l'équation la plus simple :

$$\frac{\partial f}{\partial x_1} = 0$$

Les fonctions qui vérifient cette équation sont, sur un ouvert **convexe**, les fonctions

$$(x_1, \ldots, x_p) \mapsto \phi(x_2, \ldots, x_p)$$

Un exemple : Soit $U = \mathbf{R}_*^+ \times \mathbf{R}$; utiliser le changement de variable

$$\psi \ : \ (x, y) \mapsto (y/x, x^2 + y^2)$$

pour résoudre (sur $\mathscr{U}$) l'équation aux dérivées partielles

$$x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) = \lambda f(x, y)$$

Que signifie « utiliser ψ pour résoudre » l'équation aux dérivées partielles ? On pose $\psi(x, y) = (u, v)$ (donc $u = y/x$ et $v = x^2 + y^2$), et on définit une nouvelle fonction par

$$f(x, y) = g(u, v)$$

Plus précisément, on définit g sur $\mathbf{R}_*^+ \times \mathbf{R}_*^+$ par $g = f \circ \psi^{-1}$. Si f est de classe C^1 (on le suppose implicitement dans l'énoncé, il faudrait le dire mais on ne le dit pas...), g l'est aussi. L'explicitation de ψ^{-1} est peu simple, on part donc de

$$f(x, y) = g(y/x, x^2 + y^2)$$

et on obtient :

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

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

l'équation aux dérivées partielles

$\forall (x, y) \in \mathscr{U} \quad x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) = \lambda f(x, y)$

est donc équivalente à l'équation aux dérivées partielles

$\forall (u, v) \in \mathbf{R}_*^+ \times \mathbf{R}_*^+$

C'est là le but d'un changement de variables dans une EDP : ne garder de dérivée partielle que par rapport à une seule variable, et ainsi se ramener à un problème élémentaire (voir introduction) ou, comme ici, à une équation différentielle ordinaire.

Le programme précise que les changements de variables que l'on doit savoir essayer sont les changements de variables affines ou le changement de variables en polaires.

Remarque : Soit f définie sur $\mathbf{R}^2 \setminus \{(x,0)\ ;\ x \leq 0\}$ par
si $x > 0$, $f(x, y) = 0$
si $x \leq 0$ et $y > 0$, $f(x, y) = x^2$
si $x \leq 0$ et $y < 0$, $f(x, y) = -x^2$
Montrer qu'en tout point de son ouvert de définition, $\dfrac{\partial f}{\partial y}$ est nulle. Conclusion ?

IV Condition nécessaire d'extremums; points critiques

Soit f une application de classe C^1 sur un ouvert $\mathscr{U}$ de l'espace vectoriel normé E, à valeurs réelles.

Définition : On dira que f atteint un maximum (local) en $a \in \mathscr{U}$ lorqu'il existe un voisinage V de a dans $\mathscr{U}$ tel que

$$\forall x \in V \quad f(x) \leq f(a)$$

Remarque : Un voisinage de a dans $\mathscr{U}$ est, comme $\mathscr{U}$ est ouvert, un voisinage de a inclus dans $\mathscr{U}$.
Remarque : On peut remplacer « un voisinage V de a dans $\mathscr{U}$ »par « une boule ouverte $B(a, r) \subset \mathscr{U}$ $(r > 0)$ ».
On définit de même un minimum local. On dit que le maximum est global lorsque

$$\forall x \in \mathscr{U} \quad f(x) \leq f(a)\ .$$

Lorsqu'on parle d'extremum sans précision, on parle d'extremum local.

Proposition : Si f admet des dérivées partielles (suivant une base quelconque) en a, et si f atteint un extremum en a, alors ces dérivées partielles en a sont nulles.

Définition : Soit f une application définie sur un ouvert $\mathscr{U}$ de E, à valeurs réelles. On suppose f différentiable en a. On dit que a est un point critique de f lorsque la différentielle de f en a est nulle, i.e. lorsque les dérivées de f en a suivant tous les vecteurs sont nulles, i.e. lorsque toutes les dérivées partielles de f en a (suivant une base quelconque) sont nulles.

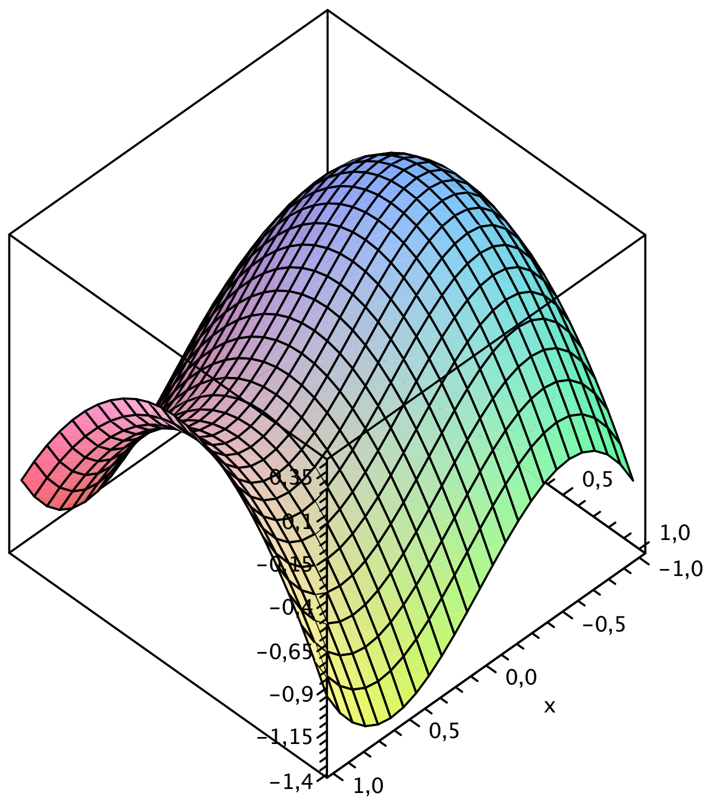
Proposition : Soit f une application définie sur un ouvert $\mathscr{U}$ de E, à valeurs réelles. Si f atteint un extremum en a, alors a est un point critique pour f.

Remarque : On n'a ici qu'une condition nécessaire d'extremum.

Remarque : Si $\mathscr{U}$ n'est pas ouvert, on peut avoir des extremums en des points non crititques.

Exemple : Déterminer les points critiques de

$$f : (x, y) \mapsto x^3 - y^2 - x$$



V Formule de Taylor-Young à l'ordre 2; application aux extremums

V.1 Matrice hessienne

Ce qu'on appelle $\mathscr{U}$ est toujours un ouvert de $\mathbf{R}^n$.

Définition Soit $f : \mathscr{U} \subset \mathbf{R}^n \longrightarrow \mathbf{R}$, de classe C^2. Pour tout $a \in \mathscr{U}$ on définit la matrice hessienne de f en a :

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right)_{1\leq i,j\leq n}$$

D'après le théorème de Schwarz, pour tout $a \in \mathscr{U}$, $H_f(a) \in \mathscr{S}_n(\mathbf{R})$.
Si $n = 2$, avec des notations habituelles,

$$H_f(a) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a) & \dfrac{\partial^2 f}{\partial y \partial x}(a) \\ \dfrac{\partial^2 f}{\partial x \partial y}(a) & \dfrac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}$$

V.2 La formule de Taylor-Young à l'ordre 2

Proposition : Soit $f : \mathscr{U} \subset \mathbf{R}^n \to \mathbf{R}$, de classe C^2. On suppose $\mathscr{U}$ ouvert. Soit $a \in \mathscr{U}$. On a alors, au voisinage de $0_{\mathbf{R}^n}$ (pour h) :

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

Que l'on peut également écrire, confondant comme d'habitude éléments de $\mathbf{R}^n$ et éléments de $\mathscr{M}_{n,1}(\mathbf{R})$:

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

Démonstration Non exigible, pas inintéressante, faisons-la pour $n = 2$ ce qui supprime les indexations.
Pour développer

$$f(a+h) - f(a)$$

on introduit la fonction

$$\phi\ :\ t \mapsto f(a+th)$$

qui permettra d'écrire $f(a+h)-f(a)=\phi(1)-\phi(0)$.

Pour que ce soit défini sur $[0,1]$, on remarque que, $\mathscr{U}$ étant ouvert, il existe une boule ouverte $B(a,r)$ $(r>0)$ incluse dans $\mathscr{U}$. On suppose dans la suite $\|h\|<r$, ce qui permet de définir ϕ sur $[0,1]$. Par théorème de composition, ϕ est de classe C^1 sur $[0,1]$, et

$$\forall t \in [0,1] \qquad \phi'(t) = df(a+th)(h)$$

ou encore

$$\forall t \in [0,1] \qquad \phi'(t) = h_1 \frac{\partial f}{\partial x}(a+th) + h_2 \frac{\partial f}{\partial y}(a+th)$$

Mais par hypothèse, $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont de classe C^1, donc par composition $t \mapsto \dfrac{\partial f}{\partial x}(a+th)$ et $t \mapsto \dfrac{\partial f}{\partial y}(a+th)$ sont de classe C^1. Et, avec la même formule qui a permis de dériver ϕ, on obtient :

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial f}{\partial x}(a+th)\right] &= h_1 \frac{\partial\left(\frac{\partial f}{\partial x}\right)}{\partial x}(a+th) + h_2 \frac{\partial\left(\frac{\partial f}{\partial x}\right)}{\partial y}(a+th)\\
&= h_1 \frac{\partial^2 f}{\partial x^2}(a+th) + h_2 \frac{\partial^2 f}{\partial y \partial x}(a+th)
\end{aligned}$$

et, de même,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial f}{\partial y}(a+th)\right] = h_1 \frac{\partial^2 f}{\partial x \partial y}(a+th) + h_2 \frac{\partial^2 f}{\partial y^2}(a+th)$$

On en déduit que ϕ est de classe C^2 sur $[0,1]$, et

$$\begin{aligned}
\forall t \in [0,1]\ \phi''(t) = h_1 &\left[h_1 \frac{\partial^2 f}{\partial x^2}(a+th) + h_2 \frac{\partial^2 f}{\partial y \partial x}(a+th)\right]\\
&+ h_2\left[h_1 \frac{\partial^2 f}{\partial x \partial y}(a+th) + h_2 \frac{\partial^2 f}{\partial y^2}(a+th)\right]
\end{aligned}$$

Appliquons le théorème de Schwarz :

$$\forall t \in [0,1]\ \phi''(t) = h_1^2 \frac{\partial^2 f}{\partial x^2}(a+th) + 2h_1 h_2 \frac{\partial^2 f}{\partial y \partial x}(a+th) + h_2^2 \frac{\partial^2 f}{\partial y^2}(a+th)$$

Ecrivons maintenant la formule de Taylor avec reste-intégrale :

$$\begin{aligned}
\phi(1) &= \phi(0) + (1-0)\phi'(0) + \int_0^1 (1-t)\phi''(t)\,dt \\
&= \phi(0) + \phi'(0) + \int_0^1 (1-t)\phi''(0)\,dt + \int_0^1 (1-t)\big[\phi''(t) - \phi''(0)\big]dt \\
&= \phi(0) + \phi'(0) + \frac{1}{2}\phi''(0) + \int_0^1 (1-t)\big[\phi''(t) - \phi''(0)\big]dt
\end{aligned}$$

ce qui donne, en réutilisant la fonction f,

$$\begin{aligned}
f(a+h) = f(a) + \big[h_1 \frac{\partial f}{\partial x}(a) + h_2 \frac{\partial f}{\partial y}(a)\big] + \frac{1}{2}\big[h_1^2 \frac{\partial^2 f}{\partial x^2}(a) \\
+ 2h_1 h_2 \frac{\partial^2 f}{\partial y \partial x}(a) + h_2^2 \frac{\partial^2 f}{\partial y^2}(a)\big] + R_a(h)
\end{aligned}$$

où

$$\begin{aligned}
R_a(h) = \int_0^1 (1-t)\Big[h_1^2\big[\frac{\partial^2 f}{\partial x^2}(a+th) - \frac{\partial^2 f}{\partial x^2}(a)\big] + 2h_1 h_2\big[\frac{\partial^2 f}{\partial y \partial x}(a+th) - \frac{\partial^2 f}{\partial y \partial x}(a)\big] \\
+ h_2^2\big[\frac{\partial^2 f}{\partial y^2}(a+th) - \frac{\partial^2 f}{\partial y^2}(a)\big]\Big]dt
\end{aligned}$$

Soit $\epsilon > 0$. Par continuité des applications $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial y \partial x}$ et $\frac{\partial^2 f}{\partial y^2}$, il existe $\eta > 0$ tel que

$$\|z-a\| \leq \eta \Rightarrow \begin{cases} |\frac{\partial^2 f}{\partial x^2}(z) - \frac{\partial^2 f}{\partial x^2}(a)| & \leq \epsilon \\ |\frac{\partial^2 f}{\partial y \partial x}(z) - \frac{\partial^2 f}{\partial y \partial x}(a)| & \leq \epsilon \\ |\frac{\partial^2 f}{\partial y^2}(z) - \frac{\partial^2 f}{\partial y^2}(a)| & \leq \epsilon \end{cases}$$

Si $\|h\| \leq \eta$, on a, pour tout $t \in [0,1]$, $\|(a+th) - a\| \leq \eta$. On peut supposer $\eta < r$ (quitte à diminuer η) et on obtient :

$$|R_a(h)| \leq \epsilon \int_0^1 (1-t)[h_1^2 + 2|h_1||h_2| + h_2^2]dt$$

Fixons $\|.\|$ égale à la norme $\|.\|_1$ usuelle, la plus adaptée ici; on vient donc de montrer

$$\|h\|_1 \leq \eta \implies |R_a(h)\| \leq \frac{\epsilon}{2}\|h\|_1^2$$

On adonc bien démontré que

$$R_a(h) = \underset{h\to 0}{\mathrm{o}}\left(\|h\|^2\right)$$

ce qui conclut.

V.3 Application à l'étude d'extremums

a. Etude

Soit $f : \mathscr{U} \subset \mathbf{R}^n \to \mathbf{R}$, de classe C^2. Soit $a \in \mathscr{U}$. On connaît déjà une condition nécessaire d'extremum en a :

$$\nabla f(a) = 0$$

(ou, de manière équivalente, toutes les dérivées partielles premières de f sont nulles en a) (autrement dit, a est un point critique).

On suppose dorénavant cette condition vérifiée.

On a alors, au voisinage de $(0,0)$ (pour h) :

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

où $Q_a(h) = h^T H_f(a)h$ (ainsi appelée car c'est une « forme quadratique », terminologie hors-programme).

1. Si $H_f(a) \in S_n^{++}(\mathbf{R})$, alors l'application

$$(x,y) \longmapsto x^T H_f(a)y$$

   est un produit scalaire sur $\mathbf{R}^n$. La norme euclidienne associée est $\sqrt{Q_a}$. Et l'équivalence des normes (nous sommes bien en dimension finie) permet d'affirmer

$$f(a+h) = f(a) + \frac{1}{2}Q_a(h) + \underset{h\to 0}{\mathrm{o}}(Q_a(h))$$

D'où l'équivalent, au voisinage de $h = 0_{\mathbf{R}^n}$:

$$f(a+h) - f(a) \sim \frac{1}{2} Q_a(h)$$

Deux termes équivalents ont strictement même signe au voisinage du point considéré, on peut donc affirmer que f atteint un minimum local strict en a.

2. Si $H_f(a) \not\in S_n^+(\mathbf{R})$, soit $\lambda < 0$ et $u \in \mathbf{R}^n$ tel que $H_f(a)u = \lambda u$. Appliquons la formule de Taylor-Young à $h = tu$, faire tendre h vers $0_{\mathbf{R}^n}$ revient à faire tendre t vers 0. On conclut que f n'atteint pas un minimum local en a.

b. Résultats

Proposition Soit f de classe C^2 sur un ouvert $\mathscr{U}$ de $\mathbf{R}^n$, à valeurs dans $\mathbf{R}$. Si f atteint un minimum (local) en a, alors a est point critique de f et $H_f(a) \in \mathscr{S}_n^+(\mathbf{R})$.
Soit f de classe C^2 sur un ouvert $\mathscr{U}$ de $\mathbf{R}^n$, à valeurs dans $\mathbf{R}$. Si a est point critique de f et $H_f(a) \in \mathscr{S}_n^{++}(\mathbf{R})$ alors f atteint un minimum local strict en a.

On voit sans peine les modifications à faire pour remplacer « minimum » par « maximum ».

c. Cas $n = 2$

On a

$$H_f(a) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

où $r = \dfrac{\partial^2 f}{\partial x^2}(a)$, $s = \dfrac{\partial^2 f}{\partial y \partial x}(a) = \dfrac{\partial^2 f}{\partial x \partial y}(a)$ (théorème de Schwarz) et $t = \dfrac{\partial^2 f}{\partial y^2}(a)$.
On suppose dans la suite que a est critique.

1. Si $\det\big(H_f(a)\big) = rt - s^2 > 0$

Alors $H_f(a) \in \mathscr{S}_n^{++}(\mathbf{R})$ (si $\mathrm{Tr}\big(H_f(a)\big)\, r+t>0$) ou $-H_f(a) \in \mathscr{S}_n^{++}(\mathbf{R})$ (si $r+t<0$). Il y a donc

2. **Si** $r\,t - s^2 < 0$

   Ici, il n'y a pas d'extremum en a, on a un « point col » (terminologie h.p.).

3. Si $r\,t - s^2 = 0$

   Il peut se passer beaucoup de choses ! la simple considération des fonctions f et $-f$, où $f(x,y) = x^4 + y^4$, en $(0,0)$, montre qu'on ne peut rien dire de général.

Conclusion

Soit $f\ :\ \mathscr{U} \subset \mathbf{R}^2 \to \mathbf{R}$, de classe C^2. Soit $a \in \mathscr{U}$ un point critique ($\nabla f(a) = 0$). Notons

$$r = \frac{\partial^2 f}{\partial x^2}(a) \quad , \quad s = \frac{\partial^2 f}{\partial y \partial x}(a) \quad , \quad t = \frac{\partial^2 f}{\partial y^2}(a)$$

Alors :

— Si $r\,t - s^2 > 0$, f atteint en a un extremum local strict.

— Si $r\,t - s^2 < 0$, f n'atteint pas un extremum en a.

— Si $r\,t - s^2 = 0$, on ne peut rien dire.

VI Vecteurs tangents à une partie d'un espace vectoriel normé

VI.1 Définition

Pour comprendre la définition suivante, imaginer le cas où X est une surface, E étant un espace euclidien de dimension 3.

Définition Soit X une partie d'un $\mathbf{R}$-espace vectoriel normé de dimension finie E, soit $x \in X$. On dit que $v \in E$ est tangent à X en x lorsqu'il existe $\epsilon > 0$ et

$$\begin{array}{cccc} \gamma : &]-\epsilon,\epsilon[& \longrightarrow & E \\ & t & \longmapsto & \gamma(t) \end{array}$$

vérifiant les propriétés suivantes :

(i) $\forall t \in]-\epsilon,\epsilon[\quad \gamma(t) \in X$

(ii) $\gamma(0) = x$

(iii) γ est dérivable en 0, et $\gamma'(0) = v$

Traduction $v \in E$ est tangent à X en x lorsqu'il existe un « petit bout d'arc » tracé sur X, passant par x, et dont le vecteur vitesse en x est v.

Notation On note T_xX l'ensemble des vecteurs tangents à X en x.

Exercice Montrer que si $v \in E$ est tangent à X en x, alors αv est aussi tangent en x à X ($\alpha \in \mathbf{K}$).

VI.2 Un exemple peu géométrique

Exercice : Quels sont les vecteurs tangents en I_n à $SO(n)$? *Indication : que peut-on dire de l'exponentielle d'une matrice antisymétrique réelle ?*

VI.3 Les exemples à connaître

Ces exemples sont dans le programme. On est donc autorisé à les utiliser, mais il faut s'attendre à devoir les redémontrer.

a. Sous -espace affine

Ici, $X = a + F$ où F est un sous-espace vectoriel de E. Soit $x \in X$. On a $T_x X =$

b. Sphère

Ici, $X = \{x \in E \,;\ \|x - a\| = R\}$ où $R > 0$ et $\|.\|$ est une norme euclidienne sur E (qui est donc un espace euclidien). Alors $T_x X =$

c. Courbe $y = f(x)$

On considère une fonction f de classe C^1 sur un intervalle ouvert I de $\mathbf{R}$. Et X est son graphe :

$$X = \{(x, y) \in \mathbf{R}^2 \,;\ y = f(x)\}$$

Soit $a = (\alpha, \beta) \in X$. Alors $T_a X =$

d. Surface d'équation $z = f(x, y)$

On considère une fonction f de classe C^1 sur un ouvert $\mathscr{U}$ de $\mathbf{R}^2$. Et X est son graphe :

$$X = \{(x, y, z) \in \mathscr{U} \times \mathbf{R} \,;\ z = f(x, y)\}$$

Soit $a = (\alpha, \beta, \gamma) \in X$. Alors $T_a X =$

VI.4 Cas particulier d'une partie définie par une équation

a. Le résultat

Proposition Soit $g : \mathscr{U} \subset E \to \mathbf{R}$, où $\mathscr{U}$ est un ouvert de l'espace vectoriel normé de dimension finie E. On suppose g de classe C^1. Soit

$$X = \{x \in \mathscr{U} \ ; \ g(x) = k\}$$

où k est un réel fixé. Alors, pour tout $x \in X$ tel que $\mathrm{d}g \neq 0$, $T_x X = \mathrm{Ker}\left(\mathrm{d}g(x)\right)$. Si E est euclidien, cela signifie que $T_x X = \left(\mathrm{Vect}\left(\nabla g(x)\right)\right)^{\perp}$.

Démonstration Contentons-nous de montrer que les vecteurs tangents à X en x_0 sont orthogonaux à $\nabla g(x_0)$. (La démonstration générale du résultat est hors-programme pour d'excellentes raisons)

Exemple On retrouve les exemples cités plus haut.

b. Tangente, normale à une courbe $F(x, y) = k$

Soit F de classe C^1 sur un ouvert $\mathscr{U}$ de $\mathbf{R}^2$, à valeurs réelles, et $(x_0, y_0) \in \mathscr{U}$ tel que $F(x_0, y_0) = k$. Soit

$$(C) = \{(x, y) \in \mathscr{U} \ ; \ F(x, y) = k\}$$

Le résultat précédent montre que

Proposition Si $\nabla F(x_0, y_0) \neq \overrightarrow{0}$, il dirige la normale à (C) en (x_0, y_0).

On peut appliquer cela à $F(x, y) = y - f(x)$ où f est une fonction C^1 sur un intervalle ouvert I.

c. Normale, plan tangent à une surface $F(x, y, z) = k$

Ici, F est définie sur un ouvert $\mathscr{U}$ de $\mathbf{R}^3$, de classe C^1. On obtient, si $F(x_0, y_0, z_0) = k$:

Proposition Si $\nabla F(x_0, y_0, z_0) \neq \overrightarrow{0}$, il dirige la normale à la surface d'équation $F(x, y, z) = k$ en (x_0, y_0, z_0).

Cette **normale** est une droite ; le **plan tangent** est le plan contenant (x_0, y_0, z_0) et orthogonal à cette normale. Son équation est donc

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

Ici aussi, l'exemple de la sphère : $x^2+y^2+z^2 = R^2$ permet de retrouver le résultat. Il est intéressant d'examiner le cas particulier des surfaces $z = f(x, y)$.

VII Optimisation sous contrainte

Théorème Soit E un $\mathbf{R}$-espace vectoriel normé de dimension finie, $\mathscr{U}$ un ouvert de E, f et g deux fonctions C^1 sur $\mathscr{U}$, à valeurs réelles. On pose

$$X = \{x \in \mathscr{U} \ ; \ g(x) = 0\}$$

On suppose que la restriction $g|_X$ de g à X atteint un extremum (local) en $x_0 \in X$. Alors, si $\mathrm{d}g(x_0) \neq 0$, il existe $\lambda \in \mathbf{R}$ tel que $\mathrm{d}f(x_0) = \lambda \mathrm{d}g(x_0)$.

Dérivées partielles En termes de dérivées partielles relatives à une bas quelconque, la conclusion s'écrit : il existe $\lambda \in \mathbf{R}$ tel que $\forall i \quad \partial_i f = \lambda \partial_i g$ (et l'hypothèse : il existe i tel que $\partial_i g(x_0) \neq 0$).

Gradient Si E est euclidien, la conclusion s'écrit : il existe $\lambda \in \mathbf{R}$ tel que $\nabla f(x_0) = \lambda \nabla g(x_0)$ (et l'hypothèse : $\nabla g(x_0) \neq 0$).

VIII Matrice jacobienne

VIII.1 Définition

Soit E, F deux espaces vectoriels réels de dimension finie, f une application de classe C^1 d'un ouvert $\mathscr{U}$ de E dans F. Soit $\mathscr{B} = (\epsilon_1, \ldots, \epsilon_p)$ une base de E,

$\mathscr{C} = (e_1, \ldots, e_n)$ une base de F. On appelle **matrice jacobienne** de f en a relativement aux bases $\mathscr{B}$ et $\mathscr{C}$ la matrice de l'application linéaire $\mathrm{d}f(a)$ dans les bases $\mathscr{B}$ et $\mathscr{C}$.

En notant $f_1, \ldots, f_n$ les applications composantes de f dans la base $\mathscr{C}$ (cést-à-dire, pour tout élément x de $\mathscr{U}$, $f(x) = \sum_{i=1}^{n} f_i(x) e_i$), et en notant ∂_j ou $\dfrac{\partial}{\partial x_j}$ la $j$ème dérivation partielle dans la base $\mathscr{B}$, on a

$$\mathrm{J}f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_p}(a) \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \cdots & \dfrac{\partial f_n}{\partial x_p}(a) \end{pmatrix} = \left(\dfrac{\partial f_i}{\partial x_j}(a) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} .$$

VIII.2 Exemples

Calculer les matrices jacobiennes (relatives aux bases canoniques) des applications

$$(\rho, \theta) \mapsto (\rho \cos\theta, \rho \sin\theta)$$

et

$$(r, \phi, \theta) \mapsto (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta)$$

VIII.3 Matrice jacobienne d'une application composée

Soit E, F, G trois espaces vectoriels réels de dimension finie, f une application de classe C^1 d'un ouvert $\mathscr{U}$ de E dans F, g une application de classe C^1 d'un ouvert $\mathscr{V}$ de F dans G. On suppose que $f(\mathscr{U}) \subset \mathscr{V}$. Soit $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$, des bases respectives de E, F, G.

$$\mathrm{J}_{\mathscr{B},\mathscr{D}}(\mathrm{g} \circ f)(a) = \mathrm{J}_{\mathscr{C},\mathscr{D}}\, g\big(f(a)\big) J_{\mathscr{B},\mathscr{C}} f(a) \ .$$

ou encore, si les dimensions respectives de E, F, G sont p, n, q, alors, pour tout i entre 1 et q et tout j entre 1 et p,

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{n} \frac{\partial g_i}{\partial y_k}\big(f(a)\big)\frac{\partial f_k}{\partial x_j}(a) \ .$$

(On désigne par $\dfrac{\partial}{\partial y_k}$ les dérivées partielles dans la base $\mathscr{C}$).
Cette formule (un peu compliquée à écrire dans le cas général, et appelée règle de la chaîne) permet le calcul des dérivées partielles d'une fonction composée. Mais la recette a déjà été vue.

Table des matières