CH20 : Calcul différentiel - aspects appliqués

Table des matières

Dans ce dernier chapitre, les fonctions considérées sont définies sur un ouvert Ω de E (un $\mathbb{R}$-evn de dimension finie) et à valeurs dans $\mathbb{R}$.

I Vecteurs tangents à une partie

Définition 1 (Vecteur tangent à une partie)
Soient X une partie de E et $x \in X$.
On dit qu'un vecteur $v \in E$ est **tangent à X en x** *lorsqu'il existe $\varepsilon > 0$ et une application $\gamma :]-\varepsilon, \varepsilon[\to X$ dérivable en 0 tels que $\gamma(0) = x$ et $\gamma'(0) = v$.*

Vocabulaire
En général, une application $\gamma : I \subset \mathbb{R} \to E$ est appelée un **arc** (ou **arc paramétré**).
Un vecteur $v \in E$ est donc tangent à X en x lorsqu'il existe un arc tracé dans X qui passe par x, avec un vecteur vitesse égal à v.

Notation
On notera $T_x X$ l'ensemble des vecteurs tangents à X en x.

Exemple (Sous-espace affine)
Si X est un sous-espace affine de direction F (un sev de E), alors pour tout $x \in X$, on a

$$T_x X = F.$$

En effet, si $v \in F$, alors en posant $\gamma : t \mapsto x + tv$, on a $\gamma(t) \in X$ pour tout $t \in \mathbb{R}$ (car $X = x + F$), et $\gamma(0) = x$, $\gamma'(0) = v$. Donc $F \subset T_x X$.
Réciproquement, si $v \in T_x X$, alors il existe $\gamma :]-\varepsilon, \varepsilon[\to X$ telle que $\gamma(0) = x$, $\gamma'(0) = v$, donc

$$v = \lim_{t \to 0} \frac{\gamma(t) - \gamma(0)}{t} \in F,$$

Mais pour tout $t \in]-\varepsilon, \varepsilon[$, $\gamma(t) \in X = \gamma(0) + F$, donc $\frac{\gamma(t)-\gamma(0)}{t} \in F$, et comme F est fermé (en tant que sev de dimension finie), la limite v reste dans F. D'où $T_x X \subset F$.

Exemple (Boule et sphère unité)
Soit E un espace euclidien. Montrer que :

1. pour tout $x \in \mathcal{B}(0,1)$, $T_x \mathcal{B}(0,1) = E$;
2. pour tout $x \in \mathcal{S}(0,1)$, $T_x \mathcal{S}(0,1) = \{x\}^{\perp}$.

Solution :

1. Par définition, $T_x \mathcal{B}(0,1) \subset E$; réciproquement, si $v \in E$, alors $\gamma(t) = x + tv$ est dans $\mathcal{B}(0,1)$ pour t assez petit (puisque la boule est ouverte et contient x), ce qui définit bien au voisinage de $t = 0$ un arc tracé dans $\mathcal{B}(0,1)$ tel que $\gamma(0) = x$ et $\gamma'(0) = v$ (une droite affine en fait). Donc $v \in T_x \mathcal{B}(0,1)$.
2. soit $x \in \mathcal{S}(0,1)$. S'il existe $\gamma :]-\varepsilon, \varepsilon[\to \mathcal{S}(0,1)$ tel que $\gamma(0) = x$ et $\gamma'(0) = v$, alors pour tout réel t :

   $$|t| < \varepsilon \implies \|\gamma(t)\| = 1 \implies (\gamma(t)|\gamma(t)) = 1,$$

   donc en dérivant cette égalité en $t = 0$, on obtient par bilinéarité du produit scalaire :

   $$2(\gamma'(0)|\gamma(0)) = 0,$$

   c'est-à-dire $(v|x) = 0$.
   Réciproquement, si $(v|x) = 0$, alors l'application

   $$\gamma : t \mapsto \frac{x + tv}{\|x + tv\|} = \frac{x + tv}{\sqrt{\|x\|^2 + 2t(v|x) + t^2\|v\|^2}} = \frac{x + tv}{\sqrt{1 + t^2\|v\|^2}}$$

   est dérivable sur $\mathbb{R}$, à valeurs dans $\mathcal{S}(0,1)$ et vérifie bien $\gamma(0) = x$ et $\gamma'(0) = v$ (faire un DL_1).

ATTENTION !
L'ensemble T_xX n'est pas toujours un sous-espace vectoriel de E ! Même si la plupart du temps, oui. Le théorème suivant donne des conditions sur X pour que T_x soit un sev de E.

Théorème 2 (Structure d'hyperplan tangent)
Soit $g : \Omega \subset E \to \mathbb{R}$ de classe $\mathcal{C}^1$. On considère "l'hypersurface" de E :

$$X = g^{-1}(\{0\}) = \{x \in \Omega,\ g(x) = 0\}.$$

Alors, pour tout $x \in X$ tel que $dg(x) \neq 0_{\mathcal{L}(E,\mathbb{R})}$, on a

$$T_xX = Ker(dg(x)).$$

Ainsi, l'ensemble des vecteurs tangents à X en x est un hyperplan de E.
De plus, si E est euclidien, alors

$$T_xX = \{\nabla g(x)\}^{\perp}.$$

Remarque

- *Ca marche aussi avec des parties du type*

$$X_\lambda = g^{-1}(\{\lambda\}) = \{x \in \Omega,\ g(x) = \lambda\}, \qquad \lambda \in \mathbb{R}$$

  (appelées **lignes de niveau de g***). En effet, les fonctions g et $g - \lambda$ ont même différentielle.*
- *On en déduit en particulier que* **∇g est orthogonal aux lignes de niveau de g** *(c'est-à-dire orthogonal aux vecteurs tangents aux lignes de niveau). C'est une autre propriété géométrique remarquable du gradient (en plus d'indiquer la direction de plus grande pente).*

Preuve (A moitié admise)
Soit $x \in X$ tel que la forme linéaire $dg(x)$ est non nulle (son noyau $Ker(dg(x))$ est donc un hyperplan de E).

- L'inclusion $T_xX \subset Ker(dg(x))$ est facile : si $v \in T_xX$, alors il existe $\gamma :]-\varepsilon, \varepsilon[\to X$ tel que $\gamma(0) = x$ et $\gamma'(0) = v$, donc pour tout réel t :

$$|t| < \varepsilon \implies \gamma(t) \in X \implies g(\gamma(t)) = 0,$$

  ce qui entraîne en dérivant en $t = 0$: $dg(x) \cdot v = 0$, i.e. $v \in Ker(dg(x))$.
- L'inclusion $Ker(dg(x)) \subset T_xX$ est difficile et hors programme (cela utilise le *théorème des fonctions implicites*, qui par exemple permet de transformer une "équation implicite" du type $g(x, y, z) = 0$ en une équation du type $z = f(x, y)$ **localement**).

Le dernier point résulte de la définition du gradient $\nabla g(x)$:

$$v \in Ker(dg(x)) \iff dg(x) \cdot v = 0 \iff (\nabla g(x)|v) = 0 \iff v \in \{\nabla g(x)\}^{\perp}.$$

Vocabulaire
Dans ce cas, on appelle en général **(hyper)plan tangent** à X en x l'hyperplan **affine**

$$\mathcal{T}_xX = x + T_xX = x + Ker(dg(x)) = x + \{\nabla f(x)\}^{\perp}.$$

Pour tout $m \in E$, on a

$$m \in \mathcal{T}_xX \iff m - x \in T_xX \iff (m - x|\nabla f(x)) = 0.$$

Exemple (Cas d'une surface du type $z = f(x, y)$)
Soit $f : \left\{ \begin{array}{ccc} U \subset \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array} \right.$ une fonction de classe $\mathcal{C}^1$.
On considère la "surface" $\mathcal{S}$ d'équation $z = f(x, y)$:

$$\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3,\ (x, y) \in U,\ z = f(x, y)\}.$$

Soit $M_0 = (x_0, y_0, z_0)$ un point de $\mathcal{S}$. Déterminer une équation cartésienne (dans le repère canonique de $\mathbb{R}^3$) du plan tangent à $\mathcal{S}$ en M_0 de deux façons :

1. *en utilisant le théorème précédent ;*
2. *"à la main".*

On trouve :

$$\mathcal{T}_{M_0}\mathcal{S} : \ z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \times (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \times (y - y_0)$$

(c'est la troncature de $f(x, y)$ à l'ordre 1).

Exemple (Cas de la sphère unité de $\mathbb{R}^3$)
Déterminer une équation cartésienne (dans le repère canonique de $\mathbb{R}^3$) du plan tangent à la sphère unité de $\mathbb{R}^3$:

$$\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3, \ x^2 + y^2 + z^2 = 1\}$$

en un point quelconque $M_0 = (x_0, y_0, z_0)$.

II Problèmes d'optimisation

E désigne toujours un $\mathbb{R}$-espace vectoriel normé de dimension finie (notée $n \in \mathbb{N}^*$), et Ω un ouvert de E.

1) Condition nécessaire d'extremum local

Définition 3 (Minimum, maximum local)
Soit $A \subset E$ (pas nécessairement un ouvert), soit $f : A \to \mathbb{R}$ et $a \in A$.
On dit que f possède un **minimum local** *(resp.* **maximum local***) en a lorsqu'il existe $r > 0$ tel que $\forall x \in B(a,r) \cap A,\ f(x) \geq f(a)$ (resp. $f(x) \leq f(a)$).*

Remarque
On peut remplacer la boule ouverte $B(a,r)$ par la boule fermée $\overline{B}(a,r)$ ou n'importe quel voisinage V de a dans E, cela ne change rien.

Vocabulaire
Ces deux termes sont regroupés sous la dénomination plus générale d'**extremum local**.
Le minimum (resp. maximum) est dit **global** lorsque l'inégalité $f(x) \geq f(a)$ (resp. $f(x) \leq f(a)$) est vraie pour tout $x \in A$.
Bien sûr, tout extremum global est en particulier un extremum local.

<u>ATTENTION !</u>
Ne pas confondre la valeur de l'extremum local ($m = f(a) \in \mathbb{R}$) avec le(s) point(s) où cette valeur est atteinte (les $x \in A$ tels que $f(x) = m$, il peut y en avoir plusieurs).

Remarque
Si A est compacte et $f : A \to \mathbb{R}$ est continue, alors f admet nécessairement un minimum global et un maximum global sur A (théorème des bornes atteintes).

Rappel (Résultats pour les fonctions $\mathbb{R} \to \mathbb{R}$)
Si $f : I \to \mathbb{R}$ est dérivable sur un intervalle I d'intérieur non vide, alors

$$\boxed{f \text{ admet un extremum local en } a \in \mathring{I} \implies f'(a) = 0}$$

(on rappelle que $\mathring{I}$ désigne l'intérieur de I).
Démonstration : supposons qu'il existe $r > 0$ tel que $\forall x \in I \cap]a-r, a+r[,\ f(x) \geq f(a)$ (cas du minimum local). Puisque $a \in \mathring{I}$, alors en diminuant suffisamment r, on a $]a-r, a+r[\subset I$, donc

$$\forall x \in]a-r, a+r[,\ f(x) \geq f(a).$$

Déterminons le signe de $f'(a) = \lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ (cette limite existe dans $\mathbb{R}$ par hypothèse).
On a $\forall x \in]a-r, a[,\ \frac{f(x)-f(a)}{x-a} \leq 0$, et donc $\lim\limits_{x \to a^-} \frac{f(x)-f(a)}{x-a} \leq 0$.
De même, en considérant les $x \in]a, a+r[$, on obtient $\lim\limits_{x \to a^+} \frac{f(x)-f(a)}{x-a} \geq 0$.
Puisque ces deux limites à gauche et à droite sont égales à $f'(a)$, on obtient $f'(a) \leq 0$ et $f'(a) \geq 0$, donc $f'(a) = 0$. Le cas du maximum local se traite de même.

Attention :

- Cette implication est fausse si $a \notin \mathring{I}$ (par ex : $f : x \mapsto x$ sur $I = [0,1]$ atteint son minimum en $x = 0$ mais $f'(0) = 1 \neq 0$).
- L'implication réciproque est fausse, même si $a \in \mathring{I}$ (par ex : $f : x \mapsto x^3$ sur $I = [-1,1]$ vérifie $f'(0) = 0$ mais n'admet pas d'extremum local en 0, puisque pour tout $r > 0$, $f(x) - f(0) = x^3$ change de signe sur $]-r, r[$).

Ce résultat se généralise aux fonctions de plusieurs variables :

Définition 4 (Point critique d'une application différentiable)
Soit Ω un ouvert de E, et $f : \Omega \to \mathbb{R}$ une fonction différentiable.
On appelle **point critique** *de f tout point $a \in \Omega$ tel que $df(a) = 0_{\mathcal{L}(E,\mathbb{R})}$.*

Remarque
La condition $df(a) = 0_{\mathcal{L}(E,\mathbb{R})}$ équivaut à $\frac{\partial f}{\partial x_j}(a) = 0$ pour tout $j \in [\![1, n]\!]$ (les dérivées partielles relativement à n'importe quelle base de E), ou encore si E est euclidien, à $\nabla f(a) = 0_E$.

Théorème 5 (Condition nécessaire d'extremum local)
Soit Ω un **ouvert** *de E, et $f : \Omega \to \mathbb{R}$ une fonction différentiable.*
Si f admet un extremum local en $a \in \Omega$, alors a est un point critique de f.

Preuve
Comme f est différentiable, ses dérivées directionnelles dans toutes les directions sont bien définies. Soit $h \in E$. Alors l'application $\varphi : t \mapsto f(a + th)$ est définie sur un voisinage $] - \varepsilon, \varepsilon[$ de $t = 0$ (car $a \in \Omega$ et Ω est ouvert, donc $a + th \in \Omega$ pour $|t|$ assez petit), est dérivable (car f est différentiable). Si f admet un minimum local en a, alors $\varphi(t) = f(a + th) \geq f(a) = \varphi(0)$ pour $|t|$ assez petit, donc φ admet un minimum local en 0, ce qui entraîne (d'après le rappel pour les fonctions réelles) que $\varphi'(0) = 0$, c'est-à-dire $D_h f(a) = 0$, ou encore $df(a) \cdot h = 0$, et ce pour tout $h \in E$, donc la forme linéaire $df(a)$ est nulle. Idem pour le cas du maximum local.

ATTENTION !
Comme pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$, la condition est nécessaire mais pas suffisante.

Exemple (Selle de cheval)
Etudier les extrema de $f : (x, y) \mapsto x^2 - y^2$ sur $\mathbb{R}^2$.
Il y a un seul point critique sur $\mathbb{R}^2$: $(0, 0)$, mais f n'admet pas d'extremum local en $(0, 0)$ (ni minimum ni maximum, on le voit en étudiant $f(x, 0) > 0$ et $f(0, y) < 0$ pour x et y non nuls), on dit que $(0, 0)$ est un **point col** (ou **point selle**).

Méthode
Dans la recherche des extrema locaux d'une fonction différentiable **sur un ouvert** *et à valeurs réelles, on commence donc par chercher les points critiques, puis on discute au cas par cas pour déterminer si ce sont réellement des extrema.*

Exemple
Etudier les extrema de $f : (x, y) \mapsto x^2 + xy + y^2 + 2x + 3y$ sur $\mathbb{R}^2$.

Remarque
Si l'on trouve un point critique $(x_0, y_0) \neq (0, 0)$, on peut utiliser le changement de variable $x = x_0 + h$, $y = y_0 + k$ avec h et k proches de 0 pour étudier le signe de $f(x, y) - f(x_0, y_0)$ au voisinage de (x_0, y_0).

ATTENTION !
Si on étudie les extrema d'une fonction $f : A \to \mathbb{R}$ avec A non ouvert, alors le lien entre extremum local et point critique ne concerne que les points **intérieurs** au domaine A (qui forment toujours un ouvert). Les points non intérieurs à A (donc situés sur la frontière) doivent donc être étudiés à part. Le cas typique est A compact, qu'on décompose en réunion disjointe :

$$A = \Omega \cup \mathcal{C},$$

avec $\Omega = \mathring{A}$ ouvert et $\mathcal{C} = Fr(A)$ (la "courbe" formée par la frontière de A).
Nous traiterons un exemple de telle situation plus loin.

2) Conditions d'ordre 2

On va maintenant donner des outils plus précis pour connaître la nature d'un point critique de f, c'est-à-dire s'il s'agit d'un minimum local, d'un maximum local, ou d'un point col. Pour cela, on va établir un développement limité à l'ordre 2 de f au voisinage du point critique.

Théorème 6 (Formule de Taylor-Young à l'ordre 2)
Soit Ω un ouvert de E et $f \in \mathcal{C}^2(\Omega, \mathbb{R})$. Pour tout $a \in \Omega$, on a dans toute base $\mathcal{B}_E$ de E :

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

où $(h_1, \cdots, h_n) = [h]_{\mathcal{B}_E}$.

Preuve
Pour $\|h\|$ suffisamment petit et $t \in [0,1]$, on a $a + th \in \Omega$ (car $a \in \Omega$ et Ω ouvert).
On applique la formule de Taylor avec reste intégral à la fontion réelle $\varphi : t \mapsto f(a+th)$ (de classe $\mathcal{C}^2$ comme f) entre les points $t = 0$ et $t = 1$:

$$f(a+h) - f(a) = \varphi(1) - \varphi(0) = \varphi'(0) + \int_0^1 (1-t)\varphi''(t)dt.$$

Or, pour tout $t \in [0,1]$, on obtient par dérivation composée :

$$\varphi'(t) = df(a+th) \cdot h = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a+th),$$

$$\varphi''(t) = \sum_{i=1}^{n} h_i \left(\sum_{j=1}^{n} h_j \frac{\partial^2 f}{\partial x_j \partial x_i}(a+th) \right) = \sum_{1 \leq i,j \leq n} h_i h_j \frac{\partial^2 f}{\partial x_j \partial x_i}(a+th).$$

Par continuité des dérivées partielles secondes de f en a, on peut écrire :

$$\forall (i,j) \in [\![1,n]\!]^2,\ \forall t \in [0,1], \qquad \frac{\partial^2 f}{\partial x_j \partial x_i}(a+th) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) + \varepsilon_{i,j}(th),$$

où $\lim\limits_{u \to 0_E} \varepsilon_{i,j}(u) = 0$. Donc par linéarité de l'intégrale :

$$\begin{aligned} f(a+h) - f(a) &= \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + \sum_{1 \leq i,j \leq n} h_i h_j \int_0^1 (1-t) \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) + \varepsilon_{i,j}(th) \right) dt \\ &= \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2} \sum_{1 \leq i,j \leq n} h_i h_j \frac{\partial^2 f}{\partial x_j \partial x_i}(a) + \sum_{1 \leq i,j \leq n} h_i h_j \int_0^1 (1-t)\varepsilon_{i,j}(th)dt. \end{aligned}$$

Reste à montrer que le dernier terme

$$R(h) = \sum_{1 \leq i,j \leq n} h_i h_j \int_0^1 (1-t)\varepsilon_{i,j}(th)dt$$

est bien négligeable devant $\|h\|^2$. Cela résulte de la majoration :

$$|R(h)| \leq \sum_{1 \leq i,j \leq n} |h_i h_j| \int_0^1 |\varepsilon_{i,j}(th)|dt \leq \|h\|^2 \sum_{1 \leq i,j \leq n} \int_0^1 |\varepsilon_{i,j}(th)|dt,$$

et pour tout couple (i,j), on a bien $\lim\limits_{h \to 0_E} \displaystyle\int_0^1 |\varepsilon_{i,j}(th)|dt = 0$ par le théorème de convergence dominée (qui s'applique car $\forall t \in [0,1]$, $\lim\limits_{h \to 0_E} \varepsilon_{i,j}(th) = 0$ et $\forall t \in [0,1]$, $\|th\| \leq \|h\|$ donc $|\varepsilon_{i,j}(th)| \leq 1$ pour $\|h\|$ assez petit (puisque $\varepsilon_{i,j}$ est de limite nulle en 0, et la fonction constante $t \mapsto 1$ est bien intégrable sur le segment $[0,1]$). Donc $\lim\limits_{h \to 0_E} \displaystyle\sum_{1 \leq i,j \leq n} \int_0^1 |\varepsilon_{i,j}(th)|dt = 0$, ce qui montre que $|R(h)| = o(\|h\|^2)$.

Exemple
Pour une fonction $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$, cette formule se réécrit :

$$f(x_0+h_1, y_0+h_2) = f(x_0,y_0) + h_1 \frac{\partial f}{\partial x}(x_0,y_0) + h_2 \frac{\partial f}{\partial y}(x_0,y_0)$$

$$+\frac{1}{2}\left(h_1^2 \frac{\partial^2 f}{\partial x^2}(x_0,y_0) + 2h_1h_2 \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + h_2^2 \frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right) + o(\|h\|^2).$$

Remarque
Si E est euclidien, alors le terme d'ordre 1 se réécrit à l'aide de $\nabla f(a)$:

$$\sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) = df(a)\cdot h = (\nabla f(a)|h).$$

Quant au terme d'ordre 2, on peut le voir comme un produit matriciel :

$$\sum_{1\leq i,j\leq n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = (h_1 \;\cdots\; h_n)\, H_f(a) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

où $H_f(a)$ est la matrice $n\times n$ des dérivées partielles secondes de f en a (relativement à la base choisie sur E). Cette matrice est donc intéressante à étudier.

Dans la suite, on se limite au cas de $E = \mathbb{R}^n$, muni de sa structure euclidienne canonique.

Définition 7 (Matrice hessienne)
Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On appelle **matrice hessienne** *de f au point $a \in \Omega$ la matrice*

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{(i,j)\in[1,n]^2} \in \mathcal{S}_n(\mathbb{R}).$$

Remarque
La symétrie de $H_f(a)$ vient du théorème de Schwarz (qui s'applique car f est de classe $\mathcal{C}^2$).
En particulier, $H_f(a)$ n'a que des valeurs propres réelles et on peut lui appliquer le théorème spectral (diagonalisation en base orthonormée).

Corollaire 8 (Formule de Taylor-Young à l'ordre 2 dans $\mathbb{R}^n$)
Soit Ω un ouvert de $\mathbb{R}^n$ et $f \in \mathcal{C}^2(\Omega,\mathbb{R})$. Pour tout $a \in \Omega$, on a :

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

ou encore

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

Preuve
Direct d'après l'expression du produit scalaire canonique de $\mathbb{R}^n$ comme un produit de matrices "ligne $\times$ colonne".

Théorème 9 (Conditions d'ordre 2)
Soit Ω un ouvert de $\mathbb{R}^n$, soit $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ et soit $a \in \Omega$.

(*i*) *Si f admet un minimum local en a, alors a est un point critique de f et $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$.*

(*ii*) *Si a est un point critique de f et si $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f admet un minimum local en a.*

De même pour les maximum locaux en remplaçant $\mathcal{S}_n^+(\mathbb{R})$ par $\mathcal{S}_n^-(\mathbb{R})$ (i.e. $-H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$) et $\mathcal{S}_n^{++}(\mathbb{R})$ par $\mathcal{S}_n^{--}(\mathbb{R})$ (i.e. $-H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$).

Preuve
Etant donné un point critique $a \in \Omega$, on a, d'après la formule de Taylor-Young à l'ordre 2 :

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

Appliquons le théorème spectral à la matrice réelle symétrique $H_f(a)$: il existe une base orthonormée de $\mathbb{R}^n$ $(e_1, \cdots, e_n)$ formée de vecteurs propres de $H_f(a)$, dont les valeurs propres associées seront notées $(\lambda_1, \cdots, \lambda_n) \in \mathbb{R}^n$.

(*i*) Si f admet un minimum local en a, alors on sait que a est critique et pour h assez petit, $f(a+h) - f(a) \geq 0$, donc

$$0 \leq \frac{1}{2}(H_f(a)h|h) + o(\|h\|^2).$$

En appliquant ceci en particulier au vecteur $h = te_i$ (avec $t \in \mathbb{R}$), on obtient

$$0 \leq \frac{1}{2}(\lambda_i te_i|te_i) + o(\|te_i\|^2),$$

soit

$$0 \leq \frac{\lambda_i}{2}t^2 + o(t^2) = t^2(\lambda_i + o(1)).$$

Donc pour $t \neq 0$ assez petit, on a $\lambda_i + o(1) \geq 0$. Par passage à la limite lorsque $t \to 0$, on obtient $\lambda_i \geq 0$, et ce pour tout $i \in [\![1, n]\!]$, donc $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$.

(*ii*) Si a est un point critique et si $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors $\lambda_i > 0$, donc pour tout vecteur $h = \sum_{i=1}^n h_i e_i$ de norme assez petite, on a

$$f(a+h)-f(a) = \frac{1}{2}\sum_{1 \leq i,j \leq n} h_i h_j (H_f(a)e_i|e_j) + o(\|h\|^2) = \frac{1}{2}\sum_{i=1}^n \lambda_i h_i^2 + o(\|h\|^2) \geq \frac{\lambda_{\min}}{2}\|h\|^2 + o(\|h\|^2),$$

où $\lambda_{\min} > 0$ désigne la plus petite valeur propre de $H_f(a)$. D'où, pour h voisin de 0_E :

$$f(a+h) - f(a) \geq \|h\|^2\left(\frac{\lambda_{\min}}{2} + o(1)\right) \geq \|h\|^2\frac{\lambda_{\min}}{4} \geq 0,$$

ce qui montre que f admet un minimum local en a.

Remarque
Pour $f : I \to \mathbb{R}$ de classe $\mathcal{C}^2$ (avec I un intervalle **ouvert** *de $\mathbb{R}$), on obtient les résultats suivants :*

- *si f admet un minimum local en a, alors $f'(a) = 0$ et $f''(a) \geq 0$.*
- *si $f'(a) = 0$ et si $f''(a) > 0$, alors f admet un minimum local en a.*

Corollaire 10 (Conditions d'ordre 2 avec le signe des valeurs propres)
Soit Ω un ouvert de $\mathbb{R}^n$, soit $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ et soit $a \in \Omega$ un point critique de f.

(*i*) *Si $Sp(H_f(a)) \subset \mathbb{R}_+^*$, alors f admet un minimum local en a.*

(*ii*) *Si $Sp(H_f(a)) \subset \mathbb{R}_-^*$, alors f admet un maximum local en a.*

(*iii*) *Si $H_f(a)$ possède deux valeurs propres non nulles de signes opposés, alors a est un point col (c'est-à-dire que f n'admet pas d'extremum local en a).*

(*iv*) *Sinon, on ne peut pas conclure.*

Preuve
Les cas (i) et (ii) sont directs d'après le théorème précédent. Pour le point (iii) : si $H_f(a)$ possède deux valeurs propres non nulles de signes opposés, alors $H_f(a)$ n'est ni dans $\mathcal{S}_n^+(\mathbb{R})$, ni dans $\mathcal{S}_n^-(\mathbb{R})$, donc f ne possède pas d'extremum local en a.

Corollaire 11 (Conditions d'ordre 2 en dimension 2)
Soit Ω un ouvert de $\mathbb{R}^2$, soit $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ et soit $a \in \Omega$ un point critique de f.

(i) *Si $\det(H_f(a)) > 0$, alors f admet un extremum local en a. Il s'agit d'un minimum lorsque $tr(H_f(a)) > 0$, d'un maximum sinon.*

(ii) *Si $\det(H_f(a)) < 0$, alors f admet un point col en a.*

(iii) *Si $\det(H_f(a)) = 0$, alors on ne peut pas conclure.*

Preuve
Si $n = 2$, alors en notant λ_1, λ_2 les deux valeurs propres de $H_f(a) \in \mathcal{S}_2(\mathbb{R})$, on a

$$\det(H_f(a)) = \lambda_1\lambda_2, \qquad tr(H_f(a)) = \lambda_1 + \lambda_2,$$

donc le résultat est direct en étudiant les signes de λ_1 et λ_2.

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

3) Optimisation sous contrainte

On revient ici au cas général où E est un $\mathbb{R}$-evn de dimension finie. On va maintenant chercher les extrema de f non pas sur un ouvert de E, mais sur une "hypersurface" de E (partie définie comme les zéros d'une fonction).

Théorème 12 (Optimisation sous contrainte)
Soit Ω un ouvert de E, et soient $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $g \in \mathcal{C}^1(\Omega, \mathbb{R})$. On considère "l'hypersurface" :

$$X = g^{-1}(\{0\}) = \{x \in \Omega,\ g(x) = 0\}.$$

Si la restriction $f|_X$ admet un extremum local en $a \in X$ et si $dg(a) \neq 0_{\mathcal{L}(E,\mathbb{R})}$, alors

$$\exists \lambda \in \mathbb{R}, \qquad df(a) = \lambda dg(a).$$

Remarque
La condition $df(a) = \lambda dg(a)$ se réécrit (en choisissant une base de E) :

$$\exists \lambda \in \mathbb{R},\ \forall i \in [\![1, n]\!], \quad \frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial g}{\partial x_i}(a).$$

Vocabulaire
Le scalaire λ est appelé **multiplicateur de Lagrange**.

Preuve
Supposons que $f|_X$ admet un minimum local en $a \in X$, c'est-à-dire qu'il existe $r > 0$ tel que $B(a, r) \subset \Omega$ et

$$\forall x \in B(a, r) \cap X, \qquad f(x) \geq f(a).$$

- Tout d'abord, on a $\boxed{T_aX \subset Ker(df(a))}$: en effet, si $v \in T_aX$, alors il existe un arc $\gamma :]-\varepsilon, \varepsilon[\to X$ tel que $\gamma(0) = a$, $\gamma'(0) = v$. Pour $|t|$ suffisamment petit, on a $\gamma(t) \in X \cap B(a, r)$ (par continuité de γ en 0), donc $f(\gamma(t)) \geq f(a) = f(\gamma(0))$, donc $f \circ \gamma$ (définie sur un voisinage de 0) atteint un minimum local en $t = 0$. On en déduit $(f \circ \gamma)'(0) = 0$, c'est-à-dire :

$$df(\gamma(0)) \cdot \gamma'(0) = 0,$$

ou encore $df(a) \cdot v = 0$, donc $v \in Ker(df(a))$.

- Puisque $dg(a) \neq 0_{\mathcal{L}(E,\mathbb{R})}$, on a d'après le théorème de l'hyperplan tangent : $T_aX = Ker(dg(a))$, donc $\boxed{Ker(dg(a)) \subset Ker(df(a))}$.
- On conclut avec des arguments d'algèbre linéaire : si $df(a)$ est nulle, alors elle est évidemment proportionnelle à $dg(a)$. Si $df(a)$ est non nulle, alors $Ker(df(a))$ est un hyperplan de E, tout comme $Ker(dg(a))$, donc l'inclusion précédemment établie est une égalité :

$$Ker(dg(a)) = Ker(df(a)).$$

Ainsi, les formes linéaires non nulles $df(a)$ et $dg(a)$ valent toutes deux 0 sur l'hyperplan $H = Ker(df(a)) = Ker(dg(a))$. En notant $e \notin H$, on a $E = H \oplus Vect(e)$, donc en posant $\lambda_1 = df(a) \cdot e$ et $\lambda_2 = dg(a) \cdot e$, on a $\lambda_1, \lambda_2 \in \mathbb{R}^*$, et

$$df(a) \cdot e = \frac{\lambda_1}{\lambda_2} dg(a) \cdot e,$$

donc $\boxed{df(a) = \frac{\lambda_1}{\lambda_2} dg(a)}$ (ces deux formes linéaires coïncident sur H et en e, donc par linéarité, elles sont égales).

ATTENTION !
La condition est nécessaire, mais pas suffisante.

Exemple
Déterminer $\min\limits_{x^2+y^2=1} xy$, *puis* $\min\limits_{x^2+y^2\leq 1} xy$.

Exemple
En considérant l'application $f : (x_1, \cdots, x_n) \mapsto x_1x_2 \cdots x_n$ *et la contrainte* $C_s = \{(x_1, \cdots, x_n) \in \mathbb{R}_+^n,\ x_1 + \cdots + x_n = s\}$, *retrouver l'inégalité arithmético-géométrique.*

III Exemples d'équations aux dérivées partielles (EDP)

Il s'agit juste ici de donner quelques exemples et quelques méthodes de résolution.

1) Ordre 1

Vocabulaire
Résoudre sur un ouvert Ω de $\mathbb{R}^n$ une **équation aux dérivées partielles** (abrégé en "EDP") d'ordre 1 en l'inconnue f, c'est déterminer toutes les fonctions $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ vérifiant une relation donnée entre f et ses dérivées partielles.

Propriété 13 (Fonctions à une dérivée nulle)
Soient I et J deux intervalles ouverts non vides de $\mathbb{R}$ et $\Omega = I \times J$ (c'est un ouvert de $\mathbb{R}^2$).

(i) *Les solutions sur Ω de l'équation $\dfrac{\partial f}{\partial x}(x,y) = 0$ sont les fonctions de la forme*

$$f : (x,y) \mapsto A(y) \text{ où } A \in \mathcal{C}^1(J, \mathbb{R}).$$

(ii) *De même, les solutions sur Ω de $\dfrac{\partial f}{\partial y}(x,y) = 0$ sont les fonctions de la forme*

$$f : (x,y) \mapsto B(x) \text{ où } B \in \mathcal{C}^1(I, \mathbb{R}).$$

Preuve
Analyse : Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ vérifiant $\dfrac{\partial f}{\partial x}(x,y) = 0$ sur Ω.
Fixons $y \in J$. Alors l'application partielle $x \mapsto f(x,y)$ a pour dérivée $\dfrac{\partial f}{\partial x}(x,y) = 0$ donc est une constante (qui dépend de y).
Il existe ainsi $A(y) \in \mathbb{R}$ tel que $\forall x \in I$, $f(x,y) = A(y)$.
Cela définit une fonction $A : J \to \mathbb{R}$ telle que pour tout $(x,y) \in I \times J$, $f(x,y) = A(y)$.
En particulier, si on fixe $a \in I$, on a pour tout $y \in J$, $A(y) = f(a,y)$. Et comme f est $\mathcal{C}^1$, l'application partielle $y \mapsto f(a,y)$ est également $\mathcal{C}^1$, donc $A \in \mathcal{C}^1(J, \mathbb{R})$.
Synthèse : on vérifie par le calcul que, réciproquement, une telle fonction f est solution de l'équation étudiée.

Exemple
Résoudre l'EDP : $\dfrac{\partial f}{\partial x}(x,y) = x$ sur $\mathbb{R}^2$.
Analyse : Soit $f \in \mathcal{C}^1(\mathbb{R}^2)$ telle que $\dfrac{\partial f}{\partial x}(x,y) = x$ sur $\mathbb{R}^2$.
Fixons $y \in \mathbb{R}$ et soit $g : \mathbb{R} \to \mathbb{R}$ l'application partielle définie par $g(x) = f(x,y)$. Alors : $g'(x) = x$ donc il existe une constante $C(y)$ telle que $f(x,y) = g(x) = \dfrac{x^2}{2} + C(y)$.
Cela permet de définir une fonction $C : \mathbb{R} \to \mathbb{R}$ est qui nécessairement $\mathcal{C}^1$ vu que g est $\mathcal{C}^1$.
Synthèse : Réciproquement, soit $C \in \mathcal{C}^1(\mathbb{R})$ et f définie sur $\mathbb{R}^2$ par $f(x,y) = \dfrac{x^2}{2} + C(y)$. Alors $f \in \mathcal{C}^1(\mathbb{R}^2)$ et on vérifie aisément qu'on a $\dfrac{\partial f}{\partial x}(x,y) = x$ sur $\mathbb{R}^2$.

Exemple
Résoudre sur $\mathbb{R}^2 \setminus \{(0,0)\}$ l'EDP :

$$(E) : x\frac{\partial f}{\partial y}(x,y) - y\frac{\partial f}{\partial x}(x,y) = 0,$$

en passant aux coordonnées polaires.

Soit $f \in \mathcal{C}^1(\mathbb{R}^2 \setminus \{(0,0)\})$.

Soit $\varphi : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{R}^2 \setminus \{(0,0)\} \\ (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta) \end{array} \right.$, qui est surjective et de classe $\mathcal{C}^1$.

On pose $g = f \circ \varphi : (r,\theta) \mapsto f(r\cos\theta, r\sin\theta)$ qui est de classe $\mathcal{C}^1$. Par composition, on peut calculer les dérivées partielles :

$$\frac{\partial g}{\partial \theta}(r,\theta) = -r\sin\theta \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta).$$

Par surjectivité de φ, on a f solution de (E) ssi $\forall (r,\theta) \in \mathbb{R}_+^* \times \mathbb{R},\ \dfrac{\partial g}{\partial \theta}(r,\theta) = 0$.

Or $\dfrac{\partial g}{\partial \theta} = 0 \iff \exists A \in \mathcal{C}^1(\mathbb{R}_+^*),\ \forall \theta \in \mathbb{R},\ g(r,\theta) = A(r) \iff \exists A \in \mathcal{C}^1(\mathbb{R}_+^*),\ \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, $f(x,y) = A(\sqrt{x^2+y^2})$.

2) Ordre 2

Vocabulaire

Résoudre sur un ouvert Ω de $\mathbb{R}^n$ une EDP d'ordre 2, c'est déterminer toutes les fonctions $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ vérifiant une relation donnée entre f et ses dérivées partielles premières et deuxièmes.

Propriété 14 (Fonctions à une dérivée seconde nulle)

Soient I et J deux intervalles ouverts non vides de $\mathbb{R}$ et $\Omega = I \times J$.

Les solutions sur Ω des équations $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 0$, $\dfrac{\partial^2 f}{\partial y^2}(x,y) = 0$ et $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = 0$ sont respectivement les fonctions $(x,y) \mapsto xA(y) + B(y)$, $(x,y) \mapsto yA(x) + B(x)$ et $(x,y) \mapsto A(x) + B(y)$, où A et B sont de classe $\mathcal{C}^2$.

Preuve

1. **Analyse :** Soit $f \in \mathcal{C}^2(\Omega)$ solution de $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 0$.

Soit $y \in J$ fixé. L'application partielle $g : I \to \mathbb{R}$ définie par $g(x) = f(x,y)$ vérifie : $g''(x) = 0$, donc est une fonction affine.

Il existe donc deux constantes $A(y), B(y) \in \mathbb{R}$ telles que, pour tout $x \in I$, $f(x,y) = g(x) = A(y)x + B(y)$. Cela permet de définir deux fonctions $A, B : J \to \mathbb{R}$ telles que, pour tout $(x,y) \in \Omega$, $f(x,y) = xA(y) + B(y)$.

Vérifions que A et B sont $\mathcal{C}^2$ sur Ω :

Soient $a < b \in I$. On peut écrire $A(y) = \dfrac{f(b,y) - f(a,y)}{b-a}$ et $B(y) = \dfrac{bf(a,y) - af(b,y)}{b-a}$, ce qui prouve, vu que f est $\mathcal{C}^2$, que A et B sont $\mathcal{C}^2$.

Synthèse : on vérifie par un calcul direct que, réciproquement, une telle fonction est solution de l'EDP.

2. **Analyse :** Soit $f \in \mathcal{C}^2(\Omega)$ solution de $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = 0 = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) = 0$.

Soit $y \in J$ fixé. L'application $g : x \mapsto \dfrac{\partial f}{\partial y}(x,y)$ vérifie $g'(x) = 0$ donc il existe $\tilde{A}(y) \in \mathbb{R}$ telle que, pour tout $x \in I$, $\dfrac{\partial f}{\partial y}(x,y) = \tilde{A}(y)$.

Cela définit une fonction $\tilde{A} : J \to \mathbb{R}$ qui est de classe $\mathcal{C}^1$. Elle admet en particulier une primitive que nous noterons A sur J et qui est de classe $\mathcal{C}^2$.

Soit $x \in I$ fixé. L'application partielle $h : J \to \mathbb{R}$ définie par $h(y) = f(x,y)$ vérifie $h'(y) = \tilde{A}(y)$ pour tout $y \in J$, donc il existe une constante $B(x) \in \mathbb{R}$ telle que pour tout $y \in J$, $f(x,y) = h(y) = A(y) + B(x)$. Cela définit une fonction $B : I \to \mathbb{R}$ qui est de classe $\mathcal{C}^2$ car f et A sont $\mathcal{C}^2$.

Synthèse : on vérifie que si f est définie par $f(x,y) = A(y) + B(x)$ où A et B sont $\mathcal{C}^2$, alors f est bien solution de l'EDP.

Exemple

L'équation des cordes vibrantes.

On considère une corde tendue horizontalement, de masse linéique homogène. On la suppose soumise à de petites oscillations, uniquement dans le plan (x,y).

On peut établir que la hauteur $y(x,t)$ à l'instant t du point de la corde d'abscisse x satisfait l'EDP : $\dfrac{\partial^2 y}{\partial x^2} = \dfrac{1}{c^2}\dfrac{\partial^2 y}{\partial t^2}$, où c est homogène à une vitesse.
On cherche les solutions y de classe $\mathcal{C}^2$.
On procède au changement de variable : $u = x + ct$, $v = x - ct$, i.e. $x = \dfrac{u+v}{2}$, $t = \dfrac{u-v}{2c}$. On pose $z\big(u(x,t), v(x,t)\big) = y(x,t)$, i.e. $z(u,v) = y\left(\dfrac{u+v}{2}, \dfrac{u-v}{2c}\right)$.
On a $\dfrac{\partial z}{\partial v} = \dfrac{1}{2}\dfrac{\partial y}{\partial x} - \dfrac{1}{2c}\dfrac{\partial y}{\partial t}$ puis $\dfrac{\partial^2 z}{\partial u \partial v} = \dfrac{1}{2}\left(\dfrac{1}{2}\dfrac{\partial^2 y}{\partial x^2} + \dfrac{1}{2c}\dfrac{\partial^2 y}{\partial t \partial x}\right) - \dfrac{1}{2c}\left(\dfrac{1}{2}\dfrac{\partial^2 y}{\partial x \partial t} + \dfrac{1}{2c}\dfrac{\partial^2 y}{\partial t^2}\right) = 0$.
D'après la proposition précédente, il existe deux fonctions A et B de classe C^2 sur $\mathbb{R}$ telles que $z(u,v) = A(u) + B(v)$, c'est-à-dire $y(x,t) = A(x+ct) + B(x-ct)$, que l'on peut interpréter comme la somme de deux ondes se propageant à la vitesse c et en sens inverse.