

Compléments d'Analyse: Calcul Différentiel

Université de Rennes 1
Isabelle Gruais

25 octobre 2023

1 Rappels de calcul différentiel

1.1 Différentielle

Définition 1.1.1 (Différentiabilité). Soit E et F deux evns et soit $U \subset E$ un ouvert de E . Une application $f : U \rightarrow F$ est dite différentiable en $x_0 \in U$ s'il existe une application linéaire continue $f'_{x_0} \in \mathcal{L}(E, F)$ appelée la différentielle de f en x_0 t.q.

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - f'_{x_0}(x - x_0)\|}{\|x - x_0\|} = 0$$

i.e. :

$$f(x) = f(x_0) + f'_{x_0}(x - x_0) + o(\|x - x_0\|).$$

cette dernière écriture signifiant que f admet un DL d'ordre 1 au voisinage de x_0 .

Remarque 1. 1. En dimension infinie, la différentiabilité d'une application dépend du choix de la norme utilisée, contrairement au cas de la dimension finie où toutes les normes sont équivalentes.

2. Par définition de la différentiabilité, la différentielle $f'_{x_0} : E \rightarrow F$ en un point x_0 est continue sur E . Si en outre l'application résultante $f' : U \rightarrow \mathcal{L}(E, F)$, $x \mapsto f'_x$, est continue sur U , alors f est dite de classe \mathcal{C}^1 sur U .

3. Si on a montré que f est différentiable en x_0 , le calcul pratique de $f'_{x_0}(h)$ pour $h \in E$ donné s'effectue grâce à la formule

$$f'_{x_0}(h) = \lim_{t \rightarrow 0} \frac{f(x_0 + th) - f(x_0)}{t}.$$

En toute rigueur, ce calcul donne la dérivée directionnelle de f en x_0 , ou différentielle au sens de Gâteaux de f en x_0 dans la direction h . Des définitions, il résulte immédiatement que si f est différentiable en x_0 , alors elle admet une dérivée directionnelle dans toutes les directions alors que la réciproque est fautive.

En résumé :

f de classe \mathcal{C}^1 en $x_0 \Rightarrow f$ est différentiable en $x_0 \Rightarrow f$ de classe \mathcal{C}^0 en x_0

\Downarrow

f dérivable en x_0 suivant toutes les directions

Exemple 1 (Contre-exemple). L'application f définie par

$$f(x, y) = \begin{cases} \frac{x^3}{x+y} & \text{si } x \neq -y, \\ 0 & \text{sinon} \end{cases}$$

admet une dérivée en $(0, 0)$ suivant toutes les directions de \mathbb{R}^2 mais n'est pas continue en $(0, 0)$

Soit $\lambda \in \mathbb{R}^*$. On a : $\forall (x, y) \in \mathbb{R}^2, x \neq -y,$

$$f(x, y) = \lambda \iff y = \frac{x^3}{\lambda} - x$$

donc

$$\lim_{(x,y) \rightarrow (0,0), y = \frac{x^3}{\lambda} - x} f(x, y) = \lambda \neq 0.$$

De plus, $\forall (x, y) \in \mathbb{R}^2, x \neq -y,$

$$\forall t \in \mathbb{R}^*, \quad f(tx, ty) = t^2 f(x, y) \Rightarrow \lim_{t \rightarrow 0} \frac{f(tx, ty)}{t} = 0 =: f'_{(0,0)}(x, y).$$

Définition 1.1.2 (Différentiabilité d'ordre supérieur). On dit que $f : U \subset E \rightarrow F$ est deux fois différentiable sur U si f est différentiable sur U de différentielle f' différentiable sur U et on note :

$$f'' := (f')'_x, \quad \forall x \in U$$

la différentielle de f' en $x \in U$, appelée différentielle d'ordre 2 en $x \in U$. Alors, on dit que f est de classe \mathcal{C}^2 sur U si f'' est continue sur U .

Théorème 1.1.1 (Schwartz). *Si $f : U \subset E \rightarrow F$ admet une différentielle d'ordre 2 sur U , alors pour tout $x \in U$, l'application $f''_x : E \times E \rightarrow F$ est une forme bilinéaire symétrique sur $E \times E$ à valeurs dans F .*

Démonstration. On peut définir les applications :

$$f' : U \subset E \rightarrow \mathcal{L}(E, F) \quad \text{et} \quad f'' : U \subset E \rightarrow \mathcal{L}(E, \mathcal{L}(E, F)).$$

Doit $x \in U$. Par définition : $f''_x \in \mathcal{L}(E, \mathcal{L}(E, F))$ et donc $h_1 \mapsto f''_x(h_1) \in \mathcal{L}(E, F)$ est définie, linéaire et continue sur E . De même, pour tout $h_1 \in E$, l'application $h_2 \mapsto f''_x(h_1)(h_2) \in F$ est définie, linéaire et continue sur E . Finalement, l'application $(h_1, h_2) \mapsto f''_x(h_1)(h_2) \in F$ est définie, bilinéaire et continue sur $E \times E$, et on note :

$$f''_x(h_1, h_2) := f''_x(h_1)(h_2), \quad \forall (h_1, h_2) \in E^2.$$

Il reste à montrer que f''_x est symétrique. Soit $(h_1, h_2) \in E^2$. On pose :

$$A_x(h_1, h_2) := f(x + h_1 + h_2) - f(x + h_1) - f(x + h_2) + f(x).$$

On a :

$$\begin{aligned} A_x(h_1, h_2) &= (f(x + h_1 + h_2) - f(x + h_1)) - (f(x + h_2) - f(x)) = \\ &= f'_{x+h_1}(h_2) - f'_x(h_2) + o(\|h_2\|) = f''_x(h_1, h_2) + o(\sqrt{\|h_1\|^2 + \|h_2\|^2}). \end{aligned}$$

De même :

$$\begin{aligned} A_x(h_1, h_2) &= (f(x + h_1 + h_2) - f(x + h_2)) - (f(x + h_1) - f(x)) = \\ &= f'_{x+h_2}(h_1) - f'_x(h_1) + o(\|h_1\|) = f''_x(h_2, h_1) + o(\sqrt{\|h_1\|^2 + \|h_2\|^2}). \end{aligned}$$

Il en résulte :

$$f''_x(h_1, h_2) = f''_x(h_2, h_1), \quad \forall h_1, h_2 \in E.$$

□

Définition 1.1.3. Soit $\mathcal{L}_2(E, F)$ l'ensemble des formes bilinéaires symétriques sur $E \times E$ à valeurs dans F . Par récurrence sur $n \geq 1$, on note $\mathcal{L}_n(E, F)$ l'ensemble des formes n -linéaires symétriques sur E^n à valeurs dans F et on définit et on note $f^{(n)} \in \mathcal{L}_n(E, F)$ la différentielle d'ordre n de f . On dit que $f : U \subset E \rightarrow F$ est de classe \mathcal{C}^n sur U si f est n fois différentiable sur U et si $f^{(n)}$ est continue sur U .

Remarque 2. Soit H un espace de Hilbert et soit $f : H \rightarrow \mathbb{R}$ différentiable en $x_0 \in H$. Alors, d'après le Théorème de Riesz, il existe $\nabla f(x_0) \in H$, appelé le gradient de f en x_0 , t.q. : $f'_{x_0}(h) = \langle \nabla f(x_0), h \rangle_H, \forall h \in H$, et le DL de f au voisinage de x_0 devient :

$$f(x_0 + h) = f(x_0) + f'_{x_0}(h) + o(\|h\|_H) = f(x_0) + \langle \nabla f(x_0), h \rangle_H + o(\|h\|_H).$$

Si en outre f est deux fois différentiable en $x_0 \in H$, alors il existe $\nabla^2 f(x_0) \in \mathcal{L}(H, H)$ t.q. :

$$f''_{x_0}(h_1, h_2) = \langle \nabla^2 f(x_0) h_1, h_2 \rangle_V, \quad \forall h_1, h_2 \in H.$$

En dimension finie, l'endomorphisme $\nabla^2 f(x_0) \in \mathcal{L}(H, H)$ s'identifie à la matrice hessienne de f en x_0 .

1.2 Application de la différentielle

Proposition 1.2.1. *On suppose que $f : U \subset E \rightarrow F$ est de classe $\mathcal{C}^{(n+1)}$ sur U . Soit $x \in U$ et soit $h \in E$ t.q. $[x, x+h] \subset U$. Alors, le développement de Taylor de f en x avec reste intégral est défini par :*

$$f(x+h) = f(x) + \sum_{k=1}^n \frac{1}{k!} \underbrace{f_x^{(k)}(h, \dots, h)}_{:= f_x^{(k)} \cdot h^k} + \int_0^1 \frac{(1-t)^n}{n!} f_{x+th}^{(n+1)} dt \cdot h^{n+1}$$

Démonstration. On pose :

$$\varphi(t) := f(x+th), \quad \forall t \in [0, 1].$$

Alors, $\varphi \in \mathcal{C}^{(n+1)}([0, 1], F)$ et on a :

$$\varphi(0) = f(x), \quad \varphi(1) = f(x+h)$$

et

$$\varphi^{(k)}(t) = f_{x+th}^{(k)} \cdot h^k, \quad \forall k \in [[0, n+1]].$$

On conclut en appliquant la formule de Taylor à φ en $t = 0$. □

Proposition 1.2.2. *On suppose que $f : U \subset E \rightarrow F$ est différentiable sur U . Si f admet un extremum en $a \in U$, alors $f'_a = 0$.*

Démonstration. Soit $r > 0$ t.q. $B(a, r) \subset U$ et soit $u \in E, \|u\| = 1$. On pose :

$$\varphi(t) := f(a+tu), \quad \forall t \in [-r, r].$$

Alors φ est dérivable sur $] -r, r[$, de dérivée définie par :

$$\varphi'(t) = f'_{a+tu}(u), \quad \forall t \in] -r, r[.$$

De plus, φ admet un extremum en $t = 0$, donc $\varphi'(0) = 0$, i.e.

$$f'_a(u) = 0.$$

Ceci étant vrai pour tout $u \in E$ t.q. $\|u\| = 1$, on en déduit, par linéarité de f'_a sur E , que :

$$f'_a(h) = 0, \quad \forall h \in E.$$

□

Définition 1.2.1. Si $f : U \subset E \rightarrow F$ est différentiable sur U , on appelle point critique de f tout $x \in U$ vérifiant $f'_x = 0$.

Remarque 3. Si $f : U \subset E \rightarrow F$ est différentiable sur U , l'ensemble de ses points critiques ne coïncide pas en général avec celui de ses extrema.

Proposition 1.2.3. On suppose que $f : U \subset E \rightarrow F$ admet une différentielle d'ordre 2 sur U . Si $a \in U$ est un point critique de f et si f''_a est positive sur E , i.e. vérifie :

$$f''_a \cdot h^2 \geq 0, \quad \forall h \in E$$

alors a est un minimum (local) de f .

Démonstration. Soit $r > 0$ t.q. $B(a, r) \subset U$ et soit $h \in B(0, r)$. On a :

$$f(a + h) = f(a) + \underbrace{f''_a \cdot h^2}_{\geq 0} + o(\|h\|^2)$$

et donc pour $\varepsilon \in]0, r[$ suffisamment petit :

$$f(a + h) \geq f(a), \quad \forall h \in B(0, \varepsilon).$$

□

Remarque 4 (Différentiabilité d'ordre supérieur). Si V est un espace de Hilbert et si $f : V \rightarrow \mathbb{R}$ est supposée différentiable en $x_0 \in V$, d'après le Théorème de Riesz, il existe $\nabla f(x_0) \in V$, appelé le gradient de f en x_0 , t.q. : $df_{x_0}(h) = \langle \nabla f(x_0), h \rangle_V, \forall h \in V$, conformément au DL : $\forall h \in V$,

$$f(x_0 + h) = f(x_0) + df_{x_0}(h) + o(\|h\|_V) = f(x_0) + \langle \nabla f(x_0), h \rangle_V + o(\|h\|_V).$$

On dit que f est deux fois différentiable en $x_0 \in V$ si $df : U \subset V \rightarrow \mathcal{L}(V, \mathbb{R}) = V' \sim V$ est différentiable en x_0 , i.e. s'il existe une application linéaire continue $L_{x_0} \in \mathcal{L}(V, V)$ t.q. :

$$df_{x_0+\xi} = df_{x_0} + L_{x_0}(\xi) + o(\|\xi\|) = df_{x_0} + d^2f_{x_0}(\xi) + o(\|\xi\|)$$

en posant $d^2f_{x_0} := L_{x_0}$, i.e. :

$$\nabla f(x_0 + \xi) = \nabla f(x_0) + \nabla^2 f(x_0)\xi + o(\|\xi\|_V)$$

en utilisant la notation d'endomorphisme $\nabla^2 f(x_0)\xi := d^2f_{x_0}(\xi)$ avec $\nabla^2 f(x_0) \in \mathcal{L}(V, V)$. On en déduit le DL à l'ordre 2 en x_0 :

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle_V + \frac{1}{2} \langle \nabla^2 f(x_0)h, h \rangle_V + o(\|h\|_V^2)$$

où $(h, \xi) \mapsto \langle \nabla^2 f(x_0)h, \xi \rangle_V$ est une forme bilinéaire continue sur $V \times V$.

En dimension finie, l'endomorphisme $\nabla^2 f(x_0) \in \mathcal{L}(V, V)$ s'identifie à la matrice hessienne de f en x_0 .

1.3 Cas particulier de la dimension finie

Dans la suite, on note (e_1, \dots, e_n) la base canonique de \mathbb{R}^n et on munit \mathbb{R}^n de sa structure euclidienne usuelle.

Proposition 1.3.1. *Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Si f est différentiable sur U , alors sa différentielle est donnée par :*

$$f'_x(h) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) =: \langle \nabla f(x), h \rangle$$

où $\nabla f(x) \in \mathbb{R}^n$ est appelé le gradient de f en x et est défini par ses composantes :

$$\nabla f(x)_i = \frac{\partial f}{\partial x_i}, \quad i = 1, \dots, n.$$

Démonstration. Soit $x \in U$ et soit $r > 0$ t.q. $B(x, r) \subset U$. Soit $h \in B(0, r)$,

$$f(x + h) = f(x) + f'_x(h) + o(\|h\|)$$

avec

$$f'_x(h) = \sum_{i=1}^n h_i f'_x(e_i).$$

Soit $i \in [[1, n]]$. On a : $\forall h \in]-r, r[, h \neq 0$.

$$f(x + he_i) = f(x) + hf'_x(e_i) + o(|h|)$$

et donc

$$f'_x(e_i) = \frac{f(x + he_i) - f(x)}{h} + o(1) := \lim_{h \rightarrow 0} \frac{f(x + he_i) - f(x)}{h} = \frac{\partial f}{\partial x_i}(x).$$

□

Proposition 1.3.2. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Si f est deux fois différentiable sur U , alors sa différentielle d'ordre 2 est donnée par :

$$f''_x(u, v) = \sum_{i=1}^n u_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x) =: \langle \nabla^2 f(x) u, v \rangle$$

où la matrice symétrique $\nabla^2 f(x) \in \mathbb{R}^{n \times n}$ est appelé la matrice hessienne de f en x et est définie par ses composantes :

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n.$$

Démonstration. Soit $x \in U$ et soit $r > 0$ t.q. $B(x, r) \subset U$. Soit $h \in B(0, r)$,

$$f(x + h) = f(x) + f'_x(h) + f''_x \cdot h^2 + o(\|h\|^2)$$

avec

$$f''_x \cdot h^2 = \sum_{i,j=1}^n h_i h_j f''_x(e_i, e_j).$$

Soit $i, j \in [[1, n]]$. En particulier :

$$f'_{x+h}(e_i) = f'_x(e_i) + f''_x(e_i, h) + o(\|h\|)$$

et donc : $\forall h \in]-r, r[, h \neq 0$,

$$f'_{x+he_j}(e_i) = f'_x(e_i) + hf''_x(e_i, e_j) + o(|h|).$$

Il en résulte :

$$\begin{aligned} f''_x(e_i, e_j) &= \frac{f'_{x+he_j}(e_i) - f'_x(e_i)}{h} + o(1) = \lim_{h \rightarrow 0} \frac{f'_{x+he_j}(e_i) - f'_x(e_i)}{h} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial x_i}(x + he_j) - \frac{\partial f}{\partial x_i}(x) \right) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \end{aligned}$$

□

Définition 1.3.1 (Fonction de classe \mathcal{C}^k). Soit $i \in [[1, n]]$ et soit $k \geq 2$. Soit U un ouvert de \mathbb{R}^n et soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

1. On dit que f admet une dérivée partielle d'indice i en x_0 si f admet une dérivée directionnelle en x_0 dans la direction de e_i .
2. On dit que f est de classe \mathcal{C}^k sur U si toutes ses dérivées partielles jusqu'à l'ordre k existent et sont continues sur U .

Remarque 5 (Contre-exemple de Peano). Si f n'est pas deux fois différentiable en x_0 alors la matrice hessienne n'est pas nécessairement symétrique, comme le montre le contre-exemple dû à Peano :

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

On en déduit :

$$\begin{aligned} \frac{|f(x, y)|}{\sqrt{x^2 + y^2}} &\leq \frac{|x||y|(x^2 + y^2)}{(x^2 + y^2)^{3/2}} = \frac{|x||y|}{(x^2 + y^2)^{1/2}} \leq \frac{(x^2 + y^2)}{2(x^2 + y^2)^{1/2}} = \\ &= \frac{1}{2}(x^2 + y^2)^{1/2} \xrightarrow{(x, y) \rightarrow (0, 0)} 0. \end{aligned}$$

et donc f est différentiable en $(0, 0)$ de différentielle $df_{(0,0)} = 0$. Des relations : $\forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$,

$$f(x, y) = xy \left(1 - \frac{2y^2}{x^2 + y^2} \right) = xy \left(\frac{2x^2}{x^2 + y^2} - 1 \right)$$

on déduit : $\forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$,

$$\frac{\partial f}{\partial x}(x, y) = y \left(1 - \frac{2y^2}{x^2 + y^2} \right) + \frac{4x^2 y^3}{(x^2 + y^2)^2},$$

$$\frac{\partial f}{\partial y}(x, y) = x \left(\frac{2x^2}{x^2 + y^2} - 1 \right) + \frac{4x^3 y^2}{(x^2 + y^2)^2}.$$

Il en résulte :

$$\lim_{y \rightarrow 0} \frac{1}{y} \left(\frac{\partial f}{\partial x}(0, y) - \frac{\partial f}{\partial x}(0, 0) \right) = \lim_{y \rightarrow 0} \frac{1}{y} \frac{\partial f}{\partial x}(0, y) = \lim_{y \rightarrow 0} \frac{-y}{y} = -1.$$

i.e. :

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1.$$

De même :

$$\lim_{x \rightarrow 0} \frac{1}{x} \left(\frac{\partial f}{\partial y}(x, 0) - \frac{\partial f}{\partial y}(0, 0) \right) = \lim_{x \rightarrow 0} \frac{1}{x} \frac{\partial f}{\partial y}(x, 0) = \lim_{x \rightarrow 0} \frac{x}{x} = 1.$$

i.e. :

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0),$$

donc la matrice hessienne de f en $(0, 0)$ n'est pas symétrique.

On rappelle les formules de Taylor à l'ordre 2.

Proposition 1.3.3 (Formule de Taylor avec reste intégral). *Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 sur un ouvert U de \mathbb{R}^n et soit $h \in \mathbb{R}^n$ t.q. $[x_0, x_0 + h] \subset U$. Alors :*

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \int_0^1 (1-t) \langle \nabla^2 f(x_0 + th)h, h \rangle dt$$

Proposition 1.3.4 (Formule de Taylor avec reste de Lagrange). *Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 sur un ouvert U de \mathbb{R}^n et soit $h \in \mathbb{R}^n$ t.q. $[x_0, x_0 + h] \subset U$. On suppose que :*

$$M := \sup_{t \in [0,1]} \frac{|\langle \nabla^2 f(x_0 + th)h, h \rangle|}{\|h\|^2} < +\infty.$$

Alors :

$$|f(x_0 + h) - f(x_0) - \langle \nabla f(x_0), h \rangle| \leq \frac{M}{2} \|h\|^2.$$

2 Existence et unicité de la solution d'un problème d'optimisation

On peut retenir comme principe général que la compacité fournit des résultats d'existence et la convexité un cadre favorable pour l'unicité.

Dans la suite, on étudie les fonctions convexes d'abord dans le cas particulier de la dimension finie. Dans la Section 2.3, on considère le cas plus général des espaces de Hilbert.

2.1 Existence en dimension finie

Soit $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ continue. On considère le problème :

$$\min_{x \in K} f(x). \quad (1)$$

Sans hypothèse supplémentaire, ce problème n'a pas de solution. Pour le voir, il suffit de prendre $f : x \mapsto e^x$ et $K = \mathbb{R}$.

Théorème 2.1.1 (Existence en dimension finie). *On suppose qu'il existe $x_0 \in \mathbb{R}^n$ t.q. $\{x \in \mathbb{R}^n, f(x) \leq f(x_0)\}$ soit borné. Alors le problème (1) admet une solution globale.*

Démonstration. On est ramené au problème :

$$\min_{x \in \tilde{K}} f(x) \quad \text{où} \quad \tilde{K} := \{x \in \mathbb{R}^n, f(x) \leq f(x_0)\}.$$

Par hypothèse, \tilde{K} est borné. De plus, $\tilde{K} = f^{-1}(]-\infty, f(x_0)])$ est fermé comme image réciproque par f continue du fermé $]-\infty, f(x_0)]$ de \mathbb{R} . On en déduit que \tilde{K} est compact et que f continue sur \tilde{K} est bornée sur \tilde{K} et y atteint ses bornes. En effet, par hypothèse, $m := \inf_{\tilde{K}} f > -\infty$ donc il existe une suite minimisante $(x_n)_{n \geq 0} \in \tilde{K}^{\mathbb{N}}$. Comme \tilde{K} est compact, on peut en extraire une suite convergente $(x_{\varphi(n)})_{n \geq 0}$. Soit x^* sa limite. Par continuité de f :

$$\lim_{n \rightarrow +\infty} f(x_{\varphi(n)}) = f(x^*).$$

Par construction :

$$\lim_{n \rightarrow +\infty} f(x_{\varphi(n)}) = m$$

donc $f(x^*) = \inf_{x \in \tilde{K}} f(x) = \min_{x \in \tilde{K}} f(x)$. □

Remarque 6 (Deux remarques très utiles en pratique...). Il reste à voir comment le Théorème 2.1.1 s'applique au problème 1. On commence par rappeler que l'hypothèse de dimension finie est fondamentale. Quand elle n'est pas vérifiée, il est facile de construire des contre-exemples.

1. Si K est compact, la continuité de f permet de conclure directement.
2. Si K est fermé et si f est coercive, i.e. si f vérifie :

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$$

alors le Théorème 2.1.1 s'applique.

Remarque 7 (Semi-continuité inférieure). Le Théorème 2.1.1 s'applique encore si on suppose seulement que f est semi-continue inférieurement sur K , i.e. si : $\forall x_0 \in K, \forall \alpha < f(x_0), f^{-1}(] \alpha, +\infty[) \in \mathcal{V}(x_0)$.

On montre que f est sci sur K ssi :

$$\forall \alpha \in \mathbb{R}, \quad f^{-1}(] - \infty, \alpha]) \quad \text{est fermé dans } \mathbb{R}^n,$$

ou, de façon équivalente, ssi : $\forall x_0 \in K, \forall \varepsilon > 0$, il existe un voisinage V de x_0 t.q. :

$$\forall x \in V, \quad f(x_0) - \varepsilon \leq f(x),$$

en abrégé : $f(x_0) \leq \liminf_{x \rightarrow x_0} f(x)$.

Démonstration. Il suffit de remarquer que si $x_n \xrightarrow[n \rightarrow +\infty]{} x^*$ alors $f(x^*) \leq \liminf_{n \rightarrow +\infty} f(x_n)$. \square

Exemple 2. Soit $I \subset \mathbb{R}$ un sous-ensemble quelconque et soit $(f_j)_{j \in I}$ une famille de fonctions linéaires de \mathbb{R}^n dans \mathbb{R} . On pose :

$$f(x) = \sup_{j \in I} f_j(x), \quad \forall x \in \mathbb{R}^n.$$

Alors f est semi-continue inférieurement. En effet : soit $\alpha \in \mathbb{R}$. Par définition de f : $f^{-1}(] \alpha, +\infty[) = \cup_{j \in I} f_j^{-1}(] \alpha, +\infty[)$ est ouvert comme réunion d'ouverts.

Exemple 3. Soit le problème : $\min_{(x,y) \in K} f(x, y)$ avec

$$f(x, y) = x^4 + y^4 - x^2, \quad K = \{(x, y) \in \mathbb{R}^2, x + y \leq 4\}.$$

On a :

$$\lim_{\|(x,y)\|_2 \rightarrow +\infty} f(x, y) = \lim_{\|(x,y)\|_2 \rightarrow +\infty} (x^4 + y^4) = +\infty$$

i.e. f est coercive. Comme de plus, f est continue et K est fermé, on déduit du Théorème 2.1.1 que le problème admet une solution.

Exemple 4. On considère une subdivision régulière de $[0, 1]$, soit :

$$u_0 = 0 < u_1 < \dots < u_{N+1} = 1,$$

définie par $u_i = ih$, $0 \leq i \leq N + 1$, $h = \frac{1}{N+1}$, où $N \geq 1$ est donné. Soit $(u_i, x_i)_{1 \leq i \leq N}$ un nuage de points de \mathbb{R}^2 . On pose $x_0 = 0$, $x_{N+1} = 1$. On pose $x = (x_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ et on note $f(x)$ la longueur de la courbe affine

par morceaux passant par les points (u_i, x_i) , $i \in [[0, N + 1]]$. On montre aisément que

$$f(x) = \sum_{i=0}^N \sqrt{(u_{i+1} - u_i)^2 + (x_{i+1} - x_i)^2} = h \sum_{i=0}^N \sqrt{1 + \frac{(x_{i+1} - x_i)^2}{h^2}}.$$

On s'intéresse au problème

$$\inf_{x \in \mathbb{R}^N} f(x). \quad (2)$$

On remarque que : $\forall k \in [[1, N + 1]]$,

$$f(x) \geq h \sum_{i=0}^{k-1} \frac{|x_{i+1} - x_i|}{h} = \sum_{i=0}^{k-1} |x_{i+1} - x_i| \geq \left| \sum_{i=0}^{k-1} (x_{i+1} - x_i) \right| = |x_k|,$$

i.e. : $f(x) \geq \|x\|_\infty$. Donc f est coercive. Comme de plus \mathbb{R}^N est fermé, on déduit du Théorème 2.1.1 que le problème (2) admet une solution.

2.2 Unicité de l'optimum

La convexité fournit une condition suffisante d'unicité dans les problèmes d'optimum.

Définition 2.2.1 (Ensembles convexes et fonctions convexes). Un ensemble $K \subset \mathbb{R}^n$ est dit convexe s'il vérifie

$$\forall (x, y) \in K^2, \quad \forall t \in [0, 1], \quad tx + (1 - t)y \in K$$

Une fonction $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ définie sur un ensemble convexe $K \subset \mathbb{R}^n$ est dite convexe si :

$$\forall (x, y) \in K^2, \quad \forall t \in [0, 1], \quad f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Une fonction convexe $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est dite strictement convexe si

$$\forall (x, y) \in K^2, \quad x \neq y \Rightarrow \forall t \in]0, 1[, \quad f(tx + (1 - t)y) < tf(x) + (1 - t)f(y).$$

Théorème 2.2.1 (Caractérisation des fonctions convexes régulières).

1. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable, les propositions suivantes sont équivalentes.

(i) f est convexe sur \mathbb{R}^n .

(ii) $\forall x, y \in \mathbb{R}^n, f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.

(iii) $\forall x, y \in \mathbb{R}^n, \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$.

2. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable, les propositions suivantes sont équivalentes.

(i) f est strictement convexe sur \mathbb{R}^n .

(ii) $\forall x, y \in \mathbb{R}^n, x \neq y \Rightarrow f(y) > f(x) + \langle \nabla f(x), y - x \rangle$.

(iii) $\forall x, y \in \mathbb{R}^n, x \neq y \Rightarrow \langle \nabla f(y) - \nabla f(x), y - x \rangle > 0$.

3. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est deux fois différentiable, les propositions suivantes sont équivalentes.

(i) f est convexe sur \mathbb{R}^n .

(ii) $\forall x \in \mathbb{R}^n$, la matrice hessienne $\nabla^2 f(x)$ est semi-définie positive.

Exemple 5 (Convexité d'une fonction quadratique). Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique réelle d'ordre n , $b \in \mathbb{R}^n$ un vecteur et soit $c \in \mathbb{R}$. On considère la fonction :

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle + c$$

En utilisant la symétrie de A , on obtient directement :

$$\forall x, h \in \mathbb{R}^n, \quad f(x+h) = f(x) + \langle Ax - b, h \rangle + \frac{1}{2} \langle Ah, h \rangle$$

ce qui conduit à :

$$\frac{1}{\|h\|} |f(x+h) - f(x) - \langle Ax - b, h \rangle| = \frac{\langle Ah, h \rangle}{2\|h\|} \leq \frac{\|A\|}{2} \|h\| \xrightarrow{h \rightarrow 0} 0$$

i.e. :

$$\forall x, h \in \mathbb{R}^n, \quad df_x(h) = \langle Ax - b, h \rangle \Rightarrow \nabla f(x) = Ax - b.$$

Il en résulte que ∇f est différentiable sur \mathbb{R}^n , de différentielle définie par : $\nabla^2 f(x) = A, \forall x \in \mathbb{R}^n$.

En particulier, on déduit de ce calcul et du Théorème 2.2.1 que f est convexe, resp. strictement convexe, ssi A est semi-définie positive, resp. strictement définie positive.

Théorème 2.2.2. On considère le problème (1) avec f convexe et K convexe, éventuellement de dimension infinie. Alors

1. Tout minimum local est global.

2. Si f est strictement convexe alors il y a au plus un minimum.

Démonstration. 1. On suppose que le problème (1) admet une solution x^* et que cette solution est un minimum local. Il existe $r > 0$ t.q.

$$\forall x \in K, \quad \|x - x^*\| < r \Rightarrow f(x^*) \leq f(x).$$

Soit $x \in K$, $\|x - x^*\| \geq r$. On pose

$$y = x^* + \frac{r}{2} \frac{x - x^*}{\|x - x^*\|}.$$

Par construction :

$$\|y - x^*\| = \frac{r}{2} < r.$$

De plus :

$$y = \left(1 - \frac{r}{2\|x - x^*\|}\right) x^* + \frac{r}{2\|x - x^*\|} x$$

avec

$$0 < \frac{r}{2\|x - x^*\|} \leq \frac{1}{2} < 1$$

donc $y \in K$ par convexité de K . On en déduit, par hypothèse sur x^* : $f(x^*) \leq f(y)$. De plus, par convexité de f sur K :

$$f(y) \leq \left(1 - \frac{r}{2\|x - x^*\|}\right) f(x^*) + \frac{r}{2\|x - x^*\|} f(x).$$

On en déduit :

$$\begin{aligned} f(x^*) &\leq \left(1 - \frac{r}{2\|x - x^*\|}\right) f(x^*) + \frac{r}{2\|x - x^*\|} f(x) \iff \\ \iff \frac{r}{2\|x - x^*\|} f(x^*) &\leq \frac{r}{2\|x - x^*\|} f(x) \iff_{r>0} f(x^*) \leq f(x). \end{aligned}$$

Donc x^* est un minimum global sur K .

2. On suppose que f est strictement convexe sur K et qu'il y a deux minima $x_1^* \neq x_2^*$. De 1., on déduit que $f(x_1^*) = f(x_2^*) = \min_K f$. Soit $t \in]0, 1[$. Par hypothèse : $x_1^* \neq x_2^* \Rightarrow$

$$f(tx_1^* + (1-t)x_2^*) < tf(x_1^*) + (1-t)f(x_2^*) = \min_K f$$

avec $tx_1^* + (1-t)x_2^* \in K$ par convexité de K . Contradiction. □

2.3 Existence en dimension infinie

Dans ce paragraphe, on énonce un résultat d'existence en dimension infinie dans le cas où f vérifie une hypothèse de convexité forte. Dans le cas général, il est beaucoup plus difficile d'établir un résultat d'existence en

dimension infinie. A titre d'exemple, on peut considérer l'espace de Hilbert de dimension infinie défini par :

$$\ell^2(\mathbb{N}) = \{u = (u_n)_{n \geq 0}, \sum_{n \geq 0} |u_n|^2 < +\infty\}.$$

muni du produit scalaire : $(x, y) \mapsto \langle x, y \rangle = \sum_{n \geq 0} x_n y_n$. On considère la fonctionnelle f définie par :

$$\forall x \in \ell^2(\mathbb{N}), \quad f(x) = (\|x\|^2 - 1)^2 + \sum_{n \geq 0} \frac{x_n^2}{n+1}$$

et on s'intéresse au problème de minimisation :

$$\inf_{x \in \ell^2(\mathbb{N})} f(x).$$

On remarque immédiatement que f est coercive. En effet :

$$\forall x \in \ell^2(\mathbb{N}), \quad f(x) \geq (\|x\|^2 - 1)^2 \xrightarrow{\|x\| \rightarrow +\infty} +\infty.$$

Pourtant f n'admet pas de minimum sur $\ell^2(\mathbb{N})$. En effet : $f \geq 0$ et

$$f(x) = 0 \iff \|x\| = 1 \quad \text{et} \quad \sum_{n \geq 0} \frac{x_n^2}{n+1} = 0 \Rightarrow x_n = 0, \quad \forall n \geq 0$$

ce qui contredit $\|x\| = 1$, donc $f > 0$ sur $\ell^2(\mathbb{N})$.

Soit $x^{(n)} \in \ell^2(\mathbb{N})$ la série définie par :

$$x_k^{(n)} = \delta_{k,n} = \begin{cases} 1 & \text{si } k = n, \\ 0 & \text{sinon} \end{cases}$$

On a :

$$\forall n \geq 0, \quad f(x^{(n)}) = \frac{1}{n+1} \xrightarrow{n \rightarrow +\infty} 0$$

i.e. $(x^{(n)})_{n \geq 0}$ est une suite minimisante de $\ell^2(\mathbb{N})$.

On remarque que la suite $(x^{(n)})_{n \geq 0}$ est bornée mais non compacte dans $\ell^2(\mathbb{N})$.

Dans ce qui suit, on étudie un cas d'existence en dimension infinie.

Définition 2.3.1 (Fonction α -elliptique). Soit V un espace de Hilbert, $\langle \cdot, \cdot \rangle$ un produit scalaire sur V . Soit $K \subset V$ un convexe. Une fonction $f : K \rightarrow \mathbb{R}$ est dite fortement convexe ou uniformément convexe ou α -convexe ou α -elliptique s'il existe $\alpha > 0$ t.q. : $\forall (x, y) \in K^2, \forall t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\alpha}{2}t(1-t)\|x - y\|^2.$$

On vérifie immédiatement que l'uniforme convexité entraîne la stricte convexité et donc la convexité. La convexité correspond au cas $\alpha = 0$.

Exemple 6 (Liens entre les variantes de la convexité). On donne des exemples élémentaires qui seront repris en détails dans la suite. En particulier, on étudiera la convexité des formes quadratiques en dimension finie.

1. Les fonctions affines de \mathbb{R} dans \mathbb{R} sont convexes mais non strictement convexes.
2. Une fonction α -elliptique est strictement convexe donc convexe.
3. La fonction $f : x \mapsto -\ln x$ est strictement convexe sur $]0, +\infty[$ mais non elliptique. En effet, f est de classe \mathcal{C}^∞ sur $]0, +\infty[$ et on a $f''(x) = \frac{1}{x^2} > 0, \forall x > 0$. On raisonne par l'absurde en supposant qu'il existe $\alpha > 0$ t.q. $\forall t \in [0, 1], \forall x, y \in]0, +\infty[$,

$$-\ln(tx + (1-t)y) \leq -t\ln(x) - (1-t)\ln(y) - \frac{\alpha}{2}t(1-t)|x - y|^2.$$

Soit $t \in]0, 1[$ et soit $y > 0$. Alors : $\forall x > y$,

$$-\frac{\ln(tx + (1-t)y)}{|x - y|^2} \leq \frac{-t\ln(x) - (1-t)\ln(y)}{|x - y|^2} - \frac{\alpha}{2}t(1-t).$$

avec

$$\frac{\ln(tx + (1-t)y)}{|x - y|^2} = \frac{\ln(x) + \ln(t) + \frac{(1-t)y}{t} + o\left(\frac{y}{x}\right)}{|x|^2 \left(1 - \frac{2y}{x} + o\left(\frac{y}{x}\right)\right)} \underset{x \rightarrow +\infty}{\sim} \frac{\ln(x)}{x^2}$$

donc

$$\lim_{x \rightarrow +\infty} -\frac{\ln(tx + (1-t)y)}{|x - y|^2} = 0.$$

De même :

$$\frac{-t\ln(x) - (1-t)\ln(y)}{|x - y|^2} = -\frac{t\ln(x) \left(1 + \frac{(1-t)\ln(y)}{t\ln(x)} + o\left(\frac{1}{\ln(x)}\right)\right)}{|x|^2 \left(1 - \frac{2y}{x} + o\left(\frac{y}{x}\right)\right)}$$

$$\underset{x \rightarrow +\infty}{\sim} -\frac{t \ln(x)}{x^2}$$

donc

$$\lim_{x \rightarrow +\infty} \frac{-t \ln(x) - (1-t) \ln(y)}{|x-y|^2} = 0.$$

On en déduit, quand $x \rightarrow +\infty$:

$$0 \leq -\frac{\alpha}{2}t(1-t)$$

en contradiction avec $-\frac{\alpha}{2}t(1-t) < 0$.

4. On vérifie aisément que $x \mapsto x^2$ est ellitique de rapport $\alpha = 2$. En effet :

$$(tx + (1-t)y)^2 = tx^2 + (1-t)^2y^2 - t(1-t)(x-y)^2.$$

Dans le cas où f est régulière, il existe des caractérisations de l'uniforme convexité qui peuvent être vues comme des conséquences du Théorème 2.2.1.

Corollaire 2.3.1 (Caractérisation des fonctions uniformément convexes).

1. Si $f : V \rightarrow \mathbb{R}$ est différentiable, alors les propositions suivantes sont équivalentes.
 - (i) f est α -elliptique.
 - (ii) $\forall (x, y) \in V^2$, $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2} \|x - y\|^2$.
 - (iii) $\forall (x, y) \in V^2$, $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \frac{\alpha}{2} \|x - y\|^2$.
2. Si $f : V \rightarrow \mathbb{R}$ est deux fois différentiable, alors les propositions suivantes sont équivalentes.
 - (i) f est α -elliptique.
 - (ii) $\forall (x, h) \in V^2$, $\langle \nabla^2 f(x)h, h \rangle \geq \alpha \|h\|^2$.

Exemple 7 (Uniforme convexité d'une fonction quadratique). Soit $A \in \mathcal{M}_n(\mathbb{R})$ symétrique d'ordre n , soit $b \in \mathbb{R}^n$ et soit $c \in \mathbb{R}$. Soit f la fonction quadratique définie par :

$$\forall x \in \mathbb{R}^n, \quad f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle + c.$$

On a vu dans l'Exemple 5 que f est convexe, resp. strictement convexe, ssi A est semi-définie positive, resp. définie positive et que de plus $\nabla^2 f(x) = A$, $\forall x \in \mathbb{R}^n$. La matrice A étant symétrique, elle est diagonalisable dans une base orthonormée de \mathbb{R}^n , i.e. A se décompose sous la forme : $A = U^T D U$ où $U \in \mathcal{O}_n(\mathbb{R})$ est orthogonale et où D est diagonale, formée par les valeurs

propres $\lambda_1^2 \leq \dots \leq \lambda_n^2$ de A , avec la convention $\lambda_i \geq 0, \forall i \in [[1, n]]$. On en déduit : $\forall h \in \mathbb{R}^n$,

$$\begin{aligned} \langle Ah, h \rangle &= \langle U^T DUh, h \rangle = \langle DUh, Uh \rangle = \sum_{i=1}^n \lambda_i^2 (Uh)_i^2 \geq \lambda_1^2 \|Uh\|^2 \\ &\stackrel{U \in \mathcal{O}_n(\mathbb{R})}{=} \lambda_1^2 \|h\|^2 \end{aligned}$$

d'où on déduit que f est λ_1^2 -elliptique dès que $\lambda_1 > 0$, i.e. dès que A est définie positive. Si $h \in \mathbb{R}^n \setminus \{0\}$ vérifie $Ah = \lambda_1^2 h$, alors $\langle Ah, h \rangle = \lambda_1^2 \|h\|^2$ et donc

$$\lambda_1^2 = \min_{h \neq 0} \frac{\langle Ah, h \rangle}{\|h\|^2}$$

i.e., λ_1^2 est la meilleure constante d'ellipticité de A .

Remarque 8. Si $f : V \rightarrow \mathbb{R}$ est α -elliptique et différentiable, alors f est coercive. En effet : soit $(x, y) \in V^2$. D'après le Corollaire 2.3.1 :

$$\begin{aligned} f(y) &\geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2} \|y - x\|^2 = \\ &= f(x) + \|x - y\| \left(\left\langle \nabla f(x), \frac{y - x}{\|x - y\|} \right\rangle + \frac{\alpha}{2} \|y - x\| \right) \end{aligned}$$

avec, pour $x \in \mathbb{R}^n$ fixé et $\|y\| \rightarrow +\infty$:

$$\begin{aligned} \|y - x\| &= \|y\| \left(1 - \frac{\langle x, y \rangle}{\|y\|^2} + o\left(\frac{1}{\|y\|}\right) \right) \underset{\|y\| \rightarrow +\infty}{\sim} \|y\| \\ \left| \left\langle \nabla f(x), \frac{y - x}{\|y - x\|} \right\rangle \right| &\leq \|\nabla f(x)\| \end{aligned}$$

donc

$$f(y) \geq f(x) + \frac{\alpha}{2} \|y\|^2 (1 + o(1)) \underset{\|y\| \rightarrow +\infty}{\rightarrow} +\infty$$

i.e. : $\lim_{\|y\| \rightarrow +\infty} f(y) = +\infty$.

On est maintenant en mesure d'énoncer le résultat d'existence en dimension infinie annoncé dans l'introduction.

Théorème 2.3.2. *Soit K un convexe fermé non vide d'un espace de Hilbert V et soit f une fonction convexe continue sur K . Alors il existe un unique minimum $x^* \in K$ de f sur K et on a*

$$\forall x \in K, \quad \|x - x^*\|^2 \leq \frac{4}{\alpha} (f(x) - f(x^*)).$$

En particulier, toute suite minimisante de f sur K converge vers x^ .*

Démonstration. La démonstration repose sur le résultat technique suivant :

Lemme 2.3.3. *Soit f une fonction α -convexe sur K . Alors, il existe deux constantes $\alpha_1 > 0$, $\alpha_2 \in \mathbb{R}$ t.q. :*

$$\forall x \in K, \quad f(x) \geq \alpha_1 \|x\|^2 + \alpha_2.$$

Démonstration du Lemme. C'est une conséquence du Théorème de Hahn-Banach géométrique appliqué à l'épigraphe de f noté $\mathcal{E}(f)$ et défini par :

$$\mathcal{E}(f) = \{(\lambda, x) \in \mathbb{R} \times K, f(x) \leq \lambda\}.$$

En effet, f étant convexe et continue, $\mathcal{E}(f)$ est fermé comme image réciproque du fermé $[0+\infty[$ de \mathbb{R} par l'application continue $(\lambda, x) \mapsto \lambda - f(x)$ et convexe par suite de la convexité de f . Soit $(\lambda_0, x_0) \in \mathbb{R} \times K$ t.q. $\lambda_0 < f(x_0)$. Alors $(\lambda_0, x_0) \notin \mathcal{E}(f)$ et d'après le Théorème de Hahn-Banach de séparation des convexes compacts et des convexes fermés, il existe une forme linéaire continue qui sépare strictement $\mathcal{E}(f)$ et le singleton $\{(\lambda_0, x_0)\}$ dans $\mathbb{R} \times V$ au sens suivant : il existe une forme linéaire continue $L \in V'$ et il existe $(\beta, \delta) \in \mathbb{R}^2$ t.q. :(3)

$$\forall (\lambda, x) \in \mathcal{E}(f), \quad \beta\lambda + L(x) > \delta > \beta\lambda_0 + L(x_0). \quad (3)$$

En considérant $\lambda \rightarrow +\infty$ dans (3), on voit que $\beta \geq 0$ et en prenant $x = x_0$ dans (3) on voit que $\beta > 0$. En remarquant que $(f(x), x) \in \mathcal{E}(f)$, $\forall x \in K$, on en déduit que

$$\forall x \in K, \quad \beta f(x) + L(x) > \delta.$$

De plus, l'uniforme convexité de f entraîne alors, compte tenu de la linéarité de L : $\forall (x, y) \in K^2$,

$$\begin{aligned} \frac{\beta}{2}(f(x) + f(y)) + \frac{1}{2}(L(x) + L(y)) - \frac{\alpha\beta}{8}\|x - y\|^2 \\ \geq_{\beta > 0} \beta f\left(\frac{x+y}{2}\right) + L\left(\frac{x+y}{2}\right) > \delta. \end{aligned}$$

Comme L est continue, il en résulte :

$$\begin{aligned} \frac{\beta}{2}(f(x) + f(y)) + \frac{1}{2}\|L\|\|x\| + \frac{1}{2}L(y) > \delta + \frac{\alpha\beta}{8}\|x - y\|^2 \\ = \delta + \frac{\alpha\beta}{8}(\|x\|^2 - 2\langle x, y \rangle + \|y\|^2) \geq \delta + \frac{\alpha\beta}{8}(\|x\|^2 - 2\|x\|\|y\| + \|y\|^2) \end{aligned}$$

i.e. :

$$f(x) > \frac{\alpha}{4}\|x\|^2 - \left(\frac{\alpha}{4}\|y\| + \frac{\|L\|}{\beta}\right)\|x\| + C \geq \frac{\alpha}{8}\|x\|^2 + C'$$

où $C, C' \in \mathbb{R}$ sont des constantes indépendantes de x . □

Du Lemme 2.3.3, on déduit que f est coercive, donc, K étant fermé par hypothèse, que le problème $\inf_K f$ admet une solution dans K .

Soit $(x_n)_{n \geq 0}$ une suite minimisante. On a : $\forall n, m \geq 0$,

$$f\left(\frac{x_m + x_n}{2}\right) \leq \frac{1}{2}f(x_m) + \frac{1}{2}f(x_n) - \frac{\alpha}{8}\|x_m - x_n\|^2. \quad (4)$$

On en déduit, dans un premier temps :

$$\begin{aligned} \inf_K f &\leq \liminf_{m,n \rightarrow +\infty} f\left(\frac{x_m + x_n}{2}\right) \leq \limsup_{m,n \rightarrow +\infty} f\left(\frac{x_m + x_n}{2}\right) \leq \\ &\leq \lim_{m \rightarrow +\infty} \frac{1}{2}f(x_m) + \lim_{n \rightarrow +\infty} \frac{1}{2}f(x_n) = \inf_K f \end{aligned}$$

i.e. :

$$f\left(\frac{x_m + x_n}{2}\right) \xrightarrow{m,n \rightarrow \infty} \inf_K f.$$

On obtient, en reportant dans (4) :

$$\begin{aligned} \inf_K f + \liminf_{m,n \rightarrow +\infty} \|x_m - x_n\|^2 &\leq \inf_K f + \limsup_{m,n \rightarrow +\infty} \|x_m - x_n\|^2 \leq \inf_K f \\ \Rightarrow \liminf_{m,n \rightarrow +\infty} \|x_m - x_n\|^2 &\leq \limsup_{m,n \rightarrow +\infty} \|x_m - x_n\|^2 \leq 0 \end{aligned}$$

donc :

$$\lim_{m,n \rightarrow +\infty} \|x_m - x_n\|^2 = 0$$

i.e. la suite $(x_n)_{n \geq 0}$ est de Cauchy dans V complet, donc convergente vers $x^* \in V$. Comme K est fermé, il en résulte que $x_n \xrightarrow{n \rightarrow +\infty} x^* \in K$, et par continuité de f : $f(x^*) = \inf_K f$.

On a d'autre part : $\forall n, p \geq 0$,

$$\frac{\alpha}{8}\|x_{n+p} - x_n\|^2 + f\left(\frac{x_{n+p} + x_n}{2}\right) \leq \frac{1}{2}f(x_{n+p}) + \frac{1}{2}f(x_n).$$

On en déduit, quand $p \rightarrow +\infty$: $\forall n \geq 0$,

$$\begin{aligned} \frac{\alpha}{8}\|x^* - x_n\|^2 + \inf_K f &\leq \frac{1}{2}\inf_K f + \frac{1}{2}f(x_n) \iff \\ \iff \frac{\alpha}{4}\|x^* - x_n\|^2 &\leq f(x_n) - \inf_K f \end{aligned}$$

□

Remarque 9. Le Théorème 2.3.2 reste vrai si on suppose f semi-continue inférieure. En effet, le même raisonnement montre que $x_n \xrightarrow{n \rightarrow +\infty} x^* \in K$ avec

$$\inf_K f \leq f(x^*) \leq \liminf_{n \rightarrow +\infty} f(x_n) = \lim_{n \rightarrow +\infty} f(x_n) = \inf_K f \Rightarrow f(x^*) = \inf_K f.$$

3 Conditions d'optimalité. Optimisation sans contrainte

Conformément au programme de l'Agrégation, on se limite au cas de la dimension finie, laissant en remarques des prolongements pour le cas de la dimension infinie.

Théorème 3.0.1 (Inéquation d'Euler). *Soit $K \subset V$ un sous-ensemble convexe d'un espace de Hilbert V . Soit $f : K \rightarrow \mathbb{R}$ différentiable sur K . Si x est un minimum local de f sur K alors x est solution de l'inéquation d'Euler :*

$$\forall y \in K, \quad \langle \nabla f(x), y - x \rangle \geq 0.$$

Si de plus f est convexe alors x est un minimum global de f sur K .

Démonstration. Soit $y \in K$. Par hypothèse sur f , l'application

$$\varphi : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f((1-t)x + ty)$$

est dérivable sur $[0, 1]$, de dérivée donnée par :

$$\forall t \in [0, 1], \quad \varphi'(t) = \langle \nabla f((1-t)x + ty), y - x \rangle.$$

De plus φ admet un minimum local en $t = 0$, i.e. il existe $r \in]0, 1[$ t.q. :

$$\forall t \in]0, r[, \quad \varphi(t) \geq \varphi(0) \Rightarrow \frac{1}{t}(\varphi(t) - \varphi(0)) \geq 0.$$

On en déduit :

$$\lim_{t \rightarrow 0} \frac{1}{t}(\varphi(t) - \varphi(0)) = \varphi'(0) \geq 0$$

i.e. $\langle \nabla f(x), y - x \rangle \geq 0$.

On suppose que de plus f est convexe. Alors, d'après le Théorème 2.2.1 :

$$\forall y \in K, \quad f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \geq f(x)$$

i.e. x est un minimum global. □

3.1 Conditions d'optimalité. Optimisation sans contrainte

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$. On considère le problème :

$$\min_{x \in \mathbb{R}^n} f(x). \quad (5)$$

On se propose d'étudier la généralisation au cas de la dimension $n \geq 2$ de la condition suffisante classique en dimension 1.

Théorème 3.1.1 (Conditions nécessaires). *Soit $x^* \in \mathbb{R}^n$ solution du problème (5).*

1. *Si f est différentiable en x^* , alors $\nabla f(x^*) = 0$ et x^* est appelé point stationnaire ou critique.*
2. *Si f est deux fois différentiable en x^* , alors la matrice $\nabla^2 f(x^*)$ est semi-définie positive.*

Démonstration. 1. On suppose f différentiable en x^* solution de (5). Soit $x \in \mathbb{R}^n$. Alors l'application

$$\varphi : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto f((1-t)x^* + tx) \quad (6)$$

admet un minimum en $t = 0$ et est dérivable en $t = 0$, de dérivée définie par : $\varphi'(0) = \langle \nabla f(x^*), x - x^* \rangle$. On a :

$$\forall t \in]0, \infty[, \quad \frac{\varphi(t) - \varphi(0)}{t} \geq 0$$

donc

$$\varphi'(0) = \lim_{t \rightarrow 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0.$$

On a de même : On a :

$$\forall t \in]-\infty, 0[, \quad \frac{\varphi(t) - \varphi(0)}{t} \leq 0$$

donc

$$\varphi'(0) = \lim_{t \rightarrow 0^-} \frac{\varphi(t) - \varphi(0)}{t} \leq 0,$$

ce qui entraîne que $\varphi'(0) = 0$, i.e. $\langle \nabla f(x^*), x - x^* \rangle = 0$. Ceci étant vrai $\forall x \in \mathbb{R}^n$, il en résulte que $\nabla f(x^*) = 0$.

2. On suppose f deux fois différentiable en x^* . Soit $x \in \mathbb{R}^n$. Alors φ définie par (6) admet une dérivée seconde en $t = 0$ définie par :

$$\varphi''(0) = \langle \nabla^2 f(x^*)(x - x^*), x - x^* \rangle.$$

De plus :

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + o(t^2)$$

$$\underset{\varphi'(0)=0}{=} \varphi(0) + \frac{t^2}{2}\varphi''(0) + o(t^2) \geq \varphi(0)$$

Si $\varphi''(0) \neq 0$, alors $\varphi(t) \sim \varphi(0) + \frac{t^2}{2}\varphi''(0)$ et $\varphi''(0) + o(1) > 0$, donc $\varphi''(0) > 0$. On en déduit que $\varphi''(0) \geq 0$ en général, i.e.

$$\langle \nabla^2 f(x^*)(x - x^*), x - x^* \rangle \geq 0.$$

Ceci étant vrai $\forall x \in \mathbb{R}^n$, il en résulte que $\nabla^2 f(x^*)$ est semi-définie positive. □

- Remarque 10.* 1. Si f est deux fois différentiable en x^* , la hessienne en x^* n'est pas en général définie positive. Par exemple, si $f(x) = x^4$, alors f atteint son minimum absolu sur \mathbb{R} en $x = 0$ et $f''(0) = 0$.
2. Le Théorème 3.1.1 fournit des conditions nécessaires mais non suffisantes. En effet, si $f(x) = x^3$, alors $f'(0) = f''(0) = 0$ et f n'a pas de minimum sur \mathbb{R} .

Théorème 3.1.2 (Conditions suffisantes). *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable en x^* t.q. $\nabla f(x^*) = 0$. On suppose en outre que :*

- (i) *soit $\nabla^2 f(x^*)$ est définie positive ;*
- (ii) *soit f est deux fois différentiable dans un voisinage de x^* , de hessienne semi-définie positive dans ce voisinage.*

Alors x^ est un minimum local pour f .*

Démonstration. (i) Soit $x \in \mathbb{R}^n$, $x \neq x^*$, et soit

$$\varphi : \mathbb{R} \rightarrow \mathbb{R}, \quad \varphi(t) = f((1-t)x^* + tx).$$

Alors φ est deux fois dérivable en 0 et on a :

$$\varphi'(0) = \langle \nabla f(x^*), x - x^* \rangle = 0$$

$$\varphi''(0) = \langle \nabla^2 f(x^*)(x - x^*), x - x^* \rangle > 0$$

On en déduit :

$$\varphi(t) = \varphi(0) + \frac{t^2}{2}\varphi''(0) + o(t^2)$$

i.e. :

$$\frac{\varphi(t) - \varphi(0)}{t^2} \underset{t \rightarrow 0}{\sim} \frac{1}{2}\varphi''(0) > 0$$

Donc il existe $\eta > 0$ t.q.

$$\varphi(t) > \varphi(0), \quad \forall t \in]-\eta, \eta[=: I_\eta$$

i.e. :

$$f((1-t)x^* + tx) \geq f(x^*), \quad \forall t \in I_\eta.$$

- (ii) Par hypothèse, f est deux fois différentiable dans un voisinage V_{x^*} de x^* . Soit $r > 0$ t.q. $B(x^*, r) \subset V_{x^*}$ et soit $x \in B(x^*, r) \setminus \{x^*\}$: $0 < \|x - x^*\| < r$. Avec les mêmes notations que dans (i), φ est deux fois différentiable dans un voisinage I_x de 0 défini par :

$$t \in I_x \iff |t|\|x - x^*\| < r \iff t \in \left] -\frac{r}{\|x - x^*\|}, \frac{r}{\|x - x^*\|} \right[=: I_x.$$

Soit $t \in I_x$. De la formule d'Euler-Mac Laurin, on déduit qu'il existe $c_t \in I_x$ t.q.

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(c_t) = \varphi(0) + \frac{t^2}{2}\varphi''(c_t) \geq \varphi(0).$$

□

On obtient un résultat optimal dans le cas convexe.

Théorème 3.1.3 (Cas convexe. Condition nécessaire et suffisante.). *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexe et différentiable sur \mathbb{R}^n . Une CNS pour que $x^* \in \mathbb{R}^n$ soit un minimum local (et donc global) de f est que x^* soit un point critique de f i.e. vérifie :*

$$\nabla f(x^*) = 0.$$

Démonstration. \Rightarrow Soit $x^* \in \mathbb{R}^n$ un point critique de f sur \mathbb{R}^n . D'après la caractérisation des fonctions convexes du Théorème 2.2.1 :

$$\forall x \in \mathbb{R}^n, \quad f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle = f(x^*)$$

i.e., x^* est un minimum global de f sur \mathbb{R}^n .

\Leftarrow Inversement on suppose que f admet un minimum local en x^* . Alors $\nabla f(x^*) = 0$ d'après le Théorème 3.1.1. □

3.2 Minimisation d'une fonctionnelle quadratique. Optimisation sans contrainte

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée réelle symétrique, soit $b \in \mathbb{R}^n$ et soit $c \in \mathbb{R}$. On définit la fonctionnelle quadratique (somme d'une forme quadratique et d'une fonction affine)

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle + c$$

et on considère le problème de minimisation :

$$\min_{x \in \mathbb{R}^n} f(x).$$

Dans l'Exemple 5, on a montré que f est deux fois différentiable sur \mathbb{R}^n , de différentielle définie par le gradient : $\nabla f(x) = Ax - b$, de matrice hessienne constante $\nabla^2 f(x) = A$. En particulier, f est convexe ssi A est semi-définie positive.

On en déduit que si A est semi-définie positive, alors f admet un minimum local (et donc global) s'il existe $x^* \in \mathbb{R}^n$ solution de $Ax^* = b$, ce qui est réalisé ssi $b \in \text{Im}(A)$ avec, en dimension finie : $\text{Im}(A) = \text{Ker}(A^T)^\perp = \text{Ker}(A)^\perp$.

Dans le cas général où A est symétrique réelle, A est diagonalisable dans une base, i.e. il existe $U \in \mathcal{O}_n(\mathbb{R})$ orthogonale t.q. $A = U^T D U$ avec D diagonale formée des valeurs propres de A , soit :

$$\lambda_1 \leq \dots \leq \lambda_n.$$

On distingue plusieurs cas suivant le signe de la plus petite valeur propre λ_1 .

(i) Si $\lambda_1 < 0$, soit $u_1 \in \mathbb{R}^n$ t.q. $Au_1 = \lambda_1 u_1$ et $\|u_1\| = 1$. On a : $\forall t \in \mathbb{R}$,

$$f(tu_1) = \frac{t^2}{2} \lambda_1 - t \langle b, u_1 \rangle + c = \frac{\lambda_1}{2} \left(t - \frac{\langle b, u_1 \rangle}{\lambda_1} \right)^2 + \frac{\langle b, u_1 \rangle^2}{2\lambda_1} + c$$

$$\text{avec } \lim_{t \rightarrow +\infty} \left| t - \frac{\langle b, u_1 \rangle}{\lambda_1} \right| = +\infty \Rightarrow$$

$$\lim_{t \rightarrow +\infty} f(tu_1) = -\infty$$

donc $\inf_{\mathbb{R}^n} f = -\infty$ et le problème (5) n'a pas de solution.

- (ii) Si $\lambda_1 = 0$ et $b \notin \text{Ker}(A)^\perp$, alors il existe $u \in \text{Ker}(A)$ t.q. $\langle b, u \rangle \neq 0$.
On a : $\forall t \in \mathbb{R}$,

$$f(tu) = -t\langle b, u \rangle + c.$$

On peut supposer, quitte à remplacer u par $-u$, que $\langle b, u \rangle > 0$. Alors

$$\lim_{t \rightarrow +\infty} f(tu) = -\infty \Rightarrow \inf_{x \in \mathbb{R}^n} f(x) = -\infty$$

et le problème (5) n'a pas de solution. Autrement dit, A étant semi-définie positive, f est convexe non minorée sur \mathbb{R}^n . En particulier, l'équation $Ax = b$ n'a pas de solution i.e. f n'a pas de point critique, car $b \notin (\text{Ker}(A))^\perp$.

- (iii) Si $\lambda_1 = 0$ et $b \in (\text{Ker}(A))^\perp$, alors on remarque que $(\text{Ker}(A))^\perp = \text{Im}(A^T) = \text{Im}(A)$ i.e. l'équation des points critiques $Ax = b$ est bien posée et admet une infinité de solutions. Plus précisément, $b \in \text{Ker}(A)^\perp = \text{Im}(A) \Rightarrow \exists x^* \in \mathbb{R}^n$ t.q. $b = Ax^*$ et alors

$$Ax = b \iff A(x - x^*) = 0 \iff x \in x^* + \text{Ker}(A).$$

Soit alors $x \in \mathbb{R}^n$ et soit $x = x' + x''$ la décomposition de x dans la somme directe $\mathbb{R}^n = \text{Ker}(A) \oplus \text{Ker}(A)^\perp$. On a

$$\langle Ax, x \rangle = \langle Ax'', x \rangle = \langle x'' Ax \rangle = \langle Ax'', x'' \rangle$$

et

$$b \in \text{Ker}(A)^\perp \Rightarrow \langle b, x \rangle = \langle b, x'' \rangle.$$

Donc $f(x) = f(x'')$ et on en déduit que $\inf_{x \in \mathbb{R}^n} f(x) = \inf_{x \in (\text{Ker}(A))^\perp} f(x)$. Or $\lambda_1 = 0 \Rightarrow \nabla^2 f(x) = A$ est semi-définie positive, $\forall x \in \mathbb{R}^n$, donc f est convexe et tout minimum local est un minimum global. De plus, l'équation $\nabla f(x) = 0 \iff Ax = b$ admet une unique solution x^* dans $(\text{Ker}(A))^\perp$. On en déduit que les solutions du problème (5) sont les vecteurs de la forme $x^* + x$, $x \in \text{Ker}(A)$ où x^* est l'unique solution de

$$Ax^* = b \quad \text{et} \quad x^* \in (\text{Ker}(A))^\perp.$$

Le calcul montre directement que :

$$\inf_{x \in \mathbb{R}^n} f(x) = f(x^*) = -\frac{1}{2}\langle b, x^* \rangle + c.$$

- (iv) Si $\lambda_1 > 0$, alors $\text{Im}(A) = \mathbb{R}^n$ et l'équation $Ax = b$ admet une unique solution x^* . De plus A est définie positive donc f est (strictement) convexe et $x^* = A^{-1}b$ est l'unique solution du problème (5). On a

$$\inf_{x \in \mathbb{R}^n} f(x) = f(x^*) = -\frac{1}{2}\langle b, A^{-1}b \rangle + c.$$

Bibliographie

- [1] J.P. Demailly, Analyse Numérique et Equations Différentielles, P.U.Grenoble, Paris, 2006.
- [2] H. Brezis, Analyse fonctionnelle : théorie et applications, Masson, Paris, 1983.
- [3] P. Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation, Masson, Dunod, Paris.
- [4] P. Lascaux, R. Théodor, Analyse numérique matricielle appliquée à l'art de l'ingénieur, Dunod, Paris.
- [5] J.-B. Hiriart-Urruty, Convex Analysis and Minimization Algorithms I, Springer-Verlag, 1996.
- [6] C. Zuily, H. Queffélec, Analyse pour l'Agrégation, 3ème Edition, Dunod, 2007.
- [7] G. Allaire, Analyse Numérique et Optimisation, Editions de l'Ecole Polytechnique, 2005.
- [8] G. Allaire, S.M. Kaber, Numerical Linear Algebra, Texts in Applied Mathematics, vol. 55, Springer, 2008.
- [9] P. Lascaux, R. Théodor, Analyse numérique matricielle appliquée à l'art de l'ingénieur. Tome 1. Dunod, Paris, 1986.
- [10] M. Bergounioux, Optimisation et contrôle des systèmes linéaires, Dunod, 2001.
- [11] J.F. Bonnans, J.C. Gilbert, C. Lemaréchal, C. Sagastizabal, Optimisation numérique, coll. SMAI Mathématiques et Applications 27, Springer, 1997.