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 $f : U \to F$ 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.

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

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 $f'_{x_0} : E \to F$ en un point x_0 est continue sur E. Si en outre l'application résultante $f' : U \to \mathcal{L}(E, F)$, $x \mapsto f'_x$, 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
$$f'_{x_0}(h) = \lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t}.$$
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 dérivéee directionnelle dans toutes les directions alors que la réciproque est fausse.

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
$$f(x,y) = \begin{cases} \dfrac{x^3}{x+y} & \text{si} \quad x \neq -y, \\ 0 & \text{sinon} \end{cases}$$
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
$$\lim_{(x,y) \to (0,0),\, y = \frac{x^3}{\lambda} - x} f(x,y) = \lambda \neq 0.$$
De plus, $\forall (x,y) \in \mathbb{R}^2$, $x \neq -y$,
$$\forall t \in \mathbb{R}^*, \quad f(tx,ty) = t^2 f(x,y) \Rightarrow \lim_{t \to 0} \frac{f(tx,ty)}{t} = 0 =: f'_{(0,0)}(x,y).$$

Définition 1.1.2 (Différentiabilité d'ordre supérieur)**.** On dit que $f : U \subset E \to F$ est deux fois différentiable sur U si f est différentiable sur U de différenielle f' différentiable sur U et on note :
$$f''_x := (f')'_x, \quad \forall, x \in U$$
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 $f : U \subset R \to F$ admet une différentielle d'ordre 2 sur U, alors pour tout $x \in U$, l'application $f''_x : E \times E \to F$ est une forme bilinéaire symétrique sur $E \times E$ à valeurs dans F.*

Démonstration. On peut définir les applications :

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

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 monrer 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 :

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

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

De même :

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

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

Il en résulte :

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

□

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 àvaleurs dans F et on définit et on note $f^{(n)} \in \mathcal{L}_n(E, F)$ la différentielle d'ordre n de f. Obn dit que $f : U \subset E \to F$ 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 $f : H \to \mathbb{R}$ 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 voisnage 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. :

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

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 $f : U \subset E \to F$ 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 :*

$$f(x+h) = f(x) + \sum_{k=1}^{n} \frac{1}{k'} \underbrace{f_x^{(k)}(h, \cdots, h)}_{:= f_x^{(k)} \cdot h^k} + \int_0^1 \frac{(1-t)^n}{n!} f_{x+th}^{(n+1)} dt \cdot h^{n+1}$$

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

Remarque 3. Si $f : U \subset E \to F$ 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 $f : U \subset E \to \mathbb{R}$ admet une différentielle d'ordre 2 sur U et que $a \in U$ est un point critique de f. Si f''_a est définie positive 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 pose :

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

Alors, φ est deux fois dérivable sur $]-r,r[$ et on a :

$$\varphi(t) = \varphi(0) + \int_0^t \varphi'(s)ds, \quad \forall t \in]-r,r[.$$

De plus : $\varphi''(0) = f''_a \cdot h^2 > 0$, donc il existe $\varepsilon \in]0,r[$, t.q. φ' soit croussnate sur $]-\varepsilon,\varepsilon[$. Soit $|t| < \varepsilon$. Si $0 < t < \varepsilon$, alors

$$\int_0^t \varphi'(s)ds \geq t\varphi'(0) = 0 \Rightarrow \varphi(t) \geq \varphi(0).$$

Si $-\varepsilon < t < 0$, alors

$$\varphi(t) = \varphi(0) - \int_t^0 \varphi'(s)ds$$

avec

$$\int_t^0 \varphi'(s)ds \leq -t\varphi'(0) = 0$$

donc $\varphi(t) \geq \varphi(0)$. Finalmentn dans tous les cas :

$$\varphi(t) \geq \varphi(0), \quad \forall t \in]-\varepsilon,\varepsilon[,$$

i.e. a est un minimum local de f. □

Remarque 4. Si $E = \mathbb{R}^N$ est de dimension finie et si f''_a est définie positive alors elle définit un produit scalaire sur E. Soit $(e_1, \cdots, e_N)$ une base orthonormée pour f''_a. Soit $r > 0$ t.q. $B(a,r) \subset U$ et soit $h = \sum_{i=1}^N h_i e_i \in B(0,r)$. On a, si a est en outre un point critique de f :

$$f(a+h) = f(a) + f''_a \cdot h^2 + o(\|h\|^2)$$

avec

$$f''_a \cdot h^2 = \sum_{i=1}^N h_i^2 f''_a \cdot e_i^2.$$

Soit $\alpha = \min_{i=1}^N f''_a \cdot e_i^2 > 0$. On a :

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

et donc :

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

On en déduit que pour $\varepsilon \in]0, r[$ suffisamment petit :

$$\frac{f(a+h) - f(a)}{\|h\|^2} \geq \frac{\alpha}{2} > 0, \quad \forall h \in B(0, \varepsilon).$$

Proposition 1.2.4. *On suppose que $f : U \subset E \to \mathbb{R}$ est de classe $\mathcal{C}^2$ sur U et que $a \in U$ est un point critique de f. Si $f''_x \geq 0$ dans un voisinage de a, alors a est un minimum local de f.*

Démonstration. Soit $r > 0$ t.q. $B(a,r) \subset U$ et $f''_x \geq 0$ dans $B(a,r)$. Soit $h \in B(0,r)$. On a :

$$fa + h) = f(a) + \int_0^1 (1-t) \underbrace{f''_{a+th} \cdot h^2}_{\geq 0} dt \geq f(a).$$

□

Remarque 5 (Différentiabilité d'ordre supérieur). Si V est un espace de Hilbert et si $f : V \to \mathbb{R}$ 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 $df : U \subset V \to \mathcal{L}(V, \mathbb{R}) = V' \sim V$ est différentiable en x_0, i.e. s'il existe une application linéaire continue $L_{x_0} \in \mathcal{L}(V, V)$ t.q. :

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

en posant $d^2 f_{x_0} := L_{x_0}$, 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 $\nabla^2 f(x_0)\xi := d^2 f_{x_0}(\xi)$ 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 $(e_1, \cdots, e_n)$ la base canonique de $\mathbb{R}^n$ et on munit $\mathbb{R}^n$ de sa structure euclidienne usuelle.

Proposition 1.3.1. *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$. 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 :

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

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

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

□

Proposition 1.3.2. *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$. 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 latrice 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 :

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

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 :

$$f''_x(e_i, e_j) = \frac{f'_{x+he_j}(e_i) - f'_x(e_i)}{h} + o(1) = \lim_{h \to 0} \frac{f'_{x+he_j}(e_i) - f'_x(e_i)}{h}$$

$$= \lim_{h \to 0} \frac{1}{h} \left(\frac{\partial f}{\partial x_i}(x + he_j) - \frac{\partial f}{\partial x_i}(x) \right) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

□

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 $f : U \subset \mathbb{R}^n \to \mathbb{R}$.

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

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

On en déduit :

$$\frac{|f(x,y)|}{\sqrt{x^2+y^2}} \leq \frac{|x||y|(x^2+y^2)}{(x^2+y^2)^{3/2}} = \frac{|x||y|}{(x^2+y^2)^{1/2}} \leq \frac{(x^2+y^2)}{2(x^2+y^2)^{1/2}} =$$

$$= \frac{1}{2}(x^2+y^2)^{1/2} \underset{(x,y)\to(0,0)}{\longrightarrow} 0.$$

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

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

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

Il en résulte :

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

i.e. :

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

De même :
$$\lim_{x\to 0}\frac{1}{x}\left(\frac{\partial f}{\partial y}(x,0)-\frac{\partial f}{\partial y}(0,0)\right)=\lim_{x\to 0}\frac{1}{x}\frac{\partial f}{\partial y}(x,0)=\lim_{x\to 0}\frac{x}{x}=1.$$
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 $f : U \subset \mathbb{R}^n \to \mathbb{R}$ 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 $f : U \subset \mathbb{R}^n \to \mathbb{R}$ 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.

2.1 Existence en dimension finie

Soit $f : K \subset \mathbb{R}^n \to \mathbb{R}$ continue. On considère le problème :
$$\min_{x\in K} f(x). \tag{1}$$
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 :

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

Par construction :

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

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

Remarque 7 (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 :

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

alors le Théorème 2.1.1 s'applique.

Remarque 8 (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 : $f(x_0) \leq \liminf_{x\to x_0} f(x)$.

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

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 :

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

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 :

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

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

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

On s'intéresse au problème

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

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

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

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 $f : K \subset \mathbb{R}^n \to \mathbb{R}$ 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 $f : K \subset \mathbb{R}^n \to \mathbb{R}$ 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 $f : \mathbb{R}^n \to \mathbb{R}$ 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 $f : \mathbb{R}^n \to \mathbb{R}$ 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 $f : \mathbb{R}^n \to \mathbb{R}$ 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.

Démonstration. 1. $(i) \Rightarrow (ii)$: On pose :

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

Par hypothèse sur f, φ 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, la définition de la convexité entraîne directement :

$$\forall t \in [0,1], \quad \varphi(t) = \varphi(t + (1-t)0) \leq t\varphi(1) + (1-t)\varphi(0). \tag{3}$$

On en déduit :

$$\forall t \in]0,1], \quad \varphi(1) \geq \varphi(0) + \frac{1}{t}(\varphi(t) - \varphi(0))$$

puis, quand $t \to 0^+$:

$$\varphi(1) \geq \varphi(0) + \varphi'(0) \tag{4}$$

i.e. (ii) :

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

$(ii) \Rightarrow (iii)$ Par hypothèse :

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \quad \text{et} \quad f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$$

d'où, par linéarité du produit scalaire :

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x) \leq \langle \nabla f(y), y - x \rangle$$

$$\Rightarrow \langle \nabla f(x), y - x \rangle \leq \langle \nabla f(y), y - x \rangle$$

i.e. (iii).
$(iii) \Rightarrow (i)$ Soit $t, s \in]0,1[$, $t \neq s$. Par hypothèse sur f, φ est continue sur $[0,1]$ et dérivable sur $]0,1[$. On a :

$$\langle \nabla f((1-t)x + ty) - \nabla f((1-s)x + sy), (t-s)(y-x) \langle \underset{(iii)}{\geq} 0$$

i.e. :

$$(t-s)(\varphi'(t) - \varphi'(s)) \geq 0.$$

Cela donne, en divisant par $|t-s|^2 > 0$:

$$\frac{\varphi'(t) - \varphi'(s)}{t - s} > 0$$

i.e. φ' est croissante sur $]0,1[$.
Soit $t \in]0,1[$. On a :

$$f((1-t)x + ty) - (1-t)f(x) - tf(y) = \varphi(t) - (1-t)\varphi(0) - t\varphi(1)$$

$$= (1-t)(\varphi(t) - \varphi(0)) + t(\varphi(t) - \varphi(1))$$

Du Théorème des valeurs intermédiaires on déduit qu'il existe $u \in]0,t[$ et $s \in]t,1[$ t.q. :

$$\varphi(t) - \varphi(0) = t\varphi'(u), \quad \varphi(t) - \varphi(1) = (t-1)\varphi'(s).$$

On en déduit :

$$f((1-t)x+ty) - (1-t)f(x) - tf(y) = \underbrace{(1-t)t}_{\geq 0}\underbrace{(\varphi'(u) - \varphi'(s))}_{\leq 0}$$

$$\leq 0$$

2. $(i) \Rightarrow (ii)$. On suppose f strictement convexe. Soit $s,t \in]0,1[$ et soit $x,y \in \mathbb{R}^n$, $x \neq y$. On a

$$\varphi(t) < (1-t)\varphi(0) + t\varphi(1) \underset{t>0}{\Rightarrow} \frac{\varphi(t)-\varphi(0)}{t} < \varphi(1) - \varphi(0).$$

De même :

$$\varphi(st) < (1-s)\varphi(0) + s\varphi(t) \underset{s>0}{\Rightarrow} \frac{\varphi(st)-\varphi(0)}{st} < \frac{\varphi(t)-\varphi(0)}{t} < \varphi(1) - \varphi(0).$$

On en déduit, quand $s \to 0^+$:

$$\varphi'(0) \leq \frac{\varphi(t)-\varphi(0)}{t} < \varphi(1) - \varphi(0)$$

i.e. (ii).

$(ii) \Rightarrow (iii)$ Soit $x \neq y$. Par hypothèse :

$$f(y) > f(x) + \langle \nabla f(x), y-x \rangle \quad \text{et} \quad f(x) > f(y) + \langle \nabla f(y), x-y \rangle$$

i.e. :

$$\langle \nabla f(x), y-x \rangle < f(y) - f(x) < \langle \nabla f(y), y-x \rangle$$

donc

$$\langle \nabla f(x), y-x \rangle < \langle \nabla f(y), y-x \rangle$$

i.e. (iii).

$(iii) \Rightarrow (i)$ Soit $t \in]0,1[$. Par le même raisonnement que dans 1., il existe $u \in]0,t[$ et $s \in]t,1[$ t.q., en posant $w = (1-u)x + uy$, $z = (1-s)x + sy$:

$$(1-t)f(x) + tf(y) = f((1-t)x+ty) + t(1-t)\langle \nabla f(w) - \nabla f(z), x-y \rangle =$$

$$= f((1-t)x+ty) + \frac{t(1-t)}{s-u}\langle \nabla f(w) - \nabla f(z), w-z \rangle \underset{(iii)}{>} f((1-t)x+ty).$$

3. $(i) \Rightarrow (ii)$ D'après le Théorème de Schwartz, f étant deux fois différentiable sur $\mathbb{R}^n$, sa matrice hessienne est symétrique en tout point de $\mathbb{R}^n$. Par hypothèse, φ est deux fois dérivable, de dérivée seconde donnée par : $\forall t \in [0, 1]$,

$$\varphi''(t) = \langle \nabla^2 f((1-t)x+ty)(y-x), y-x\rangle \Rightarrow \langle \nabla^2 f(x)(y-x), y-x\rangle = \varphi''(0)$$

avec

$$\varphi''(0) = \lim_{t\to 0^+} \frac{1}{t}(\varphi'(t) - \varphi'(0)).$$

On remarque que, par hypothèse :

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

donc, en passant à la limite quand $t \to 0^+$: $\varphi''(0) \geq 0$, i.e. (ii).
$(ii) \Rightarrow (i)$ Soit $x, y \in \mathbb{R}^n$. Par hypothèse,

$$\forall t \in [0, 1], \quad \varphi''(t) = \langle \nabla^2 f((1-t)x + ty)(y - x), y - x\rangle \geq 0,$$

donc φ' est croissante sur $[0, 1]$. On en déduit, φ étant de classe $\mathcal{C}^1$ sur $[0, 1]$:

$$\varphi(1) = \varphi(0) + \int_0^1 \varphi'(t)dt \geq \varphi(0) + \varphi'(0)$$

i.e. :

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle.$$

De 1., on déduit que f est convexe, i.e. (i).

□

Corollaire 2.2.2. *Si $f : U \subset \mathbb{R}^n \to \mathbb{R}$ est convexe et différentiable sur U, alors tout point critique $a \in U$ est un minimum local.*

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 :

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

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 :

$$\frac{1}{\|h\|}\left|f(x+h)-f(x)-\langle Ax-b,h\rangle\right| = \frac{\langle Ah,h\rangle}{2\|h\|} \le \frac{\|A\|}{2}\|h\| \underset{h\to 0}{\to} 0$$

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

1. *Tout mininum 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.

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

Soit $x\in K$, $\|x-x^*\|\ge r$. 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

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

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

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

On en déduit :

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

$$\iff \frac{r}{2\|x - x^*\|} f(x^*) \leq \frac{r}{2\|x - x^*\|} f(x) \underset{r>0}{\iff} f(x^*) \leq f(x).$$

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.

$\square$

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 :

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

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 :

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

On a :

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

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 $\ell^2(\mathbb{N})$.

Théorème 2.3.1 (Inéquation d'Euler)**.** *Soit $K \subset V$ un sous-ensemble convexe d'un espace de Hilbert V. Soit $f : K \to \mathbb{R}$ 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

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

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 :

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

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. $\square$