# Option Calcul Scientifique: Optimisation 1

Préparation Agrégation de Mathématiques
Université de Rennes 1
Isabelle Gruais

6 novembre 2020

# 1 Introduction

## 1.1 Le programme de l'Agrégation

— Optimisation et approximation
— Extrema des fonctions réelles de $n$ variables réelles sans contraintes.
— Extrema des fonctions réelles de $n$ variables réelles avec contraintes : multiplicateurs de Lagrange.
— Méthode des moindres carrés.

## 1.2 Vocabulaire

Soit $(V, \|\cdot\|)$ un espace vectoriel normé.
On considère le problème d'optimisation :

$$\inf_{x \in K} f(x) \tag{1}$$

avec $K \subset V$ et $f : K \to \mathbb{R}$ fonction coût ou critère.

Si $K = V$, resp. $K \neq V$, le problème (1) est dit sans contrainte, resp. avec contrainte.

Si $\dim K < +\infty$, resp. $\dim K = +\infty$, le problème (1) est dit d'optimisation en dimension finie, resp. en dimension infinie.

Sans perte de généralité, on peut se limiter au cas d'un problème de minimiation en remarquant que

$$\inf_{x \in K} f(x) \iff \sup_{x \in K}(-f(x))$$

On rappelle que si $x_0 \in V$ est solution de

$$x_0 \in K \quad \text{et} \quad f(x_0) = \inf_{x \in K} f(x)$$

alors on dit que la borne inf de (1) est atteinte en $x_0 \in K$ et on note

$$f(x_0) = \min_{x \in K} f(x).$$

Par abus de notation, on dit que $x_0$ est un minimum pour $f$ sur $K$.

Si $V = \mathbb{R}$, on a le résultat fondamental :

**Proposition 1.2.1.** *Toute partie $X \subset \mathbb{R}$ de $\mathbb{R}$ non vide et minorée admet une borne inférieure et on a la caractérisation :*

$$m = \inf X \iff \begin{cases} \forall x \in X,\ m \leq x, \\ et \\ \forall \varepsilon > 0,\ \exists x_\varepsilon \in X,\ t.q.\ m \leq x_\varepsilon < m + \varepsilon. \end{cases}$$

On en déduit le critère équivalent :

**Corollaire 1.2.2** (Suites minimisantes)**.** *Si $X \subset \mathbb{R}$ est une partie non vide minorée de $\mathbb{R}$, alors les propositions suivantes sont équivalentes :*

*(i) $m = \inf X$*

*(ii) $m$ est un minorant de $X$ et il existe une suite $(x_n)_{n \geq 0}$ de points de $X$ appelée minimisante qui converge vers $m$.*

Le plan d'étude est le suivant :

(i) La théorie permet de déterminer si le problème de minimisation admet une solution.

**Exemple 1.** Une application continue $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ est majorée et minorée sur $[a, b]$ et atteint ses bornes sur ce compact.

(ii) Si oui, on identifie une telle solution grâce aux conditions nécessaires d'optimalité basées sur un développement de Taylor de $f$ quand il existe : du premier ordre si $f$ est de classe $\mathcal{C}^1$, du second ordre si $f$ est de classe $\mathcal{C}^2$.

**Exemple 2.** Si $f(x) = x^2$, $\inf_{x \in [0,1]} f(x) = 0$ est atteint pour $x = 0$ solution de $f'(x) = 0$.

**Exercice 1.** Soit $p \in ]0, +\infty[$ et soit $a = (a_n)_{n \geq 0} \in \ell^p$. Montrer que si $\frac{1}{p'} = 1 - \frac{1}{p}$, le problème

$$\sup_{u \in \ell^{p'}} \frac{|\sum_{n \geq 0} a_n u_n|}{\|u\|_{p'}}$$

est bien défini et admet une solution.

(iii) Si le problème de minimisation n'admet pas de solution, on cherche une suite minimisante, i.e. une suite $(x_n)_{n\geq 0} \in K^{\mathbb{N}}$ de $K$ t.q. $f(x_n) \underset{n\to+\infty}{\to} \inf_K f$.

**Exemple 3.** La suite de terme général $u_n = 1 - \frac{1}{n}$, $n \geq 1$, est une suite minimisante pour le problème : $\sup]0,1[= 1$ qui n'a pas de solution dans $]0,1[$.

**Exercice 2.** Etudier $\sup_{u=(u_n)_{n\geq 0}\in c_0} |\sum_{n\geq 0} \frac{u_n}{2^n}|$.

(iv) Dans ce cours, on laisse de côté la question du choix de méthodes numériques pour calculer la solution du problème de minimisation quand la théorie ne permet pas de l'expliciter.

On termine par des exemples.

**Exemple 4** (Dimension finie)**.** *Déterminer le volume maximal d'un parallélépipède rectangle de surface extérieure donnée égale à* $6$.

On note $a$, $b$, $c$ les longueurs des arêtes d'un parallélogramme rectangle donné. On est ramené à résoudre :

$$\sup_K V(a,b,c) \text{ avec } V(a,b,c) = abc,\ K = \{(a,b,c) \in (\mathbb{R}^+)^3,\ ab+bc+ca = 3\}$$

i.e. un problème d'optimisation avec contrainte.

**Exemple 5.** [Dimension infinie : problème dit de la Reine Didon] *Déterminer la portion d'aire maximale enclose par une courbe de longueur donnée* $L$ *et le segment de ses extrémités données* $[a,b]$.

Soit $Y = \mathcal{C}^1([a,b])$ et soit

$$K = \{c \in Y,\ \int_a^b \sqrt{1+y'(x)^2}dx = L,\ y(a) = y(b) = 0\}.$$

D'après la théorie de l'intégration, on est ramené à résoudre :

$$\sup_{c\in K} \int_a^b c(x)dx.$$

On remarque que l'espace fonctionnel $Y$ a été choisi pour que le problème d'optimisation soit bien défini.

## 1.3 Rappels de calcul différentiel

L'exemple 5 a montré que le cadre naturel de la théorie de l'optimisation est la classe $\mathcal{C}^1$ construite sur la notion de différentiabilité.

**Définition 1.3.1** (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 $df_{x_0} \in \mathcal{L}(E, F)$ appelée la différentielle de $f$ en $x_0$ t.q.

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

i.e. :

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

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 differenielle $df_{x_0} : E \to F$ en un point $x_0$ est continue. Si en outre l'aplication résultante $df : U \to \mathcal{L}(E, F)$, $x \mapsto df_x$, 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 $df_{x_0}(h)$ pour $h \in E$ donné s'effectue grâce à la formule

$$df_{x_0}(h) = \lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t}.$$

En toute rigueur, ce calcul donne la dérivée directionnelle de $f$ en $x_0$, ou diffŕentielle 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 :

$$\begin{array}{ccccc} f \text{ de classe } \mathcal{C}^1 \text{ en } x_0 & \Rightarrow & f \text{ est différentiable en } x_0 & \Rightarrow & f \text{ de classe } \mathcal{C}^0 \text{ en } x_0 \\ & & \Downarrow & & \\ & & f \quad \text{dérivable en } x_0 \text{ suivant toutes les directions} & & \end{array}$$

**Exemple 6** (Contre-exemples)**.**

1. 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 uen 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, $\forall (x,y) \in \mathbb{R}^2$, $x \neq -y$,

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

2. L'application $f$ définie par

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

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

On a

$$\lim_{(x,y)\to(0,0),\ x\neq y^2} f(x,y) = 0 = f(0,0)$$

$$\lim_{(x,y)\to(0,0),\ x=y^2} f(x,y) = \lim_{(x,y)\to(0,0),\ x=y^2} x = 0 = f(0,0)$$

donc $f$ est continue en $(0,0)$.

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

$$\forall t \in \mathbb{R}, \quad tx = (ty)^2 \iff t(x-ty^2) = 0 \iff t = 0 \quad \text{ou} \quad x = ty^2$$

$$\Rightarrow \lim_{t\to 0,\ t\neq 0} \frac{f(tx,ty)}{t} = \lim_{t\to 0,\ t\neq 0} \frac{ty^2}{t} = y^2 \neq 0 \quad \text{si} \quad y \neq 0.$$

$$\lim_{t\to 0,\ t\neq 0} \frac{f(tx,0)}{t} = 0 \quad \text{si} \quad x \neq 0.$$

Donc $f$ admet des dérivées directionnelles en $(0,0)$ dans toutes les directions de $\mathbb{R}^2$ qui dépendent de la direction $(x,y)$. Donc $f$ n'est pas différentiable en $(0,0)$.

*Remarque* 2 (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 $x_0$.

## 1.4 Détour par la dimension finie

Le programme de l'Agrégation pour l'optimisation concerne essentiellement le cas de la dimension finie sur lequel on revient en détail. 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.

**Définition 1.4.1** (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 $U$.

Dans la suite, on considère $f : U \subset \mathbb{R}^n \to \mathbb{R}$. Soit $x_0 \in K$.

(i) On suppose $f$ différentiable en $x_0$. Alors :

$$f(x_0+h) = f(x_0) + \langle \nabla f(x_0), h\rangle_V + \underset{h\to 0}{o}(\|h\|)$$

où $\nabla f(x_0)$ est le gradient de $f$ en $x_0$, i.e. le vecteur de composantes $\frac{\partial f}{\partial x_1}(x_0), \cdots, \frac{\partial f}{\partial x_n}(x_0)$. La définition du gradient dépend du choix du produit scalaire et est une conséquence du théorème de représentation de Riesz appliqué à la différentielle de $f$ en $x_0$. En dimension finie, le choix du produit scalaire canonique conduit aux formules ci-dessus pour le gradient et la matrice hessienne et peuvent alors être considérées comme génériques.

(ii) Si $f$ est deux fois différentiable en $x_0$ alors

$$f(x_0+h) = f(x_0) + \langle \nabla f(x_0), h\rangle + \frac{1}{2}\langle \nabla^2 f(x_0)h, h\rangle + \underset{h\to 0}{o}(\|h\|^2)$$

où $\nabla^2 f(x_0)$ est la matrice hessienne de $f$ en $x_0$ formée par les dérivées partielles secondes de $f$ en $x_0$ soit :

$$(\nabla^2 f(x_0))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0), \quad i, j \in [[1, n]].$$

D'après le Théorème de Schwartz, si $f$ est deux foix différentiable en $x_0$ alors la matrice hessienne est symétrique (réelle).

*Remarque* 3 (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 :

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

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

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 1.4.1** (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 1.4.2** (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.

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

$$\min_{x \in K} f(x). \tag{2}$$

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 (2) 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* 4 (Deux remarques très utiles en pratique...). Il reste à voir comment le Théorème 2.1.1 s'applique au problème 2. On commence par rappeler que l'hypothèse de dimension finie est essentielle. 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* 5 (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 7.** 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 8.** 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 9.** 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

$$\inf_{x\in\mathbb{R}^N} f(x). \tag{3}$$

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 (3) admet une solution.

## 2.2 Unicité de l'optimum

La convexité est une condition suffisante d'unicité dans les problèmes d'optimum. On commence par rappeler les définitions.

**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

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

On rappelle qu'en dimension finie, toute fonction convexe a une régularité minimale dans le sens suivant.

**Proposition 2.2.1.** *Si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe définie sur un ouvert de $\mathbb{R}^n$, alors $f$ est continue sur $\Omega$ et lipschitzienne sur tout compact de $\Omega$.*

*Démonstration.* Voir par exemple [5] pour la démonstration dans $\mathbb{R}^n$, et [6] pour le cas $n = 1$. □

**Corollaire 2.2.2.** *Une fonction convexe $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ définie sur un ouvert de $\mathbb{R}^n$ est différentiable presque partout (au sens de la mesure de Lebesgue) sur $\Omega$.*

*Démonstration.* C'est une conséquence de la propriété de Lipschitz et du Théorème de Rademacher. □

On rappelle un résultat classique très utile en pratique mais nécessitant de la régularité.

**Théorème 2.2.3** (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 $f$, $\varphi$ 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 :

$$\forall t \in [0,1], \quad \varphi(t) = \varphi(t + (1-t)0) \leq t\varphi(1) + (1-t)\varphi(0). \tag{4}$$

On en déduit :

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

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

$$\varphi(1) \geq \varphi(0) + \varphi'(0) \tag{5}$$

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 $t \in [0,1]$. Par hypothèse sur $f$, $\varphi$ est continue sur $[0,1]$ et dérivable sur $]0,1[$. Du Théorème des valeurs intermédiaires on déduit qu'il existe $u \in ]0,t[$ et $s \in ]t,1[$ t.q. :

$$\varphi(0) = \varphi(t) - t\varphi'(u) \iff f(x) = f((1-t)x+ty) - t\langle \nabla f((1-u)x+uy), y-x \rangle,$$
$$\varphi(1) = \varphi(t) + (1-t)\varphi'(s) \iff$$
$$\iff f(y) = f((1-t)x+ty) + (1-t)\langle \nabla f((1-s)x+sy), y-x \rangle.$$

On en déduit, par combinaison linéaire des deux égalités :

$$(1-t)f(x) + tf(y) = (1-t+t)f((1-t)x+ty)+$$
$$-(1-t)t\langle \nabla f((1-u)x+uy), y-x \rangle + t(1-t)\langle \nabla f((1-s)x+sy), y-x \rangle.$$
$$= f((1-t)x+ty) + t(1-t)\langle \nabla f((1-u)x+uy) - \nabla f((1-s)x+sy), x-y \rangle.$$

On pose

$$w = (1-u)x + uy, \quad z = (1-s)x + sy.$$

Par construction $u < s$, ce qui entraîne :

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

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, $\varphi$ 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 $\varphi'$ est croissante sur $[0,1]$. On en déduit, $\varphi$ 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).

$\square$

**Exemple 10** (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 :

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

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 $\nabla f$ 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 2.2.3 que $f$ est convexe, resp. strictement convexe, ssi $A$ est semi-définie positive, resp. strictement définie positive.

**Théorème 2.2.4.** *On conidère le problème (2) 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 (2) 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 :

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

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

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.

□

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

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

**Définition 2.3.1** (Fonction $\alpha$-ellitique)**.** Soit $V$ un espace de Hilbert, $\langle \cdot, \cdot \rangle$ un produit scalaire sur $V$. Soit $K \subset V$ un convexe. Une fonction $f : K \to \mathbb{R}$ est dite fortement convexe ou uniformément convexe ou $\alpha$-convexe ou $\alpha$-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 11** (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 $\alpha$-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 $\mathcal{C}^\infty$ 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

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

donc

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

De même :

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

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

donc

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

On en déduit, quand $x \to +\infty$ :

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

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 :

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

La proposition ci-dessous précise les liens entre convexité et uniforme convexité.

**Proposition 2.3.1.** *Soit $(V, \langle\cdot,\cdot\rangle)$ un espace de Hilbert et soit $f : V \to \mathbb{R}$ une application.*

1. *La fonction $f$ est $\alpha$-elliptique ssi $f - \frac{\alpha}{2}\|\cdot\|^2$ est convexe.*
2. *On suppose $f$ continue. Alors $f$ est $\alpha$-elliptique ssi il existe $\alpha > 0$ t.q. :*

$$\forall (x,y)\in V^2, \quad f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f(x)+f(y)) - \frac{\alpha}{8}\|x-y\|^2.$$

*Démonstration.* 1. $\Rightarrow$ On suppose $f$ $\alpha$-elliptique et on pose $\varphi = f - \frac{\alpha}{2}\|\|^2$. Alors : $\forall t\in[0,1]$, $\forall x,y\in V$,

$$\varphi(tx+(1-t)y) = f(tx+(1-t)y) - \frac{\alpha}{2}\|tx+(1-t)y\|^2$$

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

avec

$$t(1-t)\|x-y\|^2 + \|tx+(1-t)y\|^2 = t(1-t)(\|x\|^2+\|y\|^2-2\langle x,y\rangle)+$$

$$+t^2\|x\|^2 + (1-t)^2\|y\|^2 + 2t(1-t)\langle x,y\rangle$$

$$= t(1-t)(\|x\|^2+\|y\|^2) + t^2\|x\|^2 + (1-t)^2\|y\|^2$$

$$= t\|x\|^2 + (1-t)\|y\|^2$$

d'où on déduit :

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

$$= t\varphi(x) + (1-t)\varphi(y).$$

$\Leftarrow$ Inversement, on suppose $\varphi$ convexe. Alors : $\forall t \in [0,1]$, $\forall x, y \in V$,

$$f(tx+(1-t)y) = \varphi(tx+(1-t)y) + \frac{\alpha}{2}\|tx+(1-t)y\|^2$$

$$\leq t\varphi(x) + (1-t)\varphi(y) + \frac{\alpha}{2}(t^2\|x\|^2 + (1-t)^2\|y\|^2 + 2t(1-t)\langle x, y\rangle)$$

$$= tf(x) + (1-t)f(y) - \frac{\alpha}{2}t(1-t)\|x-y\|^2,$$

i.e. $f$ est $\alpha$-elliptique.

2. Il suffit de montrer que $f : V \to \mathbb{R}$ continue est convexe ssi

$$\forall (x,y) \in V^2, \quad f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

Pour cela on raisonne par récurrence sur $n$ en montrant :

$$(\mathcal{P}_n) : \quad \forall (x,y) \in V^2, \quad \forall k \in \{0, \cdots, 2^n\},$$

$$f\left(\frac{k}{2^n}x + \left(1 - \frac{k}{2^n}\right)y\right) \leq \frac{k}{2^n}f(x) + \left(1 - \frac{k}{2^n}\right)f(y).$$

On voit médiatement que $(\mathcal{P}_0)$ est vraie. On suppose $(\mathcal{P}_n)$ vraie. Soit alors $k \in \{0, \cdots, 2^{n+1}\}$.
Si $2^n < k \leq 2^{n+1}$, alors :

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

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

On en déduit, par hypothèse sur $f$ :

$$f\left(\frac{k}{2^{n+1}}x + \left(1 - \frac{k}{2^{n+1}}\right)y\right) \leq \frac{1}{2}f\left(\frac{k-2^n}{2^n}x + \left(1 - \frac{k-2^n}{2^n}\right)y\right) + \frac{1}{2}f(x)$$

$$\underset{(\mathcal{P}_n)}{\leq} \frac{k-2^n}{2^{n+1}}f(x) + \frac{1}{2}\left(1 - \frac{k-2^n}{2^n}\right)f(y) + \frac{1}{2}f(x) =$$

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

On suppose que $k \in \{0, \cdots 2^n\}$. Alors :

$$\begin{aligned}
f\left(\frac{k}{2^{n+1}}x + \left(1 - \frac{k}{2^{n+1}}\right)y\right) &= f\left(\frac{k}{2^n}\frac{x}{2} + \left(1 - \frac{k}{2^n}\right)\frac{y}{2} + \frac{y}{2}\right) \\
&\leq \frac{1}{2}f\left(\frac{k}{2^n}x + \left(1 - \frac{k}{2^n}\right)y\right) + \frac{1}{2}f(y) \\
&\underset{(\mathcal{P}_n)}{\leq} \frac{k}{2^{n+1}}f(x) + \frac{1}{2}\left(1 - \frac{k}{2^n}\right)f(y) + \frac{1}{2}f(y) \\
&= \frac{k}{2^{n+1}}f(x) + \left(1 - \frac{k}{2^{n+1}}\right)f(y)
\end{aligned}$$

Finalement dans tous les cas $(\mathcal{P}_{n+1})$ est vraie. Par récurrence sur $n \geq 0$, on en déduit que $(\mathcal{P}_n)$ est vraie $\forall n \geq 0$.
Soit $t \in [0,1]$ et soit $x, y \in V$. On a : $\forall n \geq 0$,

$$[0,1] = \cup_{k=0}^{2^n-1}\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right].$$

Il existe donc une suite $(k_n)_{n\geq 0} \in \mathbb{N}^*$ t.q. :

$$\forall n \geq 0, \quad 0 \leq k_n \leq 2^n - 1 \quad \text{et} \quad \frac{k_n}{2^n} \leq t \leq \frac{k_n+1}{2^n} \Rightarrow t = \lim_{n\to+\infty}\frac{k_n}{2^n}.$$

Par continuité de $f$ :

$$f(tx + (1-t)y) = \lim_{n\to+\infty} f\left(\frac{k_n}{2^n}x + \left(1 - \frac{k_n}{2^n}\right)y\right)$$

avec $\forall n \geq 0$,

$$\begin{aligned}
f\left(\frac{k_n}{2^n}x + \left(1 - \frac{k_n}{2^n}\right)y\right) &\leq \frac{k_n}{2^n}f(x) + \left(1 - \frac{k_n}{2^n}\right)f(y) \\
&\underset{n\to+\infty}{\to} tf(x) + (1-t)f(y).
\end{aligned}$$

On en déduit, quand $n \to +\infty$ :

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Donc $f$ est convexe sur $V$.
On suppose maintenant que $f$ est $\alpha$-elliptique. On en déduit le résultat en prenant $t = 1 - t = \frac{1}{2}$ dans la définition.

Inversement, on suppose que :

$$\forall (x,y) \in V^2, \quad f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y) - \frac{\alpha}{8}\|x-y\|^2.$$

On pose :

$$\forall x \in V, \quad \varphi(x) = f(x) - \frac{\alpha}{2}\|x\|^2.$$

Alors : $\forall (x,y) \in V^2$,

$$\varphi\left(\frac{x+y}{2}\right) = f\left(\frac{x+y}{2}\right) - \frac{\alpha}{8}\|x+y\|^2$$

$$\leq \frac{1}{2}f(x) + \frac{1}{2}f(y) - \frac{\alpha}{8}\|x-y\|^2 - \frac{\alpha}{8}\|x+y\|^2 =$$

$$= \frac{1}{2}f(x) + \frac{1}{2}f(y) - \frac{\alpha}{4}(\|x\|^2 + \|y\|^2) = \frac{1}{2}\varphi(x) + \frac{1}{2}\varphi(y).$$

Du résultat préliminaire, on déduit que $\varphi$ est convexe sur $V$, i.e., d'après 1., que $f$ est $\alpha$-elliptique sur $V$.

□

Dans le cas où $f$ est régulière, il existe des caractérisations de l'uniforme convexité qui peuvent être vues comme des conséquences du Théorème 2.2.3.

**Corollaire 2.3.2** (Caractérisation des fonctions uniformément convexes)**.**

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

*Démonstration.* 1. De la Proposition 2.3.1, on déduit que $f$ est $\alpha$-elliptique ssi $\varphi$ est convexe ce qui, d'après le Théorème 2.2.3, se caractérise par les deux conditions équivalentes : $\forall (x,y) \in V^2$,

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

$$\langle \nabla\varphi(y) - \nabla\varphi(x), y-x \rangle \geq 0$$

où
$$\forall x \in V, \quad \nabla\varphi(x) = \nabla f(x) - \alpha x.$$
On conclut en remarquant que : $\forall (x,y) \in V^2$,
$$\varphi(y) \geq \varphi(x) + \langle \nabla\varphi(x), y - x\rangle$$
$$\iff f(y) - \frac{\alpha}{2}\|y\|^2 \geq f(x) - \frac{\alpha}{2}\|x\|^2 + \langle \nabla f(x) - \alpha x, y - x\rangle$$
$$\iff f(y) \geq f(x) + \frac{\alpha}{2}(\|y\|^2 - \|x\|^2 - 2\langle x, y - x\rangle) =$$
$$= f(x) + \frac{\alpha}{2}(\|x\|^2 + \|y\|^2 - 2\langle x, y\rangle) = f(x) + \frac{\alpha}{2}\|x - y\|^2$$
et que de même : $\forall (x,y) \in V^2$,
$$\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle \geq 0 \iff$$
$$\langle \nabla f(y) - \nabla f(x) - \alpha(y-x), y-x\rangle \geq 0 \iff \langle \nabla f(y) - \nabla f(x), y-x\rangle \geq \alpha\|x-y\|^2.$$

2. De même, en utilisant la Proposition 2.3.1 et le Théorème 2.2.3, $f$ $\alpha$-elliptique ssi
$$\forall (x,h) \in V^2, \quad \langle \nabla^2\varphi(x)h, h\rangle \geq 0$$
où $\nabla^2\varphi(x) = \nabla^2 f(x) - \alpha I$, ce qui donne : $\forall (x,h) \in V^2$,
$$\langle \nabla^2\varphi(x)h, h\rangle \geq 0 \iff \langle \nabla^2 f(x)h - \alpha h, h\rangle \geq 0 \iff \langle \nabla^2 f(x)h, h\rangle \geq \alpha\|h\|^2$$

□

**Exemple 12** (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 10 que $f$ est convexe, resp. strictement convexe, ssi $A$ est semi-definie 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 : $A = U^T DU$ où $U \in \mathcal{O}_n(\mathbb{R})$ est orthogonale et où $D$ est diagonale, formée par les valeurs propres $\lambda_1^2 \leq \cdots \leq \lambda_n^2$ de $A$, avec la convention $\lambda_i \geq 0$, $\forall i \in [[1, n]]$. On en déduit : $\forall h \in \mathbb{R}^n$,
$$\langle Ah, h\rangle = \langle U^T DUh, h\rangle = \langle DUh, Uh\rangle = \sum_{i=1}^{n} \lambda_i^2 (Uh)_i^2 \geq \lambda_1^2 \|Uh\|^2$$

$$\underset{U\in\mathcal{O}_n(\mathbb{R})}{=} \lambda_1^2\|h\|^2$$

d'où on déduit que $f$ est $\lambda_1^2$-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., $\lambda_1^2$ est la meilleure constante d'ellipticité de $A$.

*Remarque* 6. Si $f : V \to \mathbb{R}$ est $\alpha$-elliptique et différentiable, alors $f$ est coercive. En effet : soit $(x, y) \in V^2$. D'après le Corollaire 2.3.2 :

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|y - x\|^2 =$$

$$= f(x) + \|x - y\| \left( \left\langle \nabla f(x), \frac{y - x}{\|x - y\|} \right\rangle + \frac{\alpha}{2}\|y - x\| \right)$$

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

$$\|y - x\| = \|y\| \left( 1 - \frac{\langle x, y\rangle}{\|y\|} + o\left(\frac{1}{\|y\|}\right) \right) \underset{\|y\|\to+\infty}{\sim} \|y\|$$

$$\left| \left\langle \nabla f(x), \frac{y - x}{\|y - x\|} \right\rangle \right| \leq \|\nabla f(x)\|$$

donc

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

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

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

**Théorème 2.3.3.** *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 2.3.4.** *Soit $f$ une fonction $\alpha$-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. :(6)

$$\forall (\lambda, x) \in \mathcal{E}(f), \quad \beta\lambda + L(x) > \delta > \beta\lambda_0 + L(x_0). \tag{6}$$

En considérant $\lambda \to +\infty$ dans (6), on voit que $\beta \geq 0$ et en prenant $x = x_0$ dans (6) 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 $L$ : $\forall (x, y) \in K^2$,

$$\frac{\beta}{2}(f(x) + f(y)) + \frac{1}{2}(L(x) + L(y)) - \frac{\alpha\beta}{8}\|x - y\|^2$$

$$\underset{\beta>0}{\geq} \beta f\left(\frac{x+y}{2}\right) + L\left(\frac{x+y}{2}\right) > \delta.$$

Comme $L$ est continue, il en résulte :

$$\frac{\beta}{2}(f(x) + f(y)) + \frac{1}{2}\|L\|\|x\| + \frac{1}{2}L(y) > \delta + \frac{\alpha\beta}{8}\|x - y\|^2$$

$$= \delta + \frac{\alpha\beta}{8}(\|x\|^2 - 2\langle x, y\rangle + \|y\|^2) \geq \delta + \frac{\alpha\beta}{8}(\|x\|^2 - 2\|x\|\|y\| + \|y\|^2)$$

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 $x$. $\square$

Du Lemme 2.3.4, 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 : $\forall n, m \geq 0$,

$$f\left(\frac{x_m + x_n}{2}\right) \leq \frac{1}{2}f(x_m) + \frac{1}{2}f(x_n) - \frac{\alpha}{8}\|x_m - x_n\|^2. \tag{7}$$

On en déduit, dans un premier temps :

$$\inf_K f \leq \liminf_{m,n\to+\infty} f\left(\frac{x_m + x_n}{2}\right) \leq \limsup_{m,n\to+\infty} f\left(\frac{x_m + x_n}{2}\right) \leq$$

$$\leq \lim_{m\to+\infty} \frac{1}{2}f(x_m) + \lim_{n\to+\infty} \frac{1}{2}f(x_n) = \inf_K f$$

i.e. :

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

On obtient, en reportant dans (7) :

$$\inf_K f + \liminf_{m,n\to+\infty} \|x_m - x_n\|^2 \leq \inf_K f + \limsup_{m,n\to+\infty} \|x_m - x_n\|^2 \leq \inf_K f$$

$$\Rightarrow \liminf_{m,n\to+\infty} \|x_m - x_n\|^2 \leq \limsup_{m,n\to+\infty} \|x_m - x_n\|^2 \leq 0$$

donc :

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

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 $x_n \underset{n\to+\infty}{\to} x^* \in K$, 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 $p \to +\infty$ : $\forall n \geq 0$,

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

$$\iff \frac{\alpha}{4}\|x^* - x_n\|^2 \leq f(x_n) - \inf_K f$$

□

*Remarque* 7. Le Théorème 2.3.3 reste vrai si on suppose $f$ semi-continue inférieure. En effet, le même raisoement montre que $x_n \underset{n\to+\infty}{\to} x^* \in K$ avec

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