



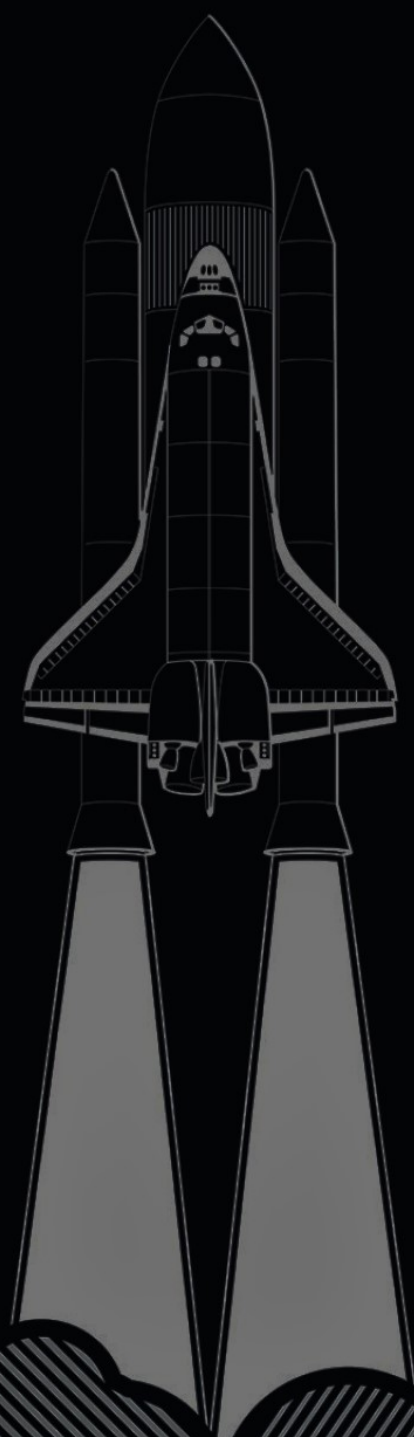
Université de Nice - Faculté des Sciences

OPTIMISATION ET ÉLÉMENTS FINIS

Jean Martin

Prises de notes d'Alexis Pic

2022



Partie I Optimisation



I PREMIERS EXEMPLES DE PROBLÈMES D'OPTIMISATION ET VOCABULAIRE

I.1 Motivations

Les problèmes d'optimisation apparaissent dans de nombreux domaines. Un ingénieur qui doit construire un bâtiment se doit de choisir les meilleurs matériaux au meilleur rapport qualité-prix afin d'avoir une structure aussi sûre que possible le moins cher possible. Un courtier doit choisir les bons investissements qui généreront les meilleures rentes tout en gardant le risque de perte sous contrôle (à son minimum). Un système physique va avoir tendance à se mettre dans un état d'énergie minimale tout en respectant les contraintes physiques qui peuvent lui être données. On pourrait même imaginer un étudiant qui cherche à minimiser le nombre d'heures à passer sur son cours pour avoir une bonne note.

Pour toutes ces situations, on peut trouver des points communs :

- Il y a un *but* ou un *objectif* à atteindre pour l'activité concernée (rendre meilleur), via un *coût* à minimiser.
- On cherche à atteindre cet objectif au regard potentiellement de toute une série de *contraintes* ou d'*attendus* qu'il faut atteindre et/ou satisfaire.
- Implicitement, on suppose qu'il y a un ou des choix possibles de paramètres pour satisfaire aux deux items précédents. Ces paramètres sont appelés *variables d'optimisation* ou de *design*. On devra sélectionner les variables importantes d'un problème d'optimisation à considérer. Ainsi, alors que le temps qu'il fait peut influencer sur la construction d'un bâtiment (ou les révisions d'un cours), ce n'est pas forcément le cas pour le banquier. Le variable « temps qu'il fait » peut donc ou non influencer l'objectif.

En résumé : la formulation d'un problème d'optimisation implique de définir et comprendre correctement le problème posé et de savoir le formaliser en une série de problèmes mathématiques pour espérer le résoudre. Plus précisément, la formalisation d'un problème d'optimisation va mettre en jeu :

- le choix d'une ou plusieurs **variables d'optimisation** vivant dans un espace général à déterminer;
- définir une **fonction objectif** (ou **fonction coût**);
- le cas échéant, identifier l'ensemble des **contraintes**.

Dans la suite, donnons quelques exemples de tels problèmes.

I.2 Premiers exemples simples

I.2.1 Révision des examens

Un étudiant doit réviser pour ses examens. Il a 4 UE à passer et a une semaine pour réviser, ce qui lui donne 42 heures de travail (6 jours et 7 heures de travail par jour). Les variables d'optimisation sont le nombre d'heures de travail par matière (il y en a donc 4) : pour $i \in \{1, \dots, 4\}$, on note x_i le nombre d'heures de révision pour la matière numéro i . L'étudiant fait face à des contraintes : la somme des heures travaillées doit être au plus égale à 42 et il doit passer un nombre d'heures positive sur chaque UE.

L'ensemble des contraintes peut ainsi se formaliser à l'aide d'un ensemble K :

$$K = \left\{ (x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \mid \forall i \in \{1, \dots, 4\}, x_i \geq 0, \sum_{i=1}^4 x_i \leq 42 \right\}.$$

Pour $x \in \mathbf{R}^4$, on note $M(x)$ la moyenne des notes (sur 20) obtenues par l'étudiant après avoir révisé x_i heures sur la matière i . L'objectif est de maximiser M ou encore de minimiser $20 - M$.

Le problème d'optimisation peut donc s'écrire :

$$\text{Trouver } x^* \in K \text{ tel que } M(x^*) = \inf_{x \in K} (20 - M(x)).$$

I.2.2 Consommation des ménages

On considère un ménage qui peut consommer n types de marchandises ($n \in \mathbf{N}^*$) dont les prix forment un vecteur $p \in \mathbf{R}_+^n$. Son revenu à dépenser est un réel $b > 0$, et ses choix de consommation sont supposés être modélisés par une fonction d'utilité u de \mathbf{R}_+^n dans \mathbf{R} (croissante et concave), qui mesure le bénéfice que le ménage tire de la consommation de la quantité $x \in \mathbf{R}^n$ des n marchandises. La consommation du ménage sera le vecteur x^* qui réalisera le maximum

$$\max_{\substack{x \in \mathbf{R}_+^n \\ \langle x, p \rangle \leq b}} u(x),$$

où $\langle x, p \rangle = \sum_{i=1}^n x_i p_i$ (ici $\langle \cdot, \cdot \rangle$ représente le produit scalaire canonique de \mathbf{R}^n). Autrement dit, c'est le vecteur qui maximise l'utilité sous une contrainte de budget maximal.

I.2.3 Problème du toboggan

On se donne deux points A et B (avec le point A plus haut que le point B qui est au sol), et on se demande quelle serait la forme d'un toboggan qui permettrait d'atteindre le point B en partant du point A, le plus rapidement possible. Mathématiquement, on cherche donc à déterminer une courbe du plan reliant A et B, telle qu'un point matériel de poids donné (donc sous l'action de la gravité) glissant le long de cette courbe partant de A arrive en B en un temps minimal. Le

toboggan est donc représenté par le graphe d'une fonction $f : [0; 1] \rightarrow \mathbf{R}$ avec $f(0) = h > 0$ (hauteur du toboggan donnée) et $f(1) = 0$. On prend donc $A = (0, h)$ et $B = (1, 0)$.

En écrivant la conservation de l'énergie, on trouve que, pour une fonction f donnée, le temps de descente $T(f)$ du toboggan est donné par la formule

$$T(f) = \int_0^1 \frac{\sqrt{1 + f'(x)^2}}{2g(h - f(x))} dx,$$

avec g la constante de gravitation.

Le problème revient donc à minimiser T avec f dans un certain espace de fonctions \mathcal{F} :

$$\min_{f \in \mathcal{F}} T(f).$$

I.2.4 Identification de paramètres

On se donne un signal $f : \mathbf{R} \rightarrow \mathbf{R}$ dont on sait qu'il dépend de 3 paramètres $(a, b, c) \in \mathbf{R}^3$ et qui a la forme

$$f(t) = a \cos(bt + c).$$

Pour identifier ces paramètres, on ne dispose que d'un nombre fini $m \in \mathbf{N}^*$ de mesures $(t_i, y_i)_{i \in \{1, \dots, m\}}$. On choisit alors de minimiser une fonctionnelle \mathcal{J} représentant une certaine distance aux mesures :

$$\mathcal{J}(a, b, c) = \sum_{i=1}^m (y_i - f(t_i))^2.$$

C'est ce qui s'appelle **calibrer les paramètres**.

I.3 Mise en forme mathématique générale et éléments de vocabulaire

I.3.1 Forme générale d'un problème d'optimisation

Tous les exemples ci-dessus rentrent dans un cadre assez général sous lequel on peut écrire un problème d'optimisation. On notera :

- V l'espace dans lequel le problème est posé et où vont vivre les variables d'optimisation (c'est-à-dire où seront à chercher les variables d'optimisation). On cherchera V comme un espace vectoriel normé d'une norme que l'on notera $\|\cdot\|$;

- on se donne aussi K qui définira l'ensemble des contraintes qui s'appliqueront sur le système;

- enfin, on se donne $\mathcal{J} : K \subset V \rightarrow \mathbf{R}$ le critère à minimiser que l'on appellera **fonction objectif**, **fonction coût** ou **fonctionnelle**.

Tout ceci sont des *données du problème*. Ensuite, on peut toujours se ramener à un problème de minimisation. En effet, si le problème consiste à maximiser une certaine fonction coût \mathcal{J} , alors étudier le problème de maximisation de \mathcal{J} revient à étudier le problème de minimisation de $-\mathcal{J}$.

Ainsi, dans la suite, on ne considèrera que des problèmes de minimisation pour développer les stratégies de résolution numérique.

Un problème d'optimisation Opt peut alors se mettre sous la forme générale très simple suivante : on cherche à savoir s'il existe un $u^* \in K$ qui réalise le minimum de \mathcal{J} sur K , $\min_{u \in K} \mathcal{J}(u)$.

Cela peut encor s'écrire :

$$\text{Trouver } u^* \in K \text{ tel que } \mathcal{J}(u^*) = \inf_{u \in K} \mathcal{J}(u).$$

L'inconnue du problème est donc u^* à chercher dans l'ensemble K .

Il se pose alors plusieurs questions mathématiques qui se traduisent sur le problème concret de départ :

- Parle-t-on de *minimum* ou d'*infimum* ?
- Y a-t-il existence d'au moins une solution u^* à ce problème ?
- Si une solution existe, est-elle unique ?
- A-t-on une caractérisation du ou des minima, s'ils existent ?
- Peut-on trouver cette (ces) solution(s) minimale(s) quand elle(s) existe(nt) ?
- Est-on capable de donner l'expression de cette solution ?

Suivant les cas, la réponse est plus ou moins simple, affirmative ou négative, voire inconnue. Pour les exemples précédents, on peut identifier les données V , K et J .

I.3.2 Définitions et éléments de vocabulaire

Soit $J : K \subset V \rightarrow \mathbf{R}$; on s'intéresse au problème de minimisation associé à J . On distingue les minima locaux des minima globaux.

Définition 1. On dit que $u \in K$ est un **point de minimum local** de J sur K si

$$\exists \delta > 0, \forall v \in K, \|v - u\| < \delta \implies J(v) \geq J(u).$$

On dit alors que la valeur $J(u)$ est un **minimum local** de J sur K .

Définition 2. On dit que $u \in K$ est un **point de minimum global** de J sur K si

$$\forall v \in K, J(v) \geq J(u).$$

On dit alors que la valeur $J(u)$ est un **minimum global** de J sur K .

Définition 3. On appelle **infimum** de J sur K , et on note $\inf_{u \in K} J(u)$:

- si J est minorée sur K , c'est la borne supérieure dans \mathbf{R} des minorants de J sur K ;
- si J n'est pas minorée sur K , c'est $-\infty$;
- si K est vide, alors c'est $+\infty$ par convention.

Définition 4. On appelle **suite minimisante** de J dans K une suite d'éléments de K telle que

$$\lim_{n \rightarrow +\infty} J(u_n) = \inf_{v \in K} J(v).$$

Par définition même de l'infimum, *une suite minimisante existe toujours*. En effet, commençons par le cas où J est minorée, et donc $\inf_{v \in K} J(v)$ est un réel donné. Soit $n \in \mathbf{N}^*$. Par définition

de l'infimum, $\inf_{v \in K} J(v) + \frac{1}{n}$ n'est plus un minorant de J sur K (puisque $\inf_{v \in K} J(v) + \frac{1}{n} > \inf_{v \in K} J(v)$) et que $\inf_{v \in K} J(v)$ est la borne supérieure des minorants de J sur K). Donc il existe $u_n \in K$ tel que $J(u_n) < \inf_{v \in K} J(v) + \frac{1}{n}$. On a donc pour tout $n \in \mathbf{N}$,

$$\inf_{v \in K} J(v) \leq J(u_n) < \inf_{v \in K} J(v) + \frac{1}{n}.$$

On en déduit le résultat par encadrement en passant à la limite lorsque $n \rightarrow +\infty$: la suite $(u_n)_{n \in \mathbf{N}^*}$ est bien une suite minimisante.

Si \mathcal{J} n'est pas minorée, alors $\inf_{v \in K} \mathcal{J}(u) = -\infty$. On sait que pour tout $m \in \mathbf{R}$, il existe $v \in K$ tel que $\mathcal{J}(v) < m$. Soit alors $n \in \mathbf{N}$. En appliquant ce qui précède à $m = -n$, on en déduit qu'il existe un point $v_n \in K$ tel que $\mathcal{J}(v_n) < -n$. En passant à la limite sur n , on a donc bien le résultat voulu : la suite $(v_n)_{n \in \mathbf{N}}$ est bien une suite minimisante.

1.3.2.1 L'espace K et type de contraintes

Si $K = V$, on parlera de **problème d'optimisation sans contrainte** (ou **problème d'optimisation libre**). Sinon, on parlera de **problème d'optimisation avec contraintes**. Très souvent, l'espace K sera donné par une série d'équations ou d'inéquations de la forme suivante :

- Soient $p \in \mathbf{N}^*$ et $h_i : V \rightarrow \mathbf{R}$ des fonctions données ($1 \leq i \leq p$). Dans le cas de

$$K = \{u \in V \mid \forall i \in \{1, \dots, p\}, h_i(u) = 0\},$$

on parlera de **problème d'optimisation avec contraintes d'égalité**. Il y a alors p contraintes données par les fonctions h_i ($1 \leq i \leq p$).

- Soient $q \in \mathbf{N}^*$ et $g_j : V \rightarrow \mathbf{R}$ des fonctions données ($1 \leq j \leq q$). Dans le cas de

$$K = \{u \in V \mid \forall j \in \{1, \dots, q\}, g_j(u) \leq 0\},$$

on parlera de **problème d'optimisation avec contraintes d'inégalité**. Il y a alors q contraintes données par les fonctions g_j ($1 \leq j \leq q$).

- Soient $p \in \mathbf{N}^*$, $q \in \mathbf{N}^*$, $h_i : V \rightarrow \mathbf{R}$ et $g_j : V \rightarrow \mathbf{R}$ des fonctions données ($1 \leq i \leq p$ et $1 \leq j \leq q$). Dans le cas de

$$K = \{u \in V \mid \forall i \in \{1, \dots, p\}, \forall j \in \{1, \dots, q\}, h_i(u) = 0, g_j(u) \leq 0\},$$

on parlera de **problème mixte**.

Si V est un espace de dimension finie, on parlera de **problème d'optimisation en dimension finie**.

1.3.3 Cas particulier où $V = \mathbf{R}^d$

On se place dans le cadre qui sera le cadre majoritaire du cours : c'est le cas où $V = \mathbf{R}^d$ pour $d \in \mathbf{N}$. On notera alors $\langle \cdot, \cdot \rangle$ le produit scalaire euclidien de \mathbf{R}^d , c'est-à-dire, pour tout $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$, $\langle u, v \rangle = \sum_{i=1}^d u_i v_i$, en notant $u = (u_i)_{1 \leq i \leq d}$ et $v = (v_i)_{1 \leq i \leq d}$.

1.3.3.1 Caractéristiques de \mathcal{J}

Définition 5. Soit $d \in \mathbf{N}^*$. On dit que $\mathcal{J} : \mathbf{R}^d \rightarrow \mathbf{R}$ est une **fonctionnelle quadratique** s'il existe $A \in \mathcal{M}_d(\mathbf{R})$, $b \in \mathbf{R}^d$ et $c \in \mathbf{R}$ tels que, pour tout $x \in \mathbf{R}^d$, on a

$$\mathcal{J}(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle + c.$$

Remarque. On peut tout aussi bien écrire

$$\mathcal{J}(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle + c$$

avec $b \in \mathbf{R}^d$ (on choisit $-b$ à la place de b dans la définition).

Ici, on fait donc une distinction entre *fonctionnelle* et *forme quadratique*. La différence entre les deux se situe dans le terme linéaire $\langle b, x \rangle$ que l'on ajoute à la forme quadratique.

Définition 6. Soit $d \in \mathbf{N}^*$. On dit que $\mathcal{J} : \mathbf{R}^d \rightarrow \mathbf{R}$ est une **forme quadratique** s'il existe $A \in \mathcal{M}_d(\mathbf{R})$, telle que, pour tout $x \in \mathbf{R}^d$, on a

$$\mathcal{J}(x) = \frac{1}{2} \langle Ax, x \rangle.$$

I.3.3.2 Les divers types de problèmes d'optimisation

Donnons dans ce cas précis où $V = \mathbf{R}^d$ les divers types de problèmes d'optimisation que l'on va distinguer suivant la nature des contraintes.

– **Problèmes sans contrainte :**

- problèmes linéaires si \mathcal{J} est *affine*⁽¹⁾;
- problèmes quadratiques si \mathcal{J} est quadratique;
- problèmes non linéaires sinon.

– **Problèmes avec contraintes linéaires (les h_i et g_i affines) :**

- problèmes avec contraintes d'égalité :
 - problèmes linéaires si \mathcal{J} est *affine*;
 - problèmes linéaires quadratiques si \mathcal{J} est quadratique;
 - problèmes non linéaires sinon;
- problèmes avec contraintes d'inégalités et mixtes :
 - programmation linéaire si \mathcal{J} est *affine*;
 - problèmes linéaires quadratiques si \mathcal{J} est quadratique;
 - problèmes non linéaires avec contraintes linéaires sinon.

– **Problèmes de programmation non linéaire.**

Cette classification peut varier légèrement suivant les personnes, et on peut donner une classification analogue dans des cas plus généraux que ceux de ce cours.

(1). Une fonction affine s'écrit $\mathcal{J}(x) = a + f(x)$ où $f : \mathbf{R}^d \rightarrow \mathbf{R}$ est une application linéaire et $a \in \mathbf{R}$.

Dans la première partie de ce cours, on étudiera en détails le cas où $V = \mathbf{R}^d$ ($d \in \mathbf{N}^*$). C'est ce qu'on appelle de l'**optimisation en dimension finie**.*

I.4 Questions d'existence et d'unicité en dimension finie ($V = \mathbf{R}^d$)

On se place dans le cadre d'un problème d'optimisation en dimension finie, et on se demande si l'on sait dire si on a existence et unicité d'un point de minimum (on dit aussi **minimiseur**). On va voir que les réponses à ces questions dépendent grandement du problème considéré, ce qui est après tout normal. On commence par donner des exemples assez simples de situations où la réponse à ces questions n'est pas oui.

I.4.1 Quelques exemples et contre-exemples simples en dimension 1

Soient a et b deux réels tels que $a < b$. On se place alors sur $V = [a; b] \subset \mathbf{R}$.

— On peut tracer une fonction \mathcal{J} continue n'ayant pas de minimum ni de maximum sur \mathbf{R} , par exemple une droite affine sur \mathbf{R} . Par contre, elle possède bien un minimum et un maximum sur tout intervalle borné $[a; b]$ aux extrémités de celui-ci.

Ce qu'on voit avec cet exemple très simple, c'est que *les espaces choisis comptent énormément*. Il y a en effet un gros rôle joué par la topologie suivant si l'espace ouvert, fermé, compact, etc.

— Essayons d'imaginer une fonction qui admet un minimum global, mais que ce minimum n'est pas unique au sens où il y a plusieurs points où le minimum global est atteint. On peut par exemple penser à la fonction sinus sur \mathbf{R} . Il y a une seule valeur de minimum, mais les points de minimum sont localisés en les points de la forme $-\frac{\pi}{2} + 2k\pi$ pour $k \in \mathbf{Z}$.

Ainsi, on peut avoir *existence du point de minimum global, sans avoir d'unicité*.

— On peut avoir le cas d'un infimum non atteint, par exemple avec la fonction $x \mapsto \frac{1}{x}$ sur $[1; +\infty[$.

— Enfin, on peut avoir une infinité de points de minima locaux, mais aucun global.

En bref, les situations peuvent être très variées et compliquées.

I.4.2 Résultat d'existence

Dans certains cas particuliers, on est capable de répondre à la question de l'existence.

I.4.2.1 Sur un espace métrique compact

On a notamment le théorème suivant si on minimise sur un compact une fonction continue. Ce résultat est connu depuis un moment.

Théorème 1. *Toute fonction continue sur un espace métrique compact est bornée et atteint ses bornes.*

Ainsi, lorsqu'on cherche à minimiser une fonction continue sur un compact, on sait répondre affirmativement à la question de l'existence. Dans ce cas là, on sait que l'infimum (ou le supremum) est atteint et que c'est un minimum (maximum); on peut donc trouver au moins un point de minimum (maximum). Autrement dit, si K est compact, et $\mathcal{J} : K \rightarrow \mathbf{R}$ continue, alors \mathcal{J} est bornée et atteint ses bornes. Il existe donc u^* et u_* deux éléments de K tels que $\mathcal{J}(u^*) = \sup_{u \in K} \mathcal{J}(u) = \max_{u \in K} \mathcal{J}(u)$ et $\mathcal{J}(u_*) = \inf_{u \in K} \mathcal{J}(u) = \min_{u \in K} \mathcal{J}(u)$.

I.4.3 Sur un espace non compact

On cherche ici à avoir un résultat un peu plus général lorsque l'on n'est pas sur un compact.

Théorème 2. Soit K un ensemble fermé et non vide de \mathbf{R}^d ($d \in \mathbf{N}^*$) et $\mathcal{J} : K \rightarrow \mathbf{R}$ une fonction continue vérifiant la propriété dite « infinie à l'infini^a » :

$$\forall (u_n) \in K^{\mathbf{N}}, \lim_{n \rightarrow +\infty} \|u_n\| = +\infty \implies \lim_{n \rightarrow +\infty} \mathcal{J}(u_n) = +\infty.$$

Alors il existe au moins un point de minimum de \mathcal{J} sur K . De plus, on peut extraire de toute suite minimisante de \mathcal{J} sur K une sous-suite convergeant vers un point de minimum sur K .

a. On dit alors aussi que \mathcal{J} est **infinie à l'infini**.

Démonstration. Soit $(u_n)_{n \in \mathbf{N}}$ une suite minimisante de \mathcal{J} sur K (on a montré dans la section I.3.2 qu'une telle suite existe). Par définition de cette suite minimisante, on sait que $(\mathcal{J}(u_n))_{n \rightarrow +\infty}$ est une suite convergente, et donc $(\mathcal{J}(u_n))_{n \in \mathbf{N}}$ est une suite bornée.

Comme \mathcal{J} est « infinie à l'infini », on en déduit que $(u_n)_{n \in \mathbf{N}}$ est bornée. Montrons cela par l'absurde. Supposons donc que $(u_n)_{n \in \mathbf{N}}$ n'est pas bornée. Alors^a

$$\forall M \geq 0, \exists n(M) \in \mathbf{N}, \|u_{n(M)}\| > M. \quad (1)$$

Soit $p \in \mathbf{N}$. En appliquant (1) à $M = p$, on en déduit qu'il existe $n(p) \in \mathbf{N}$ tel que $\|u_{n(p)}\| > p$. Posons alors, pour tout $p \in \mathbf{N}$, $v_p = u_{n(p)}$. On a alors, pour tout $p \in \mathbf{N}$, $\|v_p\| \geq p$ et donc $\|v_p\| \rightarrow +\infty$ lorsque $p \rightarrow +\infty$. Comme \mathcal{J} est infinie à l'infini, on sait que $\mathcal{J}(v_p) \rightarrow +\infty$ lorsque $p \rightarrow +\infty$. Mais on sait aussi que $(\mathcal{J}(u_n))_{n \in \mathbf{N}}$ est une suite bornée (cf. le début de la preuve), donc il existe $m > 0$ tel que pour tout $n \in \mathbf{N}$, $|\mathcal{J}(u_n)| \leq m$. Donc en particulier, pour tout $p \in \mathbf{N}$, $|\mathcal{J}(u_{n(p)})| \leq m$. Autrement dit, $(\mathcal{J}(u_{n(p)}))_{p \in \mathbf{N}}$ est bornée, ce qui est en contradiction avec le fait que $\mathcal{J}(v_p) \rightarrow +\infty$.

On a bien montré que $(u_n)_{n \in \mathbf{N}}$ est une suite bornée. Par le théorème de Bolzano-Weierstrass, on sait donc que l'on peut extraire de $(u_n)_{n \in \mathbf{N}}$ une sous-suite qui converge dans \mathbf{R}^d vers un point $u^* \in \mathbf{R}^d$. Mais comme K est fermé, on en déduit que $u^* \in K$.

Enfin, par définition de $(u_n)_{n \in \mathbf{N}}$ (c'est une suite minimisante), on sait donc, en utilisant le fait que \mathcal{J} soit continue, que

$$\mathcal{J}(u^*) = \inf_{u \in K} \mathcal{J}(u).$$

Le point u^* est donc un point de minimum de \mathcal{J} sur K . □

a. On écrit la négation de « $(u_n)_{n \in \mathbf{N}}$ est bornée » : il existe $M > 0$ tel que pour tout $n \in \mathbf{N}$, $\|u_n\| \leq M$.

On peut donner quelques exemples de fonctions qui sont « infinies à l'infini ».

Exemples.

1. $\mathcal{J} : x \in \mathbf{R}^d \mapsto \|x\|^2$ est une fonction infinie à l'infini.

2. Soit $K = \{(x, y) \in \mathbf{R}^2 \mid x + y \leq 4\} \subset \mathbf{R}^2$, et $f : (x, y) \in K \mapsto x^4 + y^4 - x^2 \in \mathbf{R}$. On peut montrer que f est infinie à l'infini. On a, pour $x \in \mathbf{R}$, $(x^2 - 1)^2 \geq 0$, donc

$$x^4 \geq 2x^2 - 1. \quad (2)$$

On en déduit que, pour tout $y \in \mathbf{R}$, $x^4 + y^4 - x^2 \geq x^2 + y^4 - 1$. On peut alors écrire, en appliquant l'inégalité (2) à y cette fois,

$$x^4 + y^4 - x^2 \geq x^2 + 2y^2 - 2.$$

Donc pour tout $x \in \mathbf{R}$, $f(x) \geq \|x\|^2 - 2$. On a donc bien f infinie à l'infini.

Remarque. Attention. Ce théorème n'est plus valable en dimension infinie, puisque les fermés bornés ne sont plus forcément compacts. On fait alors appel à l'analyse Hilbertienne (notions de convergence faible, semi-continuité de \mathcal{J} pour pouvoir conclure), mais ceci est hors en dehors du programme de ce cours.

Pour ce qui est du problème de l'unicité et du caractère global ou local du minimum, on peut avoir plein de situations différentes. La *convexité* est une notion qui aide pour les problèmes d'unicité des minimiseurs, mais aussi sur le caractère global ou local du minimiseur. C'est l'objet de la section suivante.

I.5 Existence et unicité dans le cas convexe

I.5.1 Définitions, rappels

Définition 7. On dit qu'un ensemble C est **convexe** si

$$\forall (u, v) \in C \times C, \forall t \in [0; 1], tu + (1 - t)v \in C.$$

Définition 8. Soit E un \mathbf{R} -espace vectoriel, et soit $K \subset E$ un ensemble convexe. Soit $\mathcal{J} : K \rightarrow \mathbf{R}$.

On dit que \mathcal{J} est une **fonction convexe** si elle vérifie :

$$\forall (x, y) \in K \times K, \forall t \in [0; 1], \mathcal{J}(tx + (1 - t)y) \leq t\mathcal{J}(x) + (1 - t)\mathcal{J}(y).$$

On dit que \mathcal{J} est une **fonction strictement convexe** si elle vérifie :

$$\forall (x, y) \in K \times K, \forall t \in [0; 1], x \neq y \text{ et } \mathcal{J}(tx + (1 - t)y) < t\mathcal{J}(x) + (1 - t)\mathcal{J}(y).$$

Soit $\alpha \in \mathbf{R}$. Si E est normé, on dit que \mathcal{J} est une **fonction α -convexe** si elle vérifie :

$$\forall (x, y) \in K \times K, \forall t \in [0; 1], \mathcal{J}(tx + (1 - t)y) \leq t\mathcal{J}(x) + (1 - t)\mathcal{J}(y) - \frac{\alpha}{2}t(1 - t)\|x - y\|^2.$$

On parle alors de problème d'**optimisation convexe** si \mathcal{J} est une fonction convexe et K est convexe.

On peut faire quelques remarques importantes :

- si \mathcal{J} est une fonction strictement convexe, alors \mathcal{J} est une fonction convexe;
- si \mathcal{J} est une fonction α -convexe avec $\alpha > 0$, alors \mathcal{J} est une fonction strictement convexe;
- si \mathcal{J} est une fonction α -convexe avec $\alpha \in \mathbf{R}_+$, alors \mathcal{J} est convexe.

Les réciproques sont cependant fausses. De plus,

- une fonction 0-convexe est une fonction convexe;
- une somme de fonctions convexes est une fonction convexe;
- la composition de fonctions convexes n'est *pas forcément convexe*. Par contre, si f est une fonction convexe et g est une application linéaire, alors $f \circ g$ est convexe.

On rappelle également le résultat suivant valable en dimension finie qui dit que toute fonction convexe sur un ouvert convexe inclus dans \mathbf{R}^d ($d \in \mathbf{N}^*$) est continue.

Théorème 3. Soit U un ouvert convexe de \mathbf{R}^d ($d \in \mathbf{N}^*$), et $\mathcal{J} : U \rightarrow \mathbf{R}$ une fonction convexe sur U . Alors \mathcal{J} est continue sur U .

I.5.2 Existence d'un minimiseur

Proposition 1. Soit $\mathcal{J} : \mathbf{R}^d \rightarrow \mathbf{R}$ ($d \in \mathbf{N}^*$) une fonction α -convexe avec $\alpha > 0$. Alors \mathcal{J} admet au moins un point de minimum sur \mathbf{R}^d .

Démonstration. En appliquant l'inégalité de la définition d' α -convexité, avec $u = 0_{\mathbf{R}^d}$, on obtient, pour tout $y \in \mathbf{R}^d$ et tout $\delta \in [0; 1]$,

$$\mathcal{J}(\delta y) \leq (1 - \delta)\mathcal{J}(0_{\mathbf{R}^d}) + \delta\mathcal{J}(y) - \frac{\alpha}{2}\delta(1 - \delta)\|y\|^2,$$

ce qui se réécrit

$$\forall y \in \mathbf{R}^d, \forall \delta \in]0; 1], \mathcal{J}(y) \geq \mathcal{J}(0_{\mathbf{R}^d}) + \frac{1}{\delta}(\mathcal{J}(\delta y) - \mathcal{J}(0_{\mathbf{R}^d})) + \frac{\alpha}{2}(1 - \delta)\|y\|^2.$$

On peut alors appliquer cette inégalité aux $y \in \mathbf{R}^d$ tels que $\|y\| \geq 2$, avec $\delta = \frac{1}{\|y\|}$, et on obtient

$$\mathcal{J}(y) \geq \mathcal{J}(0_{\mathbf{R}^d}) + \left[\mathcal{J}\left(\frac{y}{\|y\|}\right) - \mathcal{J}(0_{\mathbf{R}^d}) \right] \|y\| + \frac{\alpha}{2} \left(1 - \frac{1}{\|y\|}\right) \|y\|^2 \geq \mathcal{J}(0_{\mathbf{R}^d}) + k\|y\| + \frac{\alpha}{4}\|y\|^2, \quad (3)$$

où k est un minorant de $u \mapsto \mathcal{J}(u) - \mathcal{J}(0_{\mathbf{R}^d})$ sur la sphère unité (k existe car la sphère unité est compacte et \mathcal{J} est continue sur \mathbf{R}^d , cf. th. 1) et où on a utilisé que $\left(1 - \frac{1}{\|y\|}\right) \geq \frac{1}{2}$.

On déduit alors de (3) est infinie à l'infini. Donc en appliquant le théorème 2 (on a bien les hypothèses avec $\mathbf{K} = \mathbf{R}^d$), le problème d'optimisation associé à cette fonction \mathcal{J} admet au moins une solution. \square

Dans cette démonstration, on a au passage montré la proposition suivante :

Proposition 2. Si $\mathcal{J} : \mathbf{R}^d \rightarrow \mathbf{R}$ ($d \in \mathbf{N}^*$) est une fonction α -convexe avec $\alpha > 0$, alors \mathcal{J} est infinie à l'infini.

I.5.3 Unicité et caractère local ou global du minimum

Théorème 4. Soit $\mathbf{K} \subset \mathbf{R}^d$ ($d \in \mathbf{N}^*$) un ensemble convexe, et $\mathcal{J} : \mathbf{K} \rightarrow \mathbf{R}$ une fonction convexe. On a alors :

- (i) tout minimum local de \mathcal{J} sur \mathbf{K} est un minimum global de \mathcal{J} sur \mathbf{K} ;
- (ii) si \mathcal{J} est strictement convexe, alors le problème a au plus une solution.

Démonstration. Montrons (i). Soit u^* un point de minimum local de \mathcal{J} sur \mathbf{K} . Si $\mathbf{K} = \{u^*\}$, alors (i) est vérifiée (il n'y a qu'une seule valeur de \mathcal{J} sur \mathbf{K} , à savoir $\mathcal{J}(u^*)$). Supposons alors que $\mathbf{K} \neq \{u^*\}$. Soit $v \in \mathbf{K} \setminus \{u^*\}$ un élément quelconque de \mathbf{K} différent de u^* . On a, comme \mathbf{K} est convexe, pour tout $t \in [0; 1]$,

$$tu^* + (1 - t)v \in \mathbf{K}.$$

Comme u^* est un point de minimum local de \mathcal{J} , on sait qu'il existe $\delta > 0$ tel que si $w \in \mathbf{K}$ et $\|w - u^*\| \leq \delta$, alors $\mathcal{J}(u^*) \leq \mathcal{J}(w)$.

Choisissons alors $t^* \in [0; 1]$ tel que $1 > t^* > \max\left(0, 1 - \frac{\delta}{\|u^* - v\|}\right)$ (c'est possible puisque $\|u^* - v\| \neq 0$ et $\max\left(0, 1 - \frac{\delta}{\|u^* - v\|}\right) < 1$), et posons $w^* = t^*u^* + (1 - t^*)v$. On a $t^* \in [0; 1]$ et

donc $w^* \in K$. De plus

$$\|w^* - u^*\| = (1 - t^*)\|u^* - v\| \leq \frac{\delta}{\|u^* - v\|} \|u^* - v\| = \delta,$$

puisque par définition de t^* , on a $1 - t^* < \min\left(1, \frac{\delta}{\|u^* - v\|}\right)$. Donc $\mathcal{J}(w^*) \geq \mathcal{J}(u^*)$. Or, puisque \mathcal{J} est convexe,

$$\mathcal{J}(w^*) \leq t^*\mathcal{J}(u^*) + (1 - t^*)\mathcal{J}(v).$$

En combinant les deux dernière inégalités, on trouve

$$\mathcal{J}(u^*) \leq \mathcal{J}(w^*) \leq t^*\mathcal{J}(u^*) + (1 - t^*)\mathcal{J}(v).$$

Et donc les deux inégalités extrêmes donnent

$$(1 - t^*)\mathcal{J}(u^*) \leq (1 - t^*)\mathcal{J}(v).$$

Finalement, puisque $t^* < 1$, on a montré que

$$\forall v \in K, \mathcal{J}(v) \leq \mathcal{J}(u^*).$$

Cela signifie donc que u^* est bien un point de minimum global.

Montrons (ii) par l'absurde. Supposons donc qu'il existe deux points de minimum global dans K , notés $u_1 \in K$ et $u_2 \in K$ distincts. Alors dans ce cas, puisque pour tout $t \in]0; 1[$, on a $tu_1 + (1 - t)u_2 \in K$, on en déduit par stricte convexité de \mathcal{J} et définition de u_1 que

$$\mathcal{J}(u_1) \leq \mathcal{J}(tu_1 + (1 - t)u_2) < t\mathcal{J}(u_1) + (1 - t)\mathcal{J}(u_2).$$

En combinant les deux inégalités extrêmes, on trouve

$$(1 - t)\mathcal{J}(u_1) < (1 - t)\mathcal{J}(u_2).$$

Donc $\mathcal{J}(u_1) < \mathcal{J}(u_2)$ (car $t < 1$). On a donc une contradiction avec la définition de u_2 . \square

I.6 Quelques rappels de calcul différentiel

Pour pouvoir répondre à une des questions que l'on s'est posées au départ : « Peut-on avoir une caractérisation du point de minimum s'il existe ? », on a besoin de notions de calcul différentiel. On aura des conditions à vérifier sur les différentielles première (gradient) et seconde (Hessienne) de la fonction que l'on cherche à minimiser. On verra ces conditions dans le chapitre suivant. Dans cette section, on fait les rappels nécessaires pour pouvoir établir les résultats de caractérisation.

I.6.1 Cas de la dimension 1

On dispose des formules de Taylor.

Proposition 3 (Formule de Taylor-Lagrange). Soient a et b deux réels tels que $a < b$. Soit $f : [a; b] \rightarrow \mathbf{R}$ une application de classe \mathcal{C}^n ($n \in \mathbf{N}^*$) donnée telle que $f^{(n+1)}$ existe sur $]a; b[$. Alors il existe $c \in]a; b[$ tel que

$$f(b) = f(a) + (b - a)f'(a) + \dots + \frac{(b - a)^n}{n!}f^{(n)}(a) + \frac{(b - a)^{n+1}}{(n + 1)!}f^{(n+1)}(c).$$

Proposition 4 (Formule de Taylor-Young). Soient a et b deux réels tels que $a < b$. Soit $f : [a; b] \rightarrow \mathbf{R}$ une application de classe \mathcal{C}^n ($n \in \mathbf{N}^*$) donnée. Alors, lorsque $h \rightarrow 0$, on a

$$f(a+h) = f(a) + hf'(a) + \cdots + \frac{h^n}{n!} f^{(n)}(a) + o(h^n).$$

1.6.2 Rappels de résultats classiques en dimension supérieure

On peut généraliser tout ce qu'on vient de dire à la dimension supérieure, c'est-à-dire pour $d \in \mathbf{N}^*$. On notera $\langle \cdot, \cdot \rangle$ le produit scalaire euclidien sur \mathbf{R}^d . Ici, on manipule des *matrices* et des *vecteurs*, donc il faut être prudent. De plus, la visualisation en dimension 1 est relativement aisée (on peut la plupart du temps tracer le graphe d'une fonction), alors que cela devient beaucoup plus difficile et moins intuitif en dimension supérieure.

1.6.2.1 Rappel de la notion de différentiabilité et des formules de Taylor à l'ordre 2

On rappelle que l'on a une notion de *différentiabilité* d'une fonction de $U \subset \mathbf{R}^d$ dans \mathbf{R} ($d \in \mathbf{N}^*$) avec U un ouvert de \mathbf{R}^d .

Définition 9. Soit U un ouvert de \mathbf{R}^d et $a \in U$. Une application $F : U \rightarrow \mathbf{R}$ est dite **différentiable en a** s'il existe une application linéaire et continue de \mathbf{R}^d dans \mathbf{R} L telle que, lorsque $h \rightarrow 0$ (dans \mathbf{R}^d),

$$F(a+h) = F(a) + L(h) + o(\|h\|).$$

Si L existe, on l'appelle la **différentielle de F en a** et on la note $DF(a)$.

Si F est différentiable en tout point de U , on dit alors que F est **différentiable sur U** et l'application $DF : U \rightarrow \mathcal{L}_c(\mathbf{R}^d, \mathbf{R})$ est appelée l'**application différentielle de F** . Si DF est continue, on dit que F est **de classe \mathcal{C}^1** (et ainsi de suite pour \mathcal{C}^p pour $p \in \mathbf{N}$).

1.6.2.2 Gradient et Hessienne

Soit U un ouvert de \mathbf{R}^d , $a \in U$ et $F : U \rightarrow \mathbf{R}$ une application différentiable en a . On a aussi une notion de *dérivée partielle* qui est une dérivée directionnelle par rapport à chaque direction canonique $\frac{\partial}{\partial x_i}$ ($1 \leq i \leq d$) (qu'on ne rappellera pas dans ce cours), et on a

$$\forall h \in \mathbf{R}^d, DF(a)(h) = \sum_{i=1}^d \frac{\partial F}{\partial x_i}(a) h_i.$$

Une fonction $F : U \rightarrow \mathbf{R}$ est **de classe \mathcal{C}^p** ($p \in \mathbf{N}$) si toutes ses dérivées partielles jusqu'à l'ordre p existent et sont continues sur U .

Soit $a \in \mathbf{R}^d$. On appelle **gradient de F en a** , qu'on note $\nabla F(a)$, le vecteur de \mathbf{R}^d dont les coordonnées dans la base canonique de \mathbf{R}^d sont données par les dérivées partielles, c'est-à-dire :

$$\nabla F(a) = \begin{pmatrix} \frac{\partial F}{\partial x_1}(a) \\ \vdots \\ \frac{\partial F}{\partial x_d}(a) \end{pmatrix}.$$

En particulier, on a, pour tout $h \in \mathbf{R}^d$,

$$DF(a)(h) = \langle \nabla F(a), h \rangle = {}^t \nabla F(a) h.$$

Si F est de classe \mathcal{C}^2 en $a \in U$, on peut définir la différentielle seconde de F en a , $D^2F(a)$. Dans ce cas, la matrice $A = \left(\frac{\partial^2 F}{\partial x_i \partial x_j}(a) \right)_{\substack{1 \leq i \leq d \\ 1 \leq j \leq d}} \in \mathcal{M}_d(\mathbf{R})$ est appelée la **matrice Hessienne** de F en a , qu'on pourra noter $H_a(F)$. On remarque que puisque

$$\forall (i, j) \in \{1, \dots, d\}^2, \frac{\partial^2 F}{\partial x_i \partial x_j}(a) = \frac{\partial^2 F}{\partial x_j \partial x_i}(a)$$

(par le théorème de Schwarz), la matrice est symétrique.

De plus, pour tout $(h, \ell) \in \mathbf{R}^d \times \mathbf{R}^d$, on a

$$D^2F(a)(h, \ell) = \sum_{i,j=1}^d (H_a(F))_{ij} h_i \ell_j = \langle H_a(F)h, \ell \rangle.$$

Ici aussi, on dispose des formules de Taylor.

Proposition 5. Soit $d \in \mathbf{N}^*$, U un ouvert de \mathbf{R}^d et $F : U \rightarrow \mathbf{R}$ une application de classe \mathcal{C}^2 . Soit $u \in \mathbf{R}^d$ et $h \in \mathbf{R}^d$ tels que $\{u + th : t \in [0; 1]\} \subset U$. Alors on a les formules suivantes.

– **Formule de Taylor-Lagrange à l'ordre 2.** Il existe $\theta \in]0; 1[$ tel que

$$F(u + h) = F(u) + DF(u)(h) + \frac{1}{2} D^2F(u)(h, h).$$

– **Formule de Taylor avec reste intégral.** On a

$$F(u + h) = F(u) + DF(u)(h) + \int_0^1 (1-t) D^2F(u + th)(h, h) dt.$$

– **Formule de Taylor-Young.** Lorsque $h \rightarrow 0$

$$F(u + h) = F(u) + DF(u)(h) + \frac{1}{2} D^2F(u)(h, h) + o(\|h\|^2).$$

1.6.3 Le cas convexe

En utilisant les notions de calcul différentiel, on a accès à des résultats pratiques pour voir si une fonction convexe, strictement convexe ou α -convexe.

Proposition 6. Soit $d \in \mathbf{N}^*$. Soit $\mathcal{J} : \mathbf{R}^d \rightarrow \mathbf{R}$ une application de classe \mathcal{C}^1 et $\alpha \geq 0$. Alors les trois assertions suivantes sont équivalentes :

(i) \mathcal{J} est une fonctionnelle α -convexe;

(ii) pour tout $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$,

$$\mathcal{J}(v) \geq \mathcal{J}(u) + \langle \nabla \mathcal{J}(u), v - u \rangle + \frac{\alpha}{2} \|v - u\|^2;$$

(iii) pour tout $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$,

$$\langle \nabla \mathcal{J}(v) - \nabla \mathcal{J}(u), u - v \rangle \geq \alpha \|v - u\|^2.$$

Démonstration. (i) \implies (ii). Soit $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$. On a, pour tout $t \in [0; 1]$,

$$\mathcal{J}((1-t)u + tv) \leq (1-t)\mathcal{J}(u) + t\mathcal{J}(v) - \frac{\alpha}{2} t(1-t) \|u - v\|^2. \quad (4)$$

Mais de plus, en utilisant la formule de Taylor-Young, lorsque $t \rightarrow 0$, on a

$$\mathcal{J}((1-t)u + tv) = \mathcal{J}(u + t(v-u)) = \mathcal{J}(u) + \underbrace{tD\mathcal{J}(u)(v-u)}_{=\langle \nabla \mathcal{J}(u), v-u \rangle} + \|u-v\|o(t). \quad (5)$$

En rassemblant (4) et (5), on trouve

$$tD\mathcal{J}(u)(v-u) + \|u-v\|o(t) \leq t(\mathcal{J}(v) - \mathcal{J}(u)) - \frac{\alpha}{2}t(1-t)\|u-v\|^2.$$

En choisissant $t \in]0; 1[$, on peut donc écrire, lorsque $t \rightarrow 0^+$,

$$\mathcal{J}(v) \geq \mathcal{J}(u) + D\mathcal{J}(u)(v-u) + \frac{\alpha}{2}(1-t)\|u-v\|^2 + \|u-v\|o(1).$$

En faisant tendre $t \rightarrow 0^+$, on obtient donc

$$\mathcal{J}(v) \geq \mathcal{J}(u) + D\mathcal{J}(u)(v-u) + \frac{\alpha}{2}\|u-v\|^2.$$

(ii) \implies (iii). On échange le rôle de u et v dans (ii), et on trouve une inégalité (ii)'. On additionne ces deux inégalités (ii) et (ii)', et on obtient le résultat.

(iii) \implies (i). Soit $\varphi : t \mapsto \mathcal{J}(u + t(v-u))$. Vu les hypothèses sur \mathcal{J} , φ est continue, et même \mathcal{C}^1 sur \mathbf{R} , et pour tous $t \in \mathbf{R}$ et $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$,

$$\varphi'(t) = D\mathcal{J}(u + t(v-u))(v-u).$$

On a ici utilisé la différentiation de fonctions composées. De plus, par (iii) appliqué à $u + t(v-u)$ et $u + s(v-u)$, on a, pour tout $(t, s) \in \mathbf{R} \times \mathbf{R}$,

$$\varphi'(t) - \varphi'(s) \geq \alpha(t-s)\|v-u\|^2$$

si $t > s$. Soit alors $\theta \in]0; 1[$, en intégrant sur $(t, s) \in [\theta; 1] \times [0; \theta]$, on obtient

$$\int_0^\theta \int_\theta^1 \varphi'(t) dt ds - \int_\theta^1 \int_0^\theta \varphi'(s) ds dt \geq \alpha \left(\int_0^\theta \int_\theta^1 t dt ds - \int_0^\theta \int_\theta^1 s dt ds \right) \|v-u\|^2,$$

ce qui donne

$$\theta \int_\theta^1 \varphi'(t) dt - (1-\theta) \int_0^\theta \varphi'(s) ds \geq \frac{\alpha}{2} \theta(1-\theta) \|v-u\|^2.$$

Autrement dit,

$$\theta \varphi(1) + (1-\theta) \varphi(0) - \varphi(\theta) \geq \frac{\alpha}{2} \theta(1-\theta) \|v-u\|^2.$$

En remplaçant par l'expression de φ en fonction de \mathcal{J} , on trouve exactement l'expression de l' α -convexité. \square

Remarque 1. Ce théorème peut être généralisé au cas où \mathcal{J} est une fonction définie sur un ouvert $\Omega \subset \mathbf{R}^d$ et \mathcal{C}^1 . Dans ce cas là, dans les énoncés des assertions, on remplace le « pour tout $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$ » par « pour tout $(u, v) \in U \times U$ », avec $U \subset \Omega$ un ensemble convexe, et on dit que \mathcal{J} est alors α -convexe sur U .

Proposition 7. Soit $d \in \mathbf{N}^*$ et $\mathcal{J} : \mathbf{R}^d \rightarrow \mathbf{R}$ de classe \mathcal{C}^2 , et $\alpha \geq 0$. Alors \mathcal{J} est α -convexe si et seulement si

$$\forall (u, v) \in \mathbf{R}^d \times \mathbf{R}^d, D^2\mathcal{J}(u)(w, w) \geq \alpha \|w\|^2.$$

Démonstration. Supposons \mathcal{J} α -convexe. Soient $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$. Alors $\varphi : t \mapsto \mathcal{J}(u + t(v - u))$ est de classe \mathcal{C}^2 et on a $\varphi'(t) = D\mathcal{J}(u + t(v - u))(v - u) = \langle \nabla \mathcal{J}(u + t(v - u)), v - u \rangle$ et $\varphi''(t) = D^2\mathcal{J}(u + t(v - u))(v - u, v - u) = \langle H_{u+t(v-u)}(\mathcal{J})(v - u), v - u \rangle$. On dispose des équivalences de proposition précédente. On peut aussi reprendre la preuve de (iii) \implies (i) et utiliser le fait que pour tout $(s, t) \in \mathbf{R} \times \mathbf{R}$, si $t > s$,

$$\varphi'(t) - \varphi'(s) \geq \alpha(t - s)\|v - u\|^2.$$

Choisissons $s \in \mathbf{R}$ et $h \in \mathbf{R}_+^*$ et $t = s + h$; on a alors

$$\varphi'(s + h) - \varphi'(s) \geq \alpha h\|v - u\|^2.$$

Et donc comme φ est de classe \mathcal{C}^2 sur \mathbf{R} , on en déduit, en divisant par $h > 0$ et en faisant tendre h vers 0^+ , que

$$\varphi''(t) \geq \alpha\|v - u\|^2.$$

On a donc, pour tout $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$,

$$D^2\mathcal{J}(u + t(v - u))(v - u, v - u) \geq \alpha\|v - u\|^2.$$

Soit alors $(w, v) \in \mathbf{R}^d \times \mathbf{R}^d$. On choisit $v = w + u$ dans l'inégalité précédente, et on obtient

$$D^2\mathcal{J}(u + tw)(w, w) \geq \alpha\|w\|^2.$$

En prenant $t = 0$, on obtient le résultat voulu.

Supposons maintenant que

$$\forall (u, v) \in \mathbf{R}^d \times \mathbf{R}^d, D^2\mathcal{J}(u)(w, w) \geq \alpha\|w\|^2.$$

Soit $(u, v) \in \mathbf{R}^d \times \mathbf{R}^d$ et $t \in \mathbf{R}$. En appliquant cette inégalité avec $u + t(v - u)$ et $v - u$, on obtient

$$D^2\mathcal{J}(u + t(v - u))(v - u, v - u) \geq \alpha\|v - u\|^2$$

et comme précédemment, on déduit donc que

$$\varphi''(t) \geq \alpha\|u - v\|^2.$$

En intégrant entre 0 et 1, on trouve donc que

$$\varphi'(1) - \varphi'(0) \geq \alpha\|v - u\|^2$$

et donc

$$(D\mathcal{J}(v) - D\mathcal{J}(u))(v - u) \geq \alpha\|v - u\|^2,$$

ce qui veut dire que \mathcal{J} est α -convexe par le point (iii) de la proposition 6. □

Remarques 1.

1. On retrouve bien le fait connu que si la Hessienne est positive, alors \mathcal{J} est convexe (c'est le cas $\alpha = 0$ du théorème).

2. Ce théorème peut être généralisé au cas où \mathcal{J} est une fonction définie sur un ouvert $\Omega \subset \mathbf{R}^d$ et de classe \mathcal{C}^2 . Dans ce cas-là, on remplace la conclusion par

$$\forall (u, v) \in U \times U, D^2\mathcal{J}(u)(w, w) \geq \alpha\|w\|^2,$$

avec $U \subset \Omega$ un ensemble convexe. On dit alors que \mathcal{J} est α -convexe sur U .

I.6.3.1 Conséquences sur le théorème d'existence

On peut aussi montrer la proposition suivante, que l'on admettra.

Proposition 8. *Soit Ω un ouvert de \mathbf{R}^d ($d \in \mathbf{N}^*$) et $U \subset \Omega$ un ensemble convexe non borné. Soit alors $J : \Omega \rightarrow \mathbf{R}$ une fonction α -convexe (avec $\alpha > 0$) sur U et de classe \mathcal{C}^1 sur Ω . Alors J est infinie à l'infini sur U .*