Cours d'Optimisation numérique

Notes rédigées par Térence Bayen

Table des matières

1 Introduction à l'optimisation **4**

2 Calcul différentiel, espaces de Hilbert, convexité **6**
2.1 Un peu de calcul différentiel . . . 6
2.2 Fonctions semi-continue inférieurement . . . 7
2.3 Brefs rappels sur les espaces de Hilbert . . . 8
2.4 Convexité . . . 10
2.4.1 Propriétés générales . . . 10
2.4.2 Fonctions convexes de classe C^1 et C^2 . . . 12
2.4.3 Un peu de sous-différentiabilité . . . 13
2.5 Quelques résultats d'existence et d'unicité . . . 14
2.6 Lemme de Farkas . . . 16

3 Conditions d'optimalité **18**
3.1 Optimisation sans contraintes et condition d'Euler . . . 18
3.2 Théorème de Fritz-John et Karush-Kuhn-Tucker . . . 19
3.3 Critères de qualification . . . 21
3.3.1 Condition de Mangasarian-Fromowitz . . . 21
3.3.2 Qualification et cône de Bouligand . . . 22
3.3.3 Preuve de KKT par le lemme de Farkas . . . 23
3.3.4 Cas des contraintes affines . . . 24
3.3.5 Contraintes convexes . . . 24
3.3.6 Condition de qualification pour les contraintes d'inégalité . . . 25
3.3.7 Cas des contraintes d'égalité . . . 25
3.3.8 Cas général (contraintes d'égalités et d'inégalités) . . . 26
3.4 Résumé des conditions de qualification . . . 27
3.5 Théorème de Karush-Kuhn-Tucker en dimension infinie . . . 27
3.5.1 Cas des contraintes d'égalité . . . 28
3.5.2 Cas des contraintes d'inégalité . . . 28
3.6 Conditions d'optimalité au second ordre . . . 30
3.7 Dualité . . . 32
3.7.1 Généralités . . . 32
3.7.2 Points selles de Lagrangien . . . 32
3.7.3 Problèmes convexes et dualité . . . 34
3.7.4 Exemple de saut de dualité . . . 35
3.7.5 Preuve de la condition nécessaire au second ordre par dualité . . . 36

4 Algorithmes numériques **39**
4.1 Algorithmes de descente de gradient . . . 39
4.1.1 Règle d'Armijo . . . 40
4.1.2 Règle de Wolfe . . . 41
4.1.3 Conditionnement pour la méthode de gradient à pas optimal . . . 42
4.1.4 Gradient projeté . . . 43
4.1.5 Algorithme d'Uzawa . . . 44

4.2 Méthode du gradient conjugué 45
4.3 Algorithme de Newton 48
4.3.1 Méthode de Newton 48
4.3.2 Méthode de quasi-Newton 49

5 Exercices **51**
5.1 Exercices sur la convexité et les espaces de Hilbert 51
5.2 Exercices sur l'optimisation sans contraintes 53
5.3 Démonstration du lemme de Farkas 54
5.4 Exercices autour du théorème de KKT 55
5.5 Conditions d'optimalité en dimension infini 59
5.6 Exercices autour de la dualité 61
5.7 Exercices autour des algorithmes numériques 62
5.8 TP sous matlab 64

6 Sujets d'examen **66**
6.1 Session 2015-2016 66
6.1.1 Partiel (1h) 66
6.1.2 Examen final (2h) 67
6.2 Session 2016-2017 68
6.2.1 Partiel 68
6.3 Devoir Maison 70
6.3.1 Devoir Maison sur l'algorithme de Newton 70

Chapitre 1

Introduction à l'optimisation

Un problème d'optimisation met en évidence des variables d'état (paramètres) et des contraintes. Le but est de minimiser une fonction f dans C:

$$\inf_{x \in C} f(x). \tag{1.0.1}$$

Ici f est une fonction défini sur l'ensemble C (un sous-ensemble d'un espace vectoriel normé) à valeurs réelles. L'ensemble C s'appelle ensemble de contraintes : il s'agit d'une restriction sur les paramètres admissibles.

On ne s'intéressera pas dans ce cours à des problèmes multi-critères:

$$\inf_{x \in C} F(x),$$

où $F(x) := (f_1(x), ..., f_m(x))$ sont m fonctions à valeurs réelles définies sur C. Ce type de problème multi-objectif (i.e. qui consiste à vouloir minimiser simultanément plusieurs objectifs parfois antagonistes) est plus difficile et dépasse le cadre de ce cours.

Les différentes questions relatives à (1.0.1) que l'on est amené à se poser en optimisation sont les suivantes :

- 1. Existence d'une solution.
- 2. Conditions nécessaires d'optimalité (du type "$f'(x) = 0''$).
- 3. Conditions suffisantes d'optimalité.
- 4. Algorithmes numériques pour trouver une solution.
- 5. Analyse qualitative de la solution lors d'une perturbation des paramètres.

On étudiera principalement les points 1,2,3 et 4 (partiellement). Citons quelques exemples de problèmes bien étudiés dans la littérature:

- Problème isopérimétrique: parmi les courbes fermées C du plan et de longueur 1, trouver celle d'aire maximale:

  $$\max_{C\ ;\ L(C)=1} \mathcal{A}(C).$$

  Il s'agit d'un problème géométrique à transformer analytiquement.

- Le problème du voyageur de commerce: on a un graphe avec n points (les villes) et $\frac{n(n-1)}{2}$ arêtes de coût c_{ij} telles que $c_{ij} = c_{ji}$. On pose $x_{ij} = 1$ si on va du noeud i au noeud j et 0 sinon. On cherche à minimiser:

  $$\min_{x \in \mathbb{R}^{n^2}} \sum_{1 \le i,j \le n} c_{ij} x_{ij},$$

  sous les contraintes $\sum_{j=1}^n x_{ij} = \sum_{i=1}^n x_{ij} = 1$ et $\sum_{i \in I} x_{ij} \le |I| - 1$ pour tout sous-ensemble $I \subsetneqq \{1, ..., n\}$. Ce problème modélise comment un voyageur de commerce peut passer dans toutes les villes une et une seule fois sans cyclage. Son but est de minimiser son coût de trajet total possible (i.e. de trouver le meilleur chemin possible). Il s'agit d'un exemple de problème linéaire mais dont la complexité est très élevée (grand nombre de contraintes).

- Problème de la chaînette : il s'agit de trouver la forme prise par une chaînette accrochée à deux points du mur: ceci revient à étudier un problème de calcul des variations:

$$\inf_{\substack{y \in C^1([a,b],\mathbb{R}) \\ y(a)=a_a\ ;\ y(b)=y_b}} \int_a^b y(x)\sqrt{1+y'^2(x)}dx.$$

  En effet, la courbe optimale (dite chaînette) minimise son énergie potentielle de pesanteur. Ainsi, ce problème exprime la minimisation de cette énergie parmi toutes les courbes possible.

- Le problème de Newton consiste à trouver parmi tous les objets convexes de hauteur donnée et de base donnée, celui qui tombe le plus rapidement, c.a.d., celui qui en tombant minimise la résistance de l'air. Sous certaines hypothèses, ce problème se ramène à un problème de calcul des variations du même type que le précédent (mais c'est un problème très difficile à résoudre).

- Parmi les solutions d'un système linéaire $Ax = b$ avec $A \in M_{mn}(\mathbb{R})$ (avec $m < n$), on cherche celle qui minimise la semi-norme $\|x\|_0$ qui mesure le nombre de composantes non-nulles:

$$\min_{x\ Ax=b} \|x\|_0.$$

  Trouver une solution au système ayant le plus de composantes nulles est très utile dans le domaine du traitement du signal. La difficulté est que la fonction à minimiser n'est pas régulière.

Le cours suivra les grandes lignes suivantes:

- Existence de solutions en optimisation (propriétés de convexité convexité)
- Conditions nécessaires et suffisantes d'optimalité, qualification (dimension finie et infinie)
- Théorie de la dualité
- Méthodes numériques (algorithmes de descente et Newton).
- Optimisation en dimension infinie (conditions d'optimalité) et calcul des variations.

Il serait possible d'étudier d'autres aspects de l'optimisation comme par exemple la programmation dynamique (en horizon fini ou en horizon infini) et le contrôle optimal. Certains aspects seront traités plus en détails en M2 (comme le contrôle optimal qui constitue la continuation naturelle du calcul des variations).

Références: Pour certaines sections, le cours s'est inspiré entre autres des polycopiés mentionnés ci-dessous. Merci de me signaler tout oubli. Une liste plus exhaustive des diverses références (livres) ayant été utilisés pour ce polycopié est donnée à la fin du document.

- Cours de G. Carlier (calcul différentiel) : `https://www.ceremade.dauphine.fr/~carlier/calculdiff.pdf`
- Cours de G. Carlier (optimisation) : `https://www.ceremade.dauphine.fr/~carlier/progdyn.pdf`
- Cours de S. Delabrière : `https://webusers.imj-prg.fr/~sylvie.delabriere/AnaConv/ACPoly.pdf`
- Cours de P. Cardaliaguet : `https://www.ceremade.dauphine.fr/~cardalia/OptiPgrDyn15-16.pdf`
- Cours de Y. Privat : `https://www.ljll.math.upmc.fr/~privat/documents/optimAgreg.pdf`
- Cours de A. Rondepierre et P. Weiss : `http://www.math.univ-toulouse.fr/~rondep/CoursTD/polyGMM4.pdf`
- Cours, exercices, sujet d'examen par E. Trélat : `https://www.ljll.math.upmc.fr/~trelat/`

Chapitre 2

Calcul différentiel, espaces de Hilbert, convexité

On commence par quelques rappels de calcul différentiel qui seront utiles notamment pour l'optimisation en dimension infinie.

2.1 Un peu de calcul différentiel

On introduit le calcul en dimension quelconque qui sert pour étudier des fonctionnelles ou bien pour le calcul des variations: par exemple, on s'intéressera à la régularité de la fonctionnelle

$$J(x) := \int_0^1 \ell(t, x(t), x'(t))dt$$

où $l : [0,1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ et $x(\cdot) \in X$ où X est l'espace de Banach $X := C^1([0,1], \mathbb{R}^n)$ muni de la norme $\|x\| := \max_{t\in[0,1]}(|x(t)| + |x'(t)|)$. On s'intéressera en particulier à la minimisation de J sur X et à caractériser un minimum.

Définition 2.1. *Soit X et Y deux espaces de Banach. La dérivée directionnelle de f au point $x \in X$ dans la direction $h \in X$ est si elle existe la limite*

$$f'(x, h) := \lim_{t\downarrow 0} \frac{f(x+th) - f(x)}{t}.$$

Remarque 2.1. *(i) Supposons que f admette une dérivée directionnelle en x dans la direction h. Alors on a $f'(x, th) = tf'(x, h)$ pour $t > 0$ i.e. la dérivée directionnelle est positivement homogène.*
(ii) On a aussi le développement suivant: $f(x + th) = f(x) + tf'(x, h) + o(t)$, $t > 0$.

Par exemple $f(x) = |x|$ admet en $x = 0$ une dérivée directionnelle dans toute direction: $f'(0, h) = |h|$, $h \in \mathbb{R}$.

Définition 2.2. *(i) La fonction $f : X \to Y$ est dite Gâteaux-différentiable en x si la dérivée directionnelle de f au point x existe pour tout $h \in X$ et si $h \longmapsto f'(x)h$ est linéaire continue de X dans Y. On note alors $Df(x) \in \mathcal{L}(X, Y)$ sa dérivée (au sens de Gâteaux) définie par:*

$$Df(x)h = f'(x, h).$$

(ii) On dit que f est Fréchet différentiable en x si elle est Gâteaux différentiable en x et si

$$f(x + h) = f(x) + Df(x)h + o(\|h\|_X).$$

On appelle alors $Df(x)$ la dérivée au sens de Fréchet de f en x (appelé également la différentielle de f en x). La différentielle en x est unique.

Remarque 2.2. *Notons que si f est Fréchet différentiable en x alors f est Gâteaux différentiable en x avec la même dérivée.*

Propriété 2.1. *Si f est Fréchet différentiable au point x alors f est continue au point x.*

La définition de “Gâteaux-différentiable” ou “Fréchet-différentiable” dépend du choix de la norme sur les espaces considérés. Cependant, on peut montrer que pour deux normes équivalentes, on est conduit à la même définition.

Remarque 2.3. *(i) Soit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x_1, x_2) := 1$ si $x_1 > 0$ et $x_2 = x_1^2$ et $f(x_1, x_2) = 0$ sinon. Alors f n'est pas continue en $(0, 0)$. De plus $f'((0, 0), h) = 0$ pour tout $h \in \mathbb{R}^2$, donc f est Gâteaux différentiable en $(0, 0)$. Comme f n'est pas continue en $(0, 0)$ alors f n'est pas Fréchet différentiable en $(0, 0)$.*
(ii) Une fonction peut admettre des dérivées directionnelles en un point sans être Gateaux différentiable en ce point. Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x_1, x_2) = \frac{x_1^2}{x_2^2+|x_1|}$ si $(x_1, x_2) \neq (0, 0)$ et $f(0, 0) = 0$. Alors, on peut montrer que $f'((0, 0), h)$ existe pour tout $h \in \mathbb{R}^2$ mais $h \mapsto f'(0, 0)h$ n'est pas linéaire (en effet: $f'((0, 0), h) = 0$ si $h = (0, h_2)$, $h_2 \in \mathbb{R}$ et $f'((0, 0), h) = |h_1|$ si $h = (h_1, h_2)$ avec $h_1 \neq 0$).

Définition 2.3. *On dit que $f : X \to \mathbb{R}$ est de classe C^1 sur X si f est Fréchet différentiable en tout point de X et si $x \longmapsto Df(x)$ est continue de X dans X^*, le dual topologique de X.*

Dans la définition précédente, $Df(x)$ est donc une application continue de X à valeur dans l'espace des formes linéaires continues sur X.

Exercice 2.1. *1) Montrer que $A \in M_n(\mathbb{R}) \mapsto A^2$ est Gâteaux-différentiable en tout point $A \in M_n(\mathbb{R})$.*
2) Soit $f : M_n(\mathbb{R}) \to \mathbb{R}$ définie par $f(A) := e^{Tr(A)}A$. Montrer que f est Gateaux-différentiable, Fréchet différentiable et de classe C^1.

Soit X est un espace de Banach. On note X^* l'ensemble des formes linéaires continue de X dans $\mathbb{R}$ (dual topologique). On rappelle que X^* est muni de la norme:

$$\|L\|_{X^*} := \sup_{\|x\|\leq 1} |L(x)|.$$

2.2 Fonctions semi-continue inférieurement

Les résultats d'existence de solutions à un problème d'optimisation s'énoncent dans le cas où la fonction objectif f est continue, mais on peut également considérer le cas plus général où f est semi-continue inférieurement.

Soit X un espace de Banach et $K \subset X$ un sous-ensemble non vide.

Définition 2.4. *1) On dit que $f : X \to \mathbb{R}$ est semi-continue inférieurement en x_0 si et seulement si:*

$$\forall \varepsilon > 0 \; \exists r > 0 \; \forall x \in H, \; \|x - x_0\| \leq r \Rightarrow f(x) \geq f(x_0) - \varepsilon$$

2) On dit que f est semi-continue inférieurement (s.c.i.) sur H si et seulement si elle est semi-continue inférieurement en tout point de X.

Propriété 2.2. *Une fonction fonction $f : X \to \mathbb{R}$ est s.c.i. en x_0 ssi pour tout suite $(x_n) \in H^{\mathbb{N}}$ telle que $x_n \to x_0$, alors:*

$$f(x_0) \leq \liminf x_n.$$

Preuve. Supposons f s.c.i. en x_0. Soit φ l'extraction correspondant à $\liminf x_n$. Supposons par l'absurde que $\lim x_{\varphi(n)} < f(x_0)$. Donc il existe $\varepsilon > 0$ tel que $\lim x_{\varphi(n)} \leq f(x_0) - \varepsilon$. Comme f est s.c.i. on applique la définition avec $\varepsilon/2$ et donc il existe $r > 0$ tel que si n est assez grand, alors $f(x_{\varphi(n)}) \geq f(x_0) - \varepsilon/2$. D'où une contradiction.

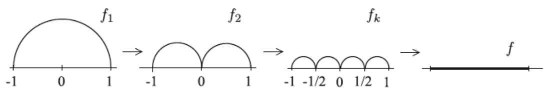
Supposons maintenant que f ne soit pas s.c.i. en x_0. Alors il existe $\varepsilon > 0$ tel que pour tout $r > 0$ il existe $x \in H$, tel que $\|x - x_0\| \leq r$ avec $f(x) < f(x_0) - \varepsilon$. On prend $r = 1/n$. Donc il existe $x_n \in H$ tel que $\|x_0 - x_n\| \leq 1/n$ et $f(x_n) < f(x_0) - \varepsilon$. Ainsi $\liminf f(x_n) < f(x_0) - \varepsilon$. □

En d'autres termes, la fonction $f : H \to \mathbb{R}$ est s.c.i. si pour tout $\lambda \in \mathbb{R}$, le sous-ensemble de niveau:

$$\{x \in H \; ; \; f(x) \leq \lambda\},$$

est fermé dans H.

Exemple: la longueur d'une courbe décrite par une fonction f, $f \mapsto L(f)$ est semi-continue inférieurement. Les



courbes décrites par les fonctions f_k ont toutes la même longueur π et donc le segment f de longueur 1 est bien tel que:

$$L(f) < \liminf_{k \to +\infty} L(f_k).$$

Un certain nombre de résultats standards qui suivent concernant l'existence d'une solution à un problème d'optimisation sont écrits dans le cas où la fonction f à minimiser est continue, mais ils peuvent être étendus au cas f s.c.i.

2.3 Brefs rappels sur les espaces de Hilbert

Dans la suite, les espaces vectoriels considérés sont définis sur $\mathbb{R}$.

Définition 2.5. *Soit H un espace vectoriel. On appelle produit scalaire sur H une forme bilinéaire symétrique $\langle \cdot, \cdot \rangle$ de $H \times H$ dans $\mathbb{R}$, $(u, v) \mapsto \langle u, v \rangle$ et définie positive:*

$$\langle u, u \rangle \geq 0 \; \forall u \in H \quad ; \quad \langle u, u \rangle = 0 \iff u = 0.$$

Un produit scalaire définit sur H une structure d'espace vectoriel normé pour la norme $u \mapsto \langle u, u \rangle^{\frac{1}{2}}$.

Définition 2.6. *On dit que X est un espace de Hilbert ssi X est un espace vectoriel normé muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ qui rend complet l'espace X pour la norme associée.*

De tels espaces constituent les exemples les plus simples d'espace vectoriels de dimension infinie et sont fondamentaux dans le domaine de l'analyse (équations différentielles, équations aux dérivées partielles). Citons les exemples suivants:

- $\mathbb{R}^n$ muni du produit scalaire usuel.
- l^2 l'espace des suites réelles telles que $\sum_{n=0}^{+\infty} |x_n|^2 < +\infty$ muni du produit scalaire $\langle x, y \rangle = \sum_{n=0}^{+\infty} x_n y_n$.
- L'espace de Lebesgue $L^2((0,T)\,;\,\mathbb{R})$ des applications mesurables $f : (0,T) \to \mathbb{R}$ telles que $\int_0^T f^2(t)dt < +\infty$, muni du produit scalaire $\langle f, g \rangle = \int_0^T f(t)g(t)dt$, (voir [9]). Mais $C^0([0,T]\,;\,\mathbb{R})$ muni de la même structure est seulement préhilbertien (considérer l'application affine définie sur $[0,1]$ par $f_n(t) = 1$ sur $[0, 1/2]$ et $f_n(t) = 0$ sur $[1/2 + 1/n, 1]$).
- Soit Ω un ouvert de $\mathbb{R}^n$. L'espace $L^2(\Omega) := \{v : \Omega \to \mathbb{R}\,;\, v \text{ mes. } \int_\Omega v^2 < +\infty\}$ est un espace de Hilbert pour la norme issue du produit scalaire $\langle u, v \rangle = \int_\Omega uv$ où $u, v \in L^2(\Omega)$, (voir [9] p.58).
- Soit Ω un ouvert de $\mathbb{R}^n$. L'espace $H^1(\Omega) := \{v : \Omega \to \mathbb{R}\,;\, v \in L^2(\Omega) \text{ et } \nabla v \in L^2(\Omega)\}$[1] est un espace de Hilbert pour la norme issue du produit scalaire $\langle u, v \rangle = \int_\Omega (\nabla u \cdot \nabla v + uv)$ où $u, v \in H^1(\Omega)$.
- L'espace $H_0^1(\Omega) := \{v \in H^1(\Omega)\,;\, v = 0 \text{ sur } \partial\Omega\}$ est un espace de Hilbert pour la norme $H^1(\Omega)$.

Propriété 2.3. *Soit X un espace de Hilbert. Alors on a:*
- L'inégalité de Cauchy-Schwarz: $\forall x, y \in X, \quad |\langle x, y \rangle| \leq \|x\|\|y\|$.
- L'égalité de la médiane: $\forall x, y, a, b \in X, \quad \|x - a\|^2 + \|x - b\|^2 = 2\|x - \frac{a+b}{2}\|^2 + 2\|\frac{a-b}{2}\|^2$.
- L'égalité du parallélogramme : $\forall x, y \in X$, $\|x - y\|^2 + \|x + y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

Exercice 2.2. *Pour $u \in L^2([0,1], \mathbb{R})$, on note $\|u\|_2 := (\int_0^1 u^2)^{\frac{1}{2}}$. Montrer l'inégalité de Poincaré en dimension 1:*

$$u \in H_0^1([0,1], \mathbb{R}) \;\Rightarrow\; \|u\|_2^2 \leq \|u'\|_2^2.$$

En déduire que $H_0^1([0,1], \mathbb{R})$ peut-être muni du produit scalaire $\langle u, v \rangle := \int_0^1 u'v'$.

[1] On définit l'espace de Sobolev $H^1(\Omega)$ comme l'ensemble des fonctions $u \in L^2(\Omega)$ telles qu'il existe $v = (v_1, \ldots, v_n) \in L^2(\Omega)^n$ vérifiant $\int_\Omega u \frac{\partial \varphi}{\partial x_i} = -\int_\Omega \varphi v_i$ pour toute fonction $\varphi \in C_c^1(\Omega)$, $1 \leq i \leq n$. On notera alors $\nabla u := v$.

On rappelle le théorème de projection sur un ensemble convexe fermé.

Théorème 2.1. *Soit C un ensemble convexe fermé non-vide d'un espace de Hilbert H. Alors pour tout $x \in H$ il existe un unique $p(x) \in H$ tel que: $\|x - p(x)\| = \inf\{\|x - y\| \ ; \ y \in C\}$. Le point $p(x)$ est caractérisé par:*

$$\forall y \in C \ \ \langle x - p(x), y - p(x)\rangle \leq 0.$$

De plus on a:

$$\forall (x, y) \in H \times H, \quad \langle x - y, p(x) - p(y)\rangle \geq 0 \quad \text{et} \quad \|p(x) - p(y)\| \leq \|x - y\|.$$

On rappelle le théorème de séparation suivant.

Théorème 2.2. *Soit H un espace de Hilbert, C un convexe fermé non-vide, et $x_0 \in H$ tel que $x_0 \notin C$. Alors il existe $p \in H$ et $\varepsilon > 0$ tel que:*

$$\forall y \in C, \ \langle p, x_0\rangle \leq \langle p, y\rangle - \varepsilon. \tag{2.3.1}$$

Preuve. On définit un ensemble convexe fermé K par: $K := C - x_0 = \{y - x_0 \ ; \ y \in C\}$. On a $0 \notin K$. Soit p la projection de 0 sur K. On a $p \neq 0$ car $0 \notin K$. Le point p est caractérisé par:

$$\langle -p, z - p\rangle \leq 0 \ \forall z \in K.$$

D'où pour tout $z \in K$, $\langle p, z\rangle \geq \|p\|^2 > 0$ et le résultat en remplaçant z par $y - c_0$, $y \in C$. □

On rappelle le théorème de représentation de Riesz permettant d'identifier H à son dual topologique.

Théorème 2.3. *Soit X' le dual de X. Alors pour tout $L \in X'$ il existe un unique $y \in X$ tel que*

$$\forall x \in X \quad \langle L, x\rangle_{X',X} = \langle y, x\rangle \,.$$

Preuve. Soit $F = \text{Ker}(L)$ et $x_0 \in H$ tel que $L(x_0) = 1$. On a donc $F \oplus \mathbb{R}x_0 = H$ (utiliser le théorème de projection pour vérifier cette égalité[2]). Comme $x_0 \notin \text{Ker}(L)$ on a $y_0 := x_0 - P_{\text{Ker}(L)}(x_0) \neq 0$. Posons $x = \frac{x_0 - y_0}{\|x_0 - y_0\|}^2 \in F^{\perp}$. Vérifions que $L(u) = \langle x, u\rangle$ pour tout $u \in H$. Si $u \in F$, alors on a bien $\langle x, u\rangle = 0$. Supposons maintenant $u \in \mathbb{R}x_0$ de sorte que $u = \lambda x_0$, $\lambda \in \mathbb{R}$. On a $L(u) = \lambda L(x_0) = \lambda$ et $\langle x, u\rangle = \left\langle \lambda x_0, \frac{x_0 - y_0}{\|x_0 - y_0\|}\right\rangle = \lambda$, d'où le résultat. □

Tout espace de Hilbert peut donc être identifié avec son dual. On définit maintenant une nouvelle topologie sur H dite topologie faible qui est la topologie la moins fine rendant continue les formes linéaires $x \mapsto \langle f, x\rangle$, $f \in H$.

Définition 2.7. *On dit qu'une suite (x_n) de H converge faiblement vers x si et seulement si pour tout $y \in H$, on a $\langle x_n, y\rangle \to \langle x, y\rangle$ lorsque $n \to +\infty$.*

On notera qu'en dimension finie, toute suite (x_n) de H converge faiblement vers un point x si et seulement si (x_n) converge fortement vers x.

Les propriétés suivantes découlent de résultats classiques en analyse fonctionnelle comme le théorème de Banach-Steinhaus (voir [9]).

Propriété 2.4. *Soit H un espace de Hilbert.*

- *(i) Soit (x_n) une suite de H qui converge faiblement vers $x \in H$. Alors on a la propriété de semi-continuité:*

$$\|x\|_H \leq \liminf_{n \to +\infty} \|x_n\|_H.$$

- *(ii) Si une suite (x_n) de H est bornée, alors il existe une sous-suite (x_{n_k}) qui converge faiblement.*
- *(iii) Toute suite de H qui converge faiblement est bornée.*

Propriété 2.5. *Si (x_n) converge faiblement vers x et $\|x_n\| \to \|x\|$ lorsque $n \to +\infty$, alors (x_n) converge fortement vers x.*

[2] Si $F \subset H$ est un sous-espace vectoriel fermé, alors $F \oplus F^{\perp} = H$. On utilise le théorème de projection et le fait que F est un espace vectoriel.

Remarque 2.4. *La preuve de (i) est très simple en développant $\|x_n - x\|^2$. La preuve de (ii) et (iii) utilise des résultats d'analyse fonctionnelle (voir [9]) comme le théorème de Banach-Steinhaus. Néanmoins, on peut démontrer (iii) par l'absurde (voir exercice suivant).*

Exercice 2.3. *On souhaite montrer que toute suite (x_n) de H qui converge faiblement est bornée.*
1) Montrer que pour tout $\xi \in H$, $M(\xi) < +\infty$ où $M(\xi) := \sup_n \langle \xi, x_n \rangle$. On raisonne par l'absurde dans toute la suite.
2) Montrer qu'il existe n_1 et $e_1 \in H$ unitaire tel que $\langle e_1, x_{n_1} \rangle \geq 1$. Soit V_1 l'orthogonal de $Vect(e_1, x_{n_1})$.
3) Montrer que la suite $P_{V_1}(x_n)$ n'est pas bornée (raisonner par l'absurde en supposant qu'elle bornée).
4) Montrer qu'il existe $e_2 \in H$ unitaire et x_{n_2} tel que $|\langle e_2, x_{n_2} \rangle| \geq 2(2 + M(e_1))$.
5) Etablir par récurrence qu'il existe une famille de vecteurs unitaires $(e_k)_{k\geq 0}$ et une suite extraite $(x_{n_k})_{k\geq 0}$ vérifiant:

$$e_n \perp Vect(e_1, ..., e_{n-1}, x_{n_1}, ..., x_{n_{k-1}}) \quad \text{et} \quad |\langle e_k, x_{n_k} \rangle| \geq k(k + M(e_1) + \frac{1}{2}M(e_2) + ... + \frac{1}{k-1}M(e_{k-1})).$$

6) Soit $f := \sum_K \frac{e_k}{k}$. Calculer $\|f\|$ et montrer que pour tout k $\langle f, x_{n_k} \rangle \geq k$. Conclure.

Remarque 2.5. *A l'aide du lemme de Riemann-Lebesgue, montrer que $u_n(t) := \cos(nt)$ converge faiblement mais pas fortement vers la fonction nulle dans $L^2([0, 2\pi]; \mathbb{R})$ (phénomène d'oscillation). On peut mentionner aussi le phénomène de concentration et d'évanescence (à l'infini).*

La propriété suivante est proposée en exercice (voir fiche d'exercice).

Propriété 2.6. *Dans un espace de Hilbert, toute suite décroissante de convexes fermés bornés non-vide est d'intersection non-vide.*

Preuve. Voir preuve en exercice. □

Pour la théorie Hilbertienne (base Hilbertienne...) et le théorème de Lax-Milgram on se reporte à [9]

2.4 Convexité

2.4.1 Propriétés générales

Soit X un espace de Hilbert et $f : H \to \mathbb{R}$.

Définition 2.8. *(i) On dit que f est convexe si*

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

(ii) On dit que f est strictement convexe si

$$\forall t \in (0, 1) \; \forall x \in X \; \forall y \in X \; x \neq y \; \Rightarrow \; f(tx + (1 - t)y) < tf(x) + (1 - t)f(y).$$

(iii) On dit que f est fortement convexe de paramètre α si il existe $\alpha > 0$ tel que:

$$\forall t \in [0, 1] \; \forall x \in X \; \forall y \in X \; f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - \alpha\frac{t(1 - t)}{2}\|x - y\|^2.$$

Remarque 2.6. *La convexité forte implique la convexité stricte qui implique la convexité.*

Plus généralement on est amené à définir des fonctions convexes à valeurs dans $\mathbb{R} \cup +\{\infty\}$.

Définition 2.9. *Soit $f : H \to \mathbb{R} \cup +\{\infty\}$ une fonction convexe. On appelle domaine de f l'ensemble défini par: $dom(f) := \{x \in H \; ; \; f(x) < +\infty\}$.*

A titre d'exemple, la fonction indicatrice d'un ensemble convexe fermé $K \subset H$ non vide et définie par:

$$\mathbb{1}_K(x) := \begin{cases} 0 & \text{si } x \in K \\ +\infty & \text{si } x \notin K \end{cases} \tag{2.4.1}$$

est convexe. Pour ce qui est de la continuité des fonctions convexes, on commence par le lemme suivant.

Lemme 2.1. *Soit H un espace de Hilbert et $f : H \to]-\infty, +\infty]$ une fonction convexe. Soit $x_0 \in Dom(f)$. Soit $\rho > 0$. On suppose que $\eta := \sup f(B(x_0, \rho)) < +\infty$. Alors:*

$$\forall \alpha \in [0,1] \; \forall x \in B(x_0, \alpha\rho), \; |f(x) - f(x_0)| \leq \alpha(\eta - f(x_0)). \tag{2.4.2}$$

Preuve. Soit $x \in B(x_0, \alpha\rho)$. En posant $x = (1-\alpha)x_0 + \alpha \frac{x-(1-\alpha)x_0}{\alpha}$ on a par convexité:

$$f(x) - f(x_0) \leq (1-\alpha)f(x_0) + \alpha f\left(\frac{x-(1-\alpha)x_0}{\alpha}\right) - f(x_0) = \alpha\left[f\left(x_0 + \frac{x-x_0}{\alpha}\right) - f(x_0)\right] \leq \alpha(\eta - f(x_0)).$$

De même en posant $x_0 = \frac{x}{1+\alpha} + \frac{\alpha}{1+\alpha}\frac{(1+\alpha)x_0 - x}{\alpha}$ on obtient par convexité:

$$f(x_0) - f(x) = f\left(\frac{x}{1+\alpha} + \frac{\alpha}{1+\alpha}\frac{(1+\alpha)x_0 - x}{\alpha}\right) - f(x) \leq \frac{\alpha}{1+\alpha}\left[f\left(x_0 + \frac{x_0 - x}{\alpha}\right) - f(x)\right] \leq \frac{\alpha}{1+\alpha}(\eta - f(x)).$$

D'où l'on obtient $f(x_0) - f(x) \leq \frac{\alpha}{1+\alpha}(\eta - f(x_0) + f(x_0) - f(x))$ qui entraîne

$$f(x_0) - f(x) \leq \alpha(\eta - f(x_0)),$$

d'où le résultat. □

Proposition 2.1. *Soit H un espace de Hilbert et $f : H \to]-\infty, +\infty]$ une fonction convexe. Soit $x_0 \in Dom(f)$. On suppose que f est bornée au voisinage de x_0. Alors f est continue en x_0.*

Preuve. Il existe $\rho > 0$ tel que $\eta := \sup f(B(x_0, \rho)) < +\infty$. Pour $\alpha \in]0,1[$ et $x \in B(x_0, \alpha\rho)$ on a $|f(x) - f(x_0)| \leq \alpha(\eta - f(x_0))$ d'où le résultat. □

Remarque 2.7. *On peut montrer l'équivalence entre la continuité de f en x_0 et le fait que f soit bornée au voisinage de x_0 (voir par exemple [3] pour plus de détails sur la question).*

Exercice 2.4. *Redémontrer la proposition 2.1 en dimension* 1 *(i.e. pour une fonction convexe $f : I \to \mathbb{R}$ en utilisant la propriété des pentes croissantes:*

$$(x,y,z) \in I^3, \; x < y < z \Rightarrow \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y}.$$

(pour tout $\varepsilon > 0$ petit on a $\frac{f(x)-f(x-r)}{r} \leq \frac{f(x+\varepsilon)-f(x)}{\varepsilon} \leq \frac{f(x+r)-f(x)}{r}$ et conclure).

On peut également montrer qu'une fonction convexe admet une minorante affine

Proposition 2.2. *Soit $f : H \to \mathbb{R}$ une fonction convexe continue sur un ensemble convexe fermé C. Alors:*

$$\exists y \in H \; \exists \gamma \in \mathbb{R} \; \forall x \in C \; f(x) \geq \langle x, y \rangle + \gamma.$$

Preuve. Voir fiche d'exercice pour la preuve (cela utilise le théorème de Hahn-Banach). □

Proposition 2.3. *Soit $f : H \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe s.c.i. Soit (x_n) une suite qui converge faiblement vers x sur H. Alors:*

$$f(x) \leq \liminf_{n \to +\infty} f(x_n).$$

Preuve. Pour tout $\lambda \in \mathbb{R}$ tel que l'ensemble $A_\lambda := \{x \in H \; ; \; f(x) \leq \lambda\}$ soit non-vide, A_λ est un ensemble convexe fermé, donc il est faiblement fermé pour la topologie faible (voir le lemme ci-dessous 2.2 et pour plus de détails on se réfère à [9]). Or f est s.c.i. si et seulement si pour tout $\lambda \in \mathbb{R}$ tel que $A_\lambda \neq \emptyset$, alors A_λ est (fortement) fermé. Donc f est faiblement s.c.i. puisque tout sous-niveau de f est faiblement fermé. D'où le résultat.

□

Lemme 2.2. *Soit C un sous-ensemble convexe non-vide de H. Alors, C est fortement fermé si et seulement si C est faiblement fermé.*

Preuve. Supposons que C soit faiblement fermé. Soit alors $(x_n) \in C^n$ et $x \in H$ telle que $x_n \to x$ lorsque $n \to +\infty$. Alors (x_n) converge faiblement vers x, et comme C est faiblement fermé, on a donc $x \in C$. Donc, C est fortement fermé.

Supposons que C soit fortement fermé. Soit $(x_n) \in C^n$ et $x \in H$ telle que (x_n) converge faiblement vers x lorsque $n \to +\infty$. On suppose par l'absurde que $x \notin C$. Alors par le théorème de Hahn-Banach, il existe $p \in H$ et $\alpha \in \mathbb{R}$ tels que:

$$\forall y \in C,\ \langle p, x\rangle < \alpha < \langle p, y\rangle\,.$$

D'où pour tout $n \in \mathbb{N}$ on a $\langle p, x\rangle < \alpha < \langle p, x_n\rangle$ et une contradiction avec la convergence faible de (x_n) vers x. $\square$

2.4.2 Fonctions convexes de classe C^1 et C^2

On étudie maintenant des caractérisations des fonctions convexes dans le cas C^1.

Propriété 2.7. *Soit $f : H \to \mathbb{R}$ Fréchet différentiable. Alors on a l'équivalence:*
(i) La fonction f est convexe.
(ii) Pour tout $x \in X$ et tout $y \in X$ on a $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$.
(iii) Pour tout $x \in X$ et tout $y \in X$ on a $\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0$.

Preuve. Pour $(i) \Rightarrow (ii)$, on a

$$f(x + t(y-x)) \leq (1-t)f(x) + tf(y) \Rightarrow \frac{f(x+t(y-x)) - f(x)}{t} \leq f(y) - f(x)$$

et le résultat en faisant $t \to 0$. Pour $(ii) \Longrightarrow (iii)$ permuter x et y et additionner. Pour $(iii) \Rightarrow (ii)$ on utilise qu'il existe $t \in (0,1)$ tel que $f(y) - f(x) = \langle \nabla f(x + t(y-x)), y - x\rangle = \langle \nabla f(x), y - x\rangle + \langle \nabla f(x + t(y-x)) - \nabla f(x), y - x\rangle$ et le dernier terme étant positif on a le résultat. Enfin pour $(ii) \Rightarrow (i)$ on écrit que:

$$f(x) \geq f(tx + (1-t)y) + (1-t)\langle \nabla f(tx + (1-t)y), x - y\rangle\ ;\ f(y) \geq f(tx + (1-t)y) + t\langle \nabla f(tx + (1-t)y), y - x\rangle\,.$$

d'où le résultat en combinant les deux expressions (multiplier la première par t et la seconde par $1-t$ et additionner). $\square$

Lemme 2.3. *Soit $f : H \to \mathbb{R}$. La fonction f est fortement convexe de paramètre $\alpha > 0$ si et seulement si $f - \frac{\alpha}{2}\|\cdot\|$ est convexe.*

Preuve. Par un calcul on a que $f - \frac{\alpha}{2}\|\cdot\|$ est convexe si et seulement si pour tout $t \in [0,1]$

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

En ordonnant tous les termes quadratiques dans le membre de droite on trouve $-t(1-t)\frac{\alpha}{2}\|x-y\|^2$, d'où le résultat. $\square$

Propriété 2.8. *Soit $f : H \to \mathbb{R}$ Fréchet différentiable. Alors on a l'équivalence:*
(i) La fonction f est fortement convexe de paramètre $\alpha > 0$.
(ii) Pour tout $x \in X$ et tout $y \in X$ on a $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|y - x\|^2$.
(iii) Pour tout $x \in X$ et tout $y \in X$ on a $\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq \alpha\|x - y\|^2$.

Preuve. A faire en exercice en reprenant la preuve de la propriété précédente et en utilisant le lemme 2.3. $\square$

Dans le cas où f est deux fois dérivable on a le résultat suivant.

Proposition 2.4. *Soit $f : H \to \mathbb{R}$ deux fois différentiable. Alors f est α-elliptique si et seulement si*

$$\forall h \in H,\ \forall x \in H,\ \left\langle D^2(f)(x)h, h\right\rangle \geq \alpha\|h\|^2.$$

Preuve. On rappelle qu'étant donnée une fonction $f : H \to \mathbb{R}$ deux fois dérivable, alors f est convexe si et seulement si pour tout $x \in H$, la Hessienne de f en x est positive. Le résultat suit alors en utilisant $f - \frac{\alpha}{2}\|\cdot\|^2$. Montrons le premier point. Supposons f convexe. Alors pour $h \in H$ et $t > 0$ on a $\langle \nabla f(x + th) - \nabla f(x), th\rangle \geq 0$, d'où le résultat en divisant par t^2. Réciproquement on a la formule exacte:

$$\exists \theta \in [0,1] \quad f(y) = f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}D^2 f(x + \theta(y-x))(y-x, y-x).$$

D'où $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$ et on déduit que f est convexe. $\square$

Exercice 2.5. *1) Soit X un espace de Hilbert et $\|\cdot\|$ la norme associé au produit scalaire. Montrer que $\|\cdot\|^2$ est fortement convexe.*
2) Si $A \in \mathcal{L}(X, X^*)$ vérifie : il existe $\alpha > 0$ tel que pour tout $x \in X$ on ait $\langle Ax, x\rangle \geq \alpha \|x\|^2$ alors $x \longmapsto \langle Ax, x\rangle$ est fortement convexe.

On a également la propriété suivante (voir [7] p.18 ou fiche d'exercice).

Lemme 2.4. *Si f est fortement convexe sur H, alors il existe $\gamma > 0$ et $\beta > 0$ tel que $f(x) \geq \gamma\|x\|^2 + \delta$ pour tout $x \in H$.*

2.4.3 Un peu de sous-différentiabilité

Les fonctions convexes ne sont pas nécessairement différentiables comme le montre l'exemple de l'indicatrice d'un ensemble convexe ou bien la fonction $x \longmapsto |x|$.

Propriété 2.9. *Si f est convexe et $x \in Dom(f)$ alors la limite $\lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t}$ existe pour tout $d \in H$.*

Preuve. Soit $0 < t_1 < t_2$. On a $f(x+t_1d) = f\left(\frac{t_1}{t_2}(x+t_2d) + \frac{t_2-t_1}{t_2}x\right) \leq \frac{t_1}{t_2}f(x+t_2d) + \frac{t_2-t_1}{t_2}f(x)$. Ainsi, l'application $t \mapsto \frac{f(x+td)-f(x)}{t}$ est croissante sur $\mathbb{R}_+^*$, et donc la limite existe dans $\mathbb{R} \cup \{\pm\infty\}$ et

$$\lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t} = \inf_{t>0} \frac{f(x+td)-f(x)}{t}.$$

□

Définition 2.10. *La limite $\lim_{t\downarrow 0^+} \frac{f(x+td)-f(x)}{t}$ est appelée dérivée directionnelle de f dans la direction d en x.*

La propriété suivante est vraie en dimension finie (voir [14]).

Propriété 2.10. *Soit $f : H \to \mathbb{R}$ une fonction convexe et $x \in H$. Alors l'application $d \mapsto f'(x, d)$ est sous-linéaire i.e.*

$$\forall t > 0 \; f(tx) = tf(x) \quad \text{et} \quad \forall d_1, d_2 \in H,\; f'(x, d_1 + d_2) \leq f'(x, d_1) + f'(x, d_2).$$

De plus, il existe $L > 0$ tel que pour tout $d \in H$ on a $|f'(x, d)| \leq L\|d\|$.

Preuve. On montre d'abord que $d \mapsto f'(x, d)$ est convexe par l'inégalité:

$$f(t(x + d_1) + (1-t)(x + d_2)) - f(x) \leq t[f(x + d_1) - f(x)] + (1-t)[f(x + d_2) - f(x)].$$

La positive homogénéité est claire par la définition de la dérivée directionnelle. D'où la sous-linéarité de $d \mapsto f'(x, d)$. En admettant que f est localement Lipschitz, on déduit la dernière inégalité (en utilisant que $|f(x + td) - f(x)| \leq t$ pour $\|d\| = 1$ puis on conclue en utilisant l'homogénéité). □

On est amené à la définition suivante.

Définition 2.11. *Soit $f : H \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe. Pour $x \in Dom(f)$, on définit le sous-différentiel de f au point x par:*

$$\partial f(x) := \{p \in H \;;\; f(y) \geq f(x) + \langle p, y - x\rangle \;\forall y \in H\}.$$

Si $x \in H$, on dit que f est sous-différentiable en x si $\partial f(x) \neq \emptyset$. Les éléments de l'ensemble $\partial f(x)$ sont appelés les sous-gradients. Si f est Gâteaux-différentiable en x, et si $\partial f(x) \neq \emptyset$, alors $\partial f(x) = Df(x)$. On a le résultat suivant.

Propriété 2.11. *Pour tout $x \in Dom(f)$, l'ensemble $\partial f(x)$ est convexe fermé.*

Exemples:
- 1. $f(x) = |x| \Rightarrow \partial f(0) = [-1, 1]$.
- 2. Le sous-différentiel de la fonction indicatrice (2.4.1) en un point $x \in K$ est le **cône normal** à K au point x défini par:

$$N_K(x) := \{p \in H \;;\; \forall y \in K \;\langle p, y - x\rangle \leq 0\}.$$

On étudiera dans la suite du cours essentiellement le cas de fonctions différentiables.

2.5 Quelques résultats d'existence et d'unicité

Soit X un espace de Banach et $K \subset X$ un sous-ensemble non vide. Soit $f : X \to \mathbb{R}$ une application.

Définition 2.12. *On appelle infimum de f la valeur $l \in [-\infty, +\infty[$ telle que:*
(i) $\forall x \in K,\ f(x) \geq l$.
(ii) Il existe (x_n) une suite telle que pour tout n on ait $x_n \in K$ et $\lim_{n\to+\infty} f(x_n) = l$.

Cette valeur est notée $\inf_{x\in K} f(x)$. L'infimum existe toujours et il est fini si et seulement si f est minorée. Si f n'est pas minorée, l'infimum vaut $-\infty$. Une suite $(x_n) \in K$ telle que $\lim f(x_n) = \inf_{x\in K} f(x)$ est appelée **suite minimisante**.

Définition 2.13. *On appelle minimum de f sur K la valeur $l \in]-\infty, \infty]$ si elle existe pour laquelle il existe un élément $\bar{x} \in K$ tel que $\forall x \in K,\ f(x) \geq l = f(\bar{x})$. On note cette valeur $\min_{x\in K} f(x)$. On dit alors que f atteint son minimum sur K en $\bar{x}$.*

Plus généralement, on dit que
(i) $\bar{x}$ est un minimum global de f sur K ssi $\forall x \in K\ f(x) \geq f(\bar{x})$.
(ii) $\bar{x}$ est un minimum local de f sur K ssi $\exists r > 0\ \forall x \in K,\ \|x - \bar{x}\| \leq r \Rightarrow f(x) \geq f(\bar{x})$
(iii) $\bar{x}$ est un minimum global strict de f sur K ssi $\forall x \in K,\ x \neq \bar{x} \Rightarrow f(x) > f(\bar{x})$.
(iv) $\bar{x}$ est un minimum local de f sur K ssi $\exists r > 0\ \forall x \in K,\ \|x - \bar{x}\| \leq r$ et $x \neq \bar{x} \Rightarrow f(x) > f(\bar{x})$

Définition 2.14. *Soit $f : X \to \mathbb{R}$ une application. On dit que f est coercive si:*

$$f(x) \to +\infty \quad \text{lorsque } \|x\| \to +\infty.$$

Exemples:
- Soit $f(x) = x^T Ax + b^T x + c$ avec $A \in \mathcal{S}_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$, et $c \in \mathbb{R}$.
- Une fonction minorée par une fonction coercive est coercive.

On peut montrer la propriété suivante (voir fiche d'exercice).

Propriété 2.12. *Si f est fortement convexe alors f est coercive.*

On commence par le résultat classique d'existence sur un compact.

Théorème 2.4. *Soit (E, d) un espace métrique compact et $f : E \to \mathbb{R}$ une fonction semi-continue inférieurement. Alors il existe $x^* \in E$ tel que:*

$$f(x^*) = \inf_{x\in E} f(x).$$

Preuve. Considérons $(x_n) \in E^{\mathbb{N}}$ une suite minimisante. Par compacité, il existe $x^* \in E$ et une extraction $(x_{\varphi(n)})$ tels que $(x_{\varphi(n)})$ converge vers x^*. Comme f est s.c.i., on déduit que $f(x^*) \leq \liminf_{n\to+\infty} f(x_{\varphi(n)}) = \inf_{x\in E} f(x)$, d'où le résultat. $\square$

En dimension finie on a le résultat suivant.

Théorème 2.5. *Soit $E \subset \mathbb{R}^n$ un sous-ensemble fermé non-vide, et soit $f : E \to \mathbb{R}$ une fonction s.c.i. et coercive. Alors il existe $x^* \in E$ tel que:*

$$f(x^*) = \inf_{x\in E} f(x).$$

Preuve. Soit $(x_n) \in E^{\mathbb{N}}$ une suite minimisante i.e. $\lim_{n\to+\infty} f(x_n) = \inf_{x\in E} f(x) < +\infty$. Si (x_n) n'est pas bornée, alors il existe une extraction $(x_{\varphi(n)})$ telle que $\|x_{\varphi(n)}\| \to +\infty$. D'où $f(x_{\varphi(n)}) \to +\infty$ et une contradiction. Donc, (x_n) est bornée, ainsi il existe $x^* \in \mathbb{R}^n$ et une extraction $(x_{\psi(n)})$ tels que $x_{\psi(n)} \to x^*$. Comme E est fermé, $x^* \in E$. Il vient alors par le caractère s.c.i. $f(x^*) \leq \liminf_{n\to+\infty} f(x_{\psi(n)}) = \inf_{x\in E} f(x)$. $\square$

Dans le cas d'un ouvert, on a la propriété suivante.

Proposition 2.5. *Soit $K \subset \mathbb{R}^n$ un ouvert borné. On suppose que f est continue sur $\bar{K}$ et on suppose:*

$$\exists x_0 \in K\ \forall x \in \partial K \quad f(x) > f(x_0).$$

Alors le problème admet une solution.

Preuve. Comme $\bar{K}$ est compact, f admet un minimum $\bar{x}$ sur $\bar{K}$. Si $\bar{x} \in \partial K$, alors $f(x_0) < f(\bar{x})$, ce qui est absurde. Donc $\bar{x} \in K$ et on obtient le résultat car $K \subset \bar{K}$. ☐

En dimension quelconque on a le résultat fondamental suivant. Soit H un espace de Hilbert et $f : H \to \mathbb{R}$.

Théorème 2.6. *Soit K un ensemble convexe fermé non vide. Si f est convexe, s.c.i. et coercive sur K, alors il existe un minimum de f sur K i.e. il existe une solution du problème de minimisation:*

$$\inf_{x \in K} f(x).$$

Preuve. La preuve suivante reprend les arguments standards d'analyse fonctionnelle (voir [9]). On verra en exercice une preuve utilisant le théorème de projection.

Soit $x_0 \in K$ tel que $f(x_0) < +\infty$ et $\tilde{K} := \{x \in K \ ; \ f(x) \leq f(x_0)\}$. Alors, $\tilde{K}$ est un convexe (par convexité de f) et est fermé (par continuité de f). Comme $\tilde{K}$ est convexe fermé pour la topologie forte, il est faiblement fermé (voir [9]). De plus, $\tilde{K}$ est borné par coercivité de f. On déduit que $\tilde{K}$ est compact pour la topologie faible. Ainsi, f atteint son minimum sur $\tilde{K}$ (f convexe $\Rightarrow$ f est faiblement s.c.i.). Ainsi, il existe $x^* \in \tilde{K}$ tel que $f(x^*) \leq f(x)$ pour tout $x \in \tilde{K}$. Si $x \in K \backslash \tilde{K}$ alors $f(x) \geq f(x_0)$ d'où le résultat. ☐

Remarque 2.8. *On peut remplacer l'hypothèse de coercivité sur f par K borné.*

Théorème 2.7. *Soit K un ensemble convexe fermé non vide. Si f est fortement convexe sur K, alors il existe un unique minimum de f sur K.*

Preuve. Soit (x_n) une suite minimisante i.e. $f(x_n) \to c := \inf_{v \in K} f(x)$. Alors on a:

$$\frac{\alpha}{8}\|x_n - x_m\|^2 + f(\frac{x_n + x_m}{2}) \leq \frac{1}{2} f(x_n) + \frac{1}{2} f(x_m),$$

et on déduit que la suite (x_n) est de Cauchy (en retranchant c à l'inégalité ci-dessus) et donc elle converge vers $x \in K$ car K est fermé. On déduit le résultat par continuité. L'unicité provient du caractère strictement convexe (conséquence du caractère fortement convexe). ☐

Le résultat suivant porte sur la question de l'unicité d'une solution (essentiellement dans le cadre convexe) pour le problème:

$$\inf_{x \in K} f(x),$$

où $K \subset H$ est un ensemble fermé non vide d'un espace de Hilbert H et $f : H \to \mathbb{R}$.

Lemme 2.5. *On suppose K convexe. Alors:*
(i) Si f est convexe, tout minimum de f sur K est un minimum global et l'ensemble des minima est un convexe de H (éventuellement vide).
(ii) Si f est strictement convexe, alors il existe au plus un minimum global.

En dimension infinie, les ensembles fermés bornés ne sont pas des compacts. L'exemple ci-dessous montre que le théorème 2.5 n'est pas vrai en dimension infinie.

Exemple. Soit le problème d'optimisation:

$$\inf_{x \in H_0^1(]0,1[)} J(x) := \int_0^1 [(|x'(t)| - 1)^2 + x^2(t)]dt.$$

On peut montrer que si (x_n) converge vers x dans $H_0^1(]0,1[)$, alors il existe $C \geq 0$ tel que

$$\left| \int_0^1 ((|x_n'(t)| - 1)^2 - (|x'(t)| - 1)^2)dt \right| \leq \int_0^1 |x_n'(t) - x'(t)|(|x_n'(t)| + |x'(t)| + 2) \leq C\|x_n - x\|_{H_0^1(]0,1[)}.$$

Donc, on déduit que J est bien s.c.i. et on peut également vérifier que J est coercive. Considérons maintenant une suite $x_n : [0,1] \to \mathbb{R}$, dérivable presque partout telle que $\dot{x}_n(t) = \pm 1$ et $|x_n(t)| \leq \frac{1}{2n}$ Alors, $J(x_n) = \int_0^1 x_n(t)^2 dt \leq \frac{1}{4n^2} \to 0$, de plus $\alpha := \inf_{x \in H_0^1(]0,1[)} J(x) \geq 0$. Mais l'infimum 0 n'est pas atteint car on aurait $x'(t) = \pm 1$ presque partout et $x(t) = 0$ presque partout (ces deux conditions sont contradictoires).

2.6 Lemme de Farkas

Cette section porte sur le théorème de Farkas qui est un outil clef en optimisation notamment pour démontrer le théorème de Karush-Kuhn-Tucker. On montre d'abord le résultat d'enveloppe conique.

Proposition 2.6. *Soit E un espace vectoriel normé, $q \in \mathbb{N}^*$ et $a_1, ..., a_q$ q vecteurs de E. Soit l'ensemble:*

$$K := \left\{ \sum_{i=1}^{q} \lambda_i a_i \ ; \ \lambda_i \geq 0, \ 1 \leq i \leq q \right\}.$$

Alors K est un cône convexe fermé de E.

Preuve. L'ensemble K est un cône par construction. Montrons par récurrence qu'il est fermé. Pour $q = 1$, le résultat est clair. On suppose le résultat au rang $q \geq 1$. Soit $a_1, ..., a_q, a_{q+1}$ $q+1$ vecteurs de E et

$$K := \left\{ \sum_{i=1}^{q+1} \lambda_i a_i \ ; \ \lambda_i \geq 0, \ 1 \leq i \leq q+1 \right\}.$$

1er cas: $-a_i \in K$ pour tout $1 \leq i \leq q+1$. Alors $K = Vect(a_1, ..., a_{q+1})$ est fermé.
2ème cas: $\exists 1 \leq i \leq q+1 \ -a_i \notin K$. On peut supposer $i = q+1$ i.e. $-a_{q+1} \notin K$. Soit $(y_n) \in K^{\mathbb{N}}$ une suite qui converge vers y. Montrons que $y \in K$. Par définition:

$$y_n = \sum_{i=1}^{q} \lambda_{i,n} a_i + \mu_n a_{q+1} = z_n + \mu_n a_{q+1}.$$

Montrons que (μ_n) est bornée. Sinon, $\mu_n \to +\infty$ (à sous-suite près). Donc $-a_{q+1} = \lim \frac{z_n}{\mu_n}$ car (y_n) est bornée. Donc, par l'hypothèse de récurrence, on aurait que $-a_{q+1} \in K$ ce qui est absurde. Donc, (μ_n) est bornée. Ainsi, il existe $\mu \in \mathbb{R}$ tel que $\mu_n \to \mu$ (à sous-suite près). Comme (y_n) converge, il vient que z_n converge vers $z := y - \mu a_{q+1}$. Par hypothèse de réccurence, $z \in K$ et donc $y = z + \mu a_{q+1} \in K$. □

Cette propriété permet de montrer le lemme de Farkas.

Lemme 2.6. *Soit H un espace de Hilbert et $(a, a_1, ..., a_q) \in H^{q+1}$. On a alors l'équivalence suivante:*
(i) $\forall x \in H$, $\langle a_i, x \rangle \leq 0$, $1 \leq i \leq q \Rightarrow \langle a, x \rangle \leq 0$.
(ii) $\exists (\lambda_1, ..., \lambda_q) \in \mathbb{R}_+^q$ $a = \sum_{i=1}^{q} \lambda_i a_i$.

Preuve. On montre seulement $(i) \Rightarrow (ii)$. Soit $K := \{\sum_{i=1}^{q+1} \lambda_i a_i \ ; \ \lambda_i \geq 0, \ 1 \leq i \leq q\}$. Supposons que $a \notin K$. Par le théorème de séparation (voir (2.3.1)):

$$\exists \varepsilon > 0 \ \exists p \neq 0 \ \forall y \in K \quad \langle -p, y \rangle \leq \langle -p, a \rangle - \varepsilon. \tag{2.6.1}$$

En posant $q = -p$, on a donc $\sup_{y \in K} \langle q, y \rangle \leq \langle q, a \rangle - \varepsilon$. Fixons $y \in K$. Il vient donc $\sup_{t>0} \langle q, ty \rangle < +\infty$ et donc nécessairement $\langle q, y \rangle \leq 0$. Comme y est arbitraire dans K et $0 \in K$ on a donc $\sup_{y \in K} \langle q, y \rangle = 0$. Par (2.6.1) il vient donc $0 \leq \langle q, a \rangle - \varepsilon$ i.e. $\langle q, a \rangle \geq \varepsilon > 0$. Ceci contredit (i). En effet en prenant $y = a_i$, $1 \leq i \leq q$ on a $\langle q, a_i \rangle \leq 0$, $1 \leq i \leq q$. □

On peut montrer le corollaire suivant dans le cas d'inégalités et d'égalités.

Corollaire 2.1. *Soit H un espace de Hilbert et $(a_1, ..., a_p, a_{p+1}, ..., a_{p+q}, a) \in H^{p+q+1}$. On a alors l'équivalence suivante:*
(i) $\forall x \in H$, $\langle a_i, x \rangle \leq 0$, $1 \leq i \leq p$ et $\langle a_{p+i}, x \rangle = 0$, $1 \leq i \leq q \Rightarrow \langle a, x \rangle \leq 0$.
(ii) $\exists (\lambda_1, ..., \lambda_q) \in \mathbb{R}_+^q$ et $\exists (\lambda_{p+1}, ..., \lambda_{p+q}) \in \mathbb{R}^q$ $a = \sum_{i=1}^{p+q} \lambda_i a_i$.

Références

[1] G. Allaire, *Analyse numérique et optimisation*, Editions de l'ecole Polytechnique, 2005.

[2] D. Azé, J.-B. Hiriart-Urruty, *Analyse variationnelle et optimisation*, ellipses, 2010.

[3] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[4] M. Bergounioux, *Optimisation et contrôle des systèmes linéaires*, Dunod, 2001.

[5] D. M. Bertsekas, *Dynamic Prorgamming and Stochastic Control*, cours en ligne du MIT, 2011.

[6] F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer, 2000.

[7] F. Bonnans, *Optimisation Continue, Cours et Problèmes Corrigés*, Dunod, 2006.

[8] F. Bonnans, J.-C. Gilbert, C. Lemaréchal, C. Sagastizábal, *Numerical Optimization: Theoretical And Practical Aspects.*

[9] H. Brézis, *Analyse fonctionnelle Théorie et Applications*, Masson, 1987.

[10] P.G. Ciarlet, *Introduction à l'analyse matricielle et à l'optimisation*, 5ème édition, Dunod, 2007.

[11] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[12] O. Güler, *Foundations of Optimization*, Springer, 2010.

[13] J.-C. Culioli, *Introduction à l'optimisation*, Ellipses, 1994.

[14] J.-B. Hiriart-Urruty, *Les mathématiques du mieux faire, vol. 1, premiers pas en optimisation*, ellipses, 2010.

[15] J.-B. Hiriart-Urruty, *Optimisation et analyse convexe*, Mathématiques, PUF, 1998.

[16] J.-B. Hiriart-Urruty, *L'optimisation*, Que sais-je?, PUF, 1996.

[17] M. Minoux, *Programmation mathématique, théorie et algorithmes (tomes 1 et 2)*, Dunod 1983.

[18] T. Rockafellar, *Fundamentals of Optimization*, Lectures Note 2007,
http://www.math.washington.edu/$\sim$ rtr/mypage.html

[19] A. Ruszczynski, *Nonlinear Optimization*, Princeton University Press, 2006.

Chapitre 3

Conditions d'optimalité

L'objectif de cette section est de donner le théorème de Karush-Kuhn-Tucker en dimension quelconque. On étudiera en détails le cas de la dimension finie.

3.1 Optimisation sans contraintes et condition d'Euler

Théorème 3.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe C^1, et K un ensemble ouvert. Si $\bar{x}$ est une solution du problème*

$$\inf_{x \in K} f(x),$$

alors $\bar{x}$ vérifie

$$\nabla f(\bar{x}) = 0.$$

Preuve. Soit $v \in \mathbb{R}^n$. Comme K est ouvert, il existe $h_0 > 0$ tel que pour tout $h \in [0, h_0]$ on ait $\bar{x} + hv \in K$. Ainsi, $f(\bar{x} + hv) - f(\bar{x}) \geq 0$ et donc $h \langle \nabla f(\bar{x}), v \rangle + h\varepsilon(hv) \geq 0$. En divisant par $h > 0$ on obtient $\langle \nabla f(\bar{x}), v \rangle \geq 0$. En considérant $-v$, on déduit que $\langle \nabla f(\bar{x}), v \rangle = 0$. Comme ceci vaut pour tout $v \in \mathbb{R}^n$, on déduit le résultat. □

Exercice 3.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^2.*
a) Montrer que si $\bar{x}$ est un minimum local de f alors $D^2 f(\bar{x})$ est positive.
b) Montrer que si $\bar{x}$ est tel que $\nabla f(\bar{x}) = 0$ et si $D^2 f(\bar{x})$ est définie positive (i.e. il existe $\alpha > 0$ tel que $D^2 f(\bar{x})(h, h) \geq \alpha \|h\|^2$ pour tout $h \in \mathbb{R}^n$), alors $\bar{x}$ est un minimum local.

Solution: Etant donné $h \in \mathbb{R}^n$, on a $f(\bar{x} + h) = f(\bar{x}) + \langle \nabla f(\bar{x}), h \rangle + D^2 f(\bar{x})(h, h) + o(\|h\|^2)$

On s'intéresse maintenant au cas avec contraintes et on donne la condition nécessaire d'optimalité d'ordre 1. Soit $K \subset H$ un ensemble fermé non vide d'un espace de Hilbert H et $f : H \to \mathbb{R}$ une fonction continue. On s'intéresse au problème:

$$\inf_{x \in K} f(x). \tag{3.1.1}$$

Théorème 3.2. *On suppose K* **convexe** *fermé non-vide et f. On suppose que f est différentiable en un point u. Si u est un minimum local de f sur K, alors*

$$\forall v \in K, \quad \langle \nabla f(u), v - u \rangle \geq 0. \tag{3.1.2}$$

Si f est convexe et si $u \in K$ vérifie (3.1.2), alors u est un minimum global de f sur K.

Preuve. On a pour $0 < t < 1$, $u + t(v - u) \in K$ et donc $f(u + t(v - u)) - f(u) \geq 0$ d'où le résultat en divisant par t. De plus, lorsque f est convexe on a $f(v) - f(u) \geq \langle \nabla f(u), v - u \rangle \geq 0$ d'où le résultat. □

Remarque 3.1. *La direction $v - u$ est rentrante et la condition d'Euler s'interprète en disant que dans ces directions, localement la fonction ne peut que croître. De plus, si K est tout l'espace ou si u est dans l'intérieur de K, alors on obtient $\nabla f(u) = 0$.*

On exprime la même condition dans le cas où K n'est pas nécessairement convexe. On a la définition suivante.

Définition 3.1. *Soit K un sous-ensemble non-vide de H. Le cône tangent à K en x est alors définie par:*

$$T_K(x) := \{h \in H \ ; \ \exists t_n \downarrow 0 \ \exists h_n \to h \ ; \ x + t_n h_n \in K\}.$$

Remarque 3.2. *Il s'agit bien d'un cône fermé (la vérification est faite plus loin) parfois appelé cône des directions admissibles.*

L'inéquation d'Euler s'exprime ainsi.

Théorème 3.3. *Soit $f : H \to \mathbb{R}$ une fonction de classe C^1. Si x^* est un minimum local pour (3.1.1) alors:*

$$\forall h \in T_K(x), \ \langle \nabla f(x^*), h \rangle \geq 0. \tag{3.1.3}$$

Si K est un ensemble fermé convexe non-vide et si f est convexe sur K et si de plus x^* vérifie (3.1.3) alors x^* est un minimum global de f sur K.

Preuve. Soit $h \in T_K(x)$. Pour n assez grand $x + t_n h_n \in K$. Donc,

$$f(x^* + t_n h_n) - f(x^*) = t_n \langle \nabla f(x^*), h_n \rangle + t_n \|h_n\| \varepsilon(t_n h_n) \geq 0.$$

D'où le résultat en divisant par $t_n > 0$ et en faisant $n \to +\infty$. Réciproquement, comme f est convexe et de classe C^1, on a $f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle$ pour tout $x \in K$. Soit $t_n \downarrow 0$. Par convexité de K, $x^* + t_n(x - x^*) \in K$, et donc $x - x^* \in T_K(x^*)$, ainsi $\langle \nabla f(x^*), x - x^* \rangle \geq 0$, d'où le résultat. $\square$

Exercice 3.2. *Soit $K := \{x \in H \ ; \ \langle x, a_i \rangle = 0 \ ; \ 1 \leq i \leq m\}$ où a_i, $i = 1, ..., m$ sont fixés dans H. Montrer que* (3.1.3) *équivaut à*

$$\forall 1 \leq i \leq m \quad \exists \lambda_i \in \mathbb{R} \quad \nabla f(x) + \sum_{i=1}^{m} \lambda_i a_i = 0.$$

Solution: Comme K est un sous-espace vectoriel de H, on a par la condition d'Euler $\langle \nabla f(x), z \rangle = 0$ pour tout $z \in K$ i.e. z est dans l'orthogonal de K. Or $K^\perp = (H_1 \cap \cdots \cap H_m)^\perp = \sum_{1 \leq i \leq m} H_i^\perp$ où H_i est l'hyperplan de vecteur normal a_i. D'où le résultat en utilisant que $H_i^\perp = \mathbb{R} a_i$.

Exercice 3.3. *Soit A une matrice rectangulaire de taille (m, n), $m \geq n$ de rang n. Calculer l'unique solution du problème:*

$$\inf_{x \in \mathbb{R}^n} f(x) := \|Ax - b\|_2^2.$$

Solution: On note x^T le transposé du vecteur $x \in \mathbb{R}^n$ et A^T la transposée de la matrice A. On a $f(x) = \langle x, Bx \rangle - 2 \left\langle A^T b, x \right\rangle + \|b\|^2$ où $B = A^T A$ est de taille n et inversible (sinon $x^T A^T A x = 0$ implique $Ax = 0$ et $x = 0$ par le théorème du rang). D'où $\nabla f(x) = 2Bx - 2A^T b$ et le résultat.

3.2 Théorème de Fritz-John et Karush-Kuhn-Tucker

Soit $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq l$ et $h_j : \mathbb{R}^n \to \mathbb{R}$, $1 \leq j \leq m$ des fonctions de classe C^1. Soit K la contrainte fermée définie par :

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0, \ i \in I \ ; \ h_j(x) = 0 \ ; \ j \in J\},$$

et soit le problème de minimisation :

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

Dans toute la suite on notera $I := \{1, ..., l\}$ l'ensemble d'indices pour les contraintes d'égalité et $J := \{1, ..., m\}$ l'ensemble d'indices pour les contraintes d'innégalité.

Théorème 3.4. *(Fritz-John). Si un point $x^\star$ est un minimum du problème (P), alors il existe $p_0 \in \mathbb{R}_+$, $p \in \mathbb{R}_+^l$ et $q \in \mathbb{R}^m$ tels que :*

(i) $\sum_{i \in I} p_i g_i(x^\star) = 0$,

(ii) $(p_0, p, q) \neq 0$,

(iii) $p_0 \nabla f(x^\star) + \sum_{i \in I} p_i \nabla g_i(x^\star) + \sum_{j \in J} q_j \nabla h_j(x^\star) = 0.$

Preuve. Première étape. On considère la pénalisation suivante :

$$f_N(x) := f(x) + \|x - x^\star\|^2 + \frac{N}{2}\left[\sum_{i\in I}\max(0, g_i(x))^2 + \sum_{j\in J} h_j(x)^2\right]$$

La fonction f_N vérifie $f_N(x^\star) = f(x^\star)$. De plus, la fonction f_N est de classe C^1 en utilisant que $x \longmapsto \max(0, g_i(x))^2$ est de classe C^1 (pour ce dernier point, utiliser que g_i est de classe C^1 et que $\nabla(g_i(x)^2) = 2g_i(x)\nabla g_i(x)$, ce qui implique le résultat aux éventuels points x où $g_i(x) = 0$).
Seconde étape.

Lemme 3.1. *Il existe $\varepsilon_0 > 0$ tel que pour tout $\varepsilon \in]0, \varepsilon_0[$ il existe $N_\varepsilon > 0$ tel que pour tout $x \in \mathbb{R}^n$ vérifiant $\|x - x^\star\| = \varepsilon$ on ait $f_{N_\varepsilon}(x) > f(x^\star)$.*

Preuve du lemme. Comme $x^\star$ est un minimum local de f sur K il existe $\varepsilon_0 > 0$ tel que pour tout $x \in K$, si $\|x - x^\star\| \leq \varepsilon_0$ alors $f(x) \geq f(x^\star)$. Fixons $\varepsilon \in]0, \varepsilon_0[$. On raisonne par l'absurde. Supposons que pour tout N il existe x_N tel que $\|x_N - x^\star\| = \varepsilon$ et $f_N(x_N) \leq f(x^\star)$. Comme (x_N) est une suite bornée, elle converge à une sous-suite près vers un point noté $\bar{x}$. Montrons que $\bar{x}$ appartient à l'ensemble K et vérifie $\|\bar{x} - x^\star\| = \varepsilon$ et $f(\bar{x}) < f(x^\star)$. On aura ainsi une contradiction. Le fait que $\|\bar{x} - x^\star\| = \varepsilon$ est immédiat. D'autre part :

$$f(x_N) + \|x_N - x^\star\|^2 + \frac{N}{2}\left[\sum_{i\in I}\max(0, g_i(x_N))^2 + \sum_{j\in J} h_j(x_N)^2\right] \leq f(x^\star). \tag{3.2.1}$$

Donc :

$$0 \leq \sum_{i\in I}\max(0, g_i(x_N))^2 + \sum_{j\in J} h_j(x_N)^2 \leq \frac{2}{N}[f(x^\star) - f(x_N) - \varepsilon^2].$$

On fait $N \to \infty$ ce qui donne $0 \leq \sum_{i\in I}\max(0, g_i(\bar{x}))^2 + \sum_{j\in J} h_j(\bar{x})^2 \leq 0$. Ainsi on a bien $\bar{x} \in K$. De plus par (3.7.4) on a $f(x_N) + \varepsilon^2 \leq f(x^\star)$. Il vient donc en faisant $N \to \infty$

$$f(\bar{x}) \leq f(x^\star) - \varepsilon^2 < f(x^\star).$$

Or par définition de ε_0 il est impossible d'avoir $\bar{x} \in I$ et $\|\bar{x} - x^\star\| \leq \varepsilon_0$ et $f(\bar{x}) < f(x^\star)$. D'où une contradiction. □

Fixons maintenant $\varepsilon \in]0, \varepsilon_0[$. On sait alors que f_{N_ε} admet un minimum local sur l'ouvert $\{x \in \mathbb{R}^n \ ; \ \|x - x^\star\| < \varepsilon\}$ en un point noté x_ε (voir Proposition d'existence d'un minimum sur un ouvert au début du cours). En ce point la condition nécessaire d'optimalité s'applique et donc

$$\nabla f_{N_\varepsilon}(x_\varepsilon) = 0.$$

D'où :

$$\nabla f(x_\varepsilon) + 2(x_\varepsilon - x^\star) + N\left[\sum_{i\in I}\max(0, g_i(x_\varepsilon))\nabla g_i(x_\varepsilon) + \sum_{j\in J} h_j(x_\varepsilon)\nabla h_j(x_\varepsilon)\right] = 0.$$

On pose

$$\rho_\varepsilon := \left(1 + N^2\sum_{i\in I}\max(0, g_i(x_\varepsilon))^2 + N^2\sum_{j\in J} h_j(x_\varepsilon)^2\right)^{\frac{1}{2}}, \ p_0^\varepsilon := \frac{1}{\rho_\varepsilon}, \ p_i^\varepsilon := Np_0^\varepsilon\max(0, g_i(x_\varepsilon)), \ q_j^\varepsilon := Np_0^\varepsilon h_j(x_\varepsilon).$$

Lorsque ε tend vers 0, x_ε tend vers $x^\star$ et $(p_0^\varepsilon, p^\varepsilon, q^\varepsilon)$ qui est de norme 1 dans $\mathbb{R}^{1+l+m}$ (donc non-nul) tend à une sous-suite près vers un vecteur de norme 1 noté (p_0, p, q). En divisant la condition nécessaire obtenue par ρ_ε on obtient en faisant $\varepsilon \to 0$

$$p_0\nabla f(x^\star) + \left[\sum_{i\in I} p_i\nabla g_i(x^\star) + \sum_{j\in J} q_j\nabla h_j(x^\star)\right] = 0.$$

De plus, la condition de complémentarité est satisfaite. En effet, si $g_i(x^\star) < 0$ pour un certain i, alors on a $g_i(x^\star) < 0$ pour $\varepsilon > 0$ suffisamment petit. Donc $p_i^\varepsilon = 0$ par définition de p_i^ε. D'où à la limite $p_i = 0$. □

Une fois démontré le théorème de Fritz-John, on peut aisément démontrer le théorème de Karush-Kuhn-Tucker lorsque les contraintes sont qualifiées. Ceci permet de garantir que le multiplicateur λ_0 est non nul:

$$\lambda_0 \neq 0.$$

On considère donc le même problème que ci-dessus (avec les mêmes notations) :

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

Définition 3.2. *On dira que les contraintes sont qualifiées au point $x^\star$ si et seulement si pour tout $\lambda \in \mathbb{R}^l_+$ et tout $\mu \in \mathbb{R}^m$ les deux conditions*

$$\begin{cases} \sum_{1 \le i \le l} \lambda_i g_i(x^\star) = 0, \\ \sum_{1 \le i \le l} \lambda_i \nabla g_i(x^\star) + \sum_{1 \le j \le m} \mu_j \nabla h_j(x^\star) = 0, \end{cases} \tag{3.2.2}$$

impliquent $\lambda = 0$ et $\mu = 0$.

On ne peut pas utiliser cette définition lorsque les contraintes sont affines. Lorsque les contraintes sont affines, on peut montrer que celles-ci sont qualifiées : on utilise alors la définition 3.5.

Remarque 3.3. *Deux ensembles de contraintes peuvent définir le même ensemble, l'un étant qualifié et pas l'autre.*

Théorème 3.5. *(Karush-Kuhn-Tucker). Si un point $x^\star$ est un minimum du problème (P) et si les contraintes sont* **qualifiées**, *alors il existe $\lambda_i \ge 0$, $i \in I$, il existe $\mu_j \in \mathbb{R}$ tels que*

$$\begin{array}{ll} (i) & \sum_i \lambda_i g_i(x^\star) = 0, \\ (ii) & \nabla f(x^\star) + \sum_{i=1}^l \lambda_i \nabla g_i(x^\star) + \sum_{j=1}^m \mu_j \nabla h_j(x^\star) = 0. \end{array}$$

Preuve. On applique le théorème de Fritz-John. Si $p_0 = 0$ on a nécessairement $p = q = 0$ d'après la définition de la qualification des contraintes. D'où $(p_0, p, q) = 0$ et une contradiction avec le théorème de Fritz-John. Ainsi $p_0 \neq 0$. On pose ensuite $\lambda = \frac{p}{p_0}$ et $\mu = \frac{q}{p_0}$. On déduit (i) et (ii) du théorème de Karush-Kuhn-Tucker par (i) et (iii) du théorème de Fritz-John. □

Définition 3.3. *On appelle Lagrangien associé au problème (P), la fonction $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ définie par:*

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \sum_{i=1}^l \lambda_i g_i(x) + \sum_{j=1}^m \mu_j h_j(x).$$

La condition (i) s'appelle condition de **complémentarité** *tandis que la condition (ii) s'appelle* **stationarité du lagrangien** *et se réécrit:*

$$\nabla_x \mathcal{L}(x^*, \lambda, \mu) = 0.$$

Le vecteur (p, q) est appelé **multiplicateur de Lagrange** *du problème associé à la solution x^*.*

3.3 Critères de qualification

Pour vérifier que la condition de qualification est vérifiée en un point x, on peut utiliser (3.2.2). Néanmoins, il existe des conditions plus simples à vérifier lorsque par exemple les données sont convexes ou affines. Par ailleurs, dans le cas général (avec contraintes d'égalité et d'inégalité), on utilise plutôt la condition de Mangasarian-Fromowitz (voir Proposition 3.1 ci-dessous). L'objet de cette section est de détailler diverses conditions qui assurent la qualification des contraintes en fonction de la régularité des données.

3.3.1 Condition de Mangasarian-Fromowitz

Si cela n'est pas reprécisé, on raisonne avec l'ensemble de contraintes K défini par:

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \le 0, \ i \in I \ ; \ h_j(x) = 0 \ ; \ j \in J\},$$

Proposition 3.1. *(Mangasarian-Fromowitz). On suppose qu'au point $x \in K$ les deux conditions suivantes sont satisfaites:*

$$\begin{array}{ll} (i) & \{\nabla h_1(x), ..., \nabla h_m(x)\} \text{ est libre.} \\ (ii) & \exists v \in \mathbb{R}^n,\ \forall j \in J\ \ \langle \nabla h_j(x), v\rangle = 0 \text{ et } \forall i \in I(x^*)\ \ \langle \nabla g_i(x^*), v\rangle < 0. \end{array} \tag{3.3.1}$$

Alors la contrainte est qualifiée.

Ces deux conditions réunies portent le nom de **Mangasarian-Fromowitz**.

Preuve. On suppose

$$\begin{cases} \sum_{1\leq i\leq l} \lambda_i g_i(x) = 0, \\ \sum_{1\leq i\leq l} \lambda_i \nabla g_i(x) + \sum_{1\leq j\leq m} \mu_j \nabla h_j(x) = 0. \end{cases}$$

On a $\lambda_i = 0$ pour $i \notin I(x)$. D'où en multipliant par v, on obtient $\sum_{i\in I(x)} \lambda_i \langle \nabla g_i(x), v\rangle = 0$ et donc $\lambda_i = 0$ pour $i \in I(x)$. Donc $\sum_{1\leq j\leq m} \mu_j \nabla h_j(x) = 0$. Donc par indépendance linéaire $\mu_j = 0$ pour $j \in J$. $\square$

3.3.2 Qualification et cône de Bouligand

Définition 3.4. *Etant donné un ensemble $K \subset \mathbb{R}^n$ non vide et $x \in K$, on appelle cône de Bouligand à K au point x l'ensemble:*

$$T_K(x) := \{v \in H\ ;\ \exists s_k \downarrow 0\ \exists v_k \to v\ ;\ x + t_k v_k \in K\}.$$

Propriété 3.1. *L'ensemble $T_K(x)$ est un cône fermé.*

Preuve. On a $0 \in T_K(x)$. Soit maintenant $v \in T_K(x)$ et $\lambda > 0$. Alors: $x + \frac{t_k}{\lambda}\lambda v_k \in K$, d'où $\lambda v \in T_K(x)$.

Montrons la propriété de fermeture. Soit $v^p \to v$ avec $v^p \in T_K(x)$. Montrons que $v \in T_K(x)$. Par définition, pour tout $p \in \mathbb{N}$, $\exists s_k^p \downarrow 0$ et $\exists v_k^p \to v^p$ avec $x + s_k^p v_k^p \in K$. Pour tout $l \in \mathbb{N}$ il existe $p_l \in \mathbb{N}$ et $k_l \in \mathbb{N}$ tels que

$$\|v^{p_l} - v\| \leq \frac{1}{2l}, \quad \|v^{p_l} - v_{k_l}^{p_l}\| \leq \frac{1}{2l}, \quad s_{k_l}^{p_l} \leq \frac{1}{l},$$

et donc $\|v_{k_l}^{p_l} - v\| \leq \frac{1}{l}$, donc $w_l := v_{k_l}^{p_l}$ converge vers v. De plus, $s_l := s_{k_l}^{p_l} \downarrow 0$ et $x + s_l w_l = x + s_{k_l}^{p_l} v_{k_l}^{p_l} \in K$. $\square$

Ce cône est parfois appelé cône des directions admissibles.

Exercice 3.4. *Calculer $T_K(x,y)$ dans les cas suivants:*
1) K est le carré de $\mathbb{R}^2$ et $(x,y) \in K$.
2) $K := \{(x, y \in \mathbb{R}^2)\ ;\ x = \cos(1/t)\ y = t\sin(1/t)\ t > 0\} \cup \{(0,0)\}$ et $(x,y) = (0,0)$.
3) $K := \{(x,y) \in \mathbb{R}^2\ ;\ x^2 \leq y \leq 2x^2\}$ et $(x,y) = (0,0)$.

On rappelle le résultat.

Théorème 3.6. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1 et x^* une solution de*

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

Alors:

$$\forall v \in T_K(x),\ \langle \nabla f(x^*), h\rangle \geq 0. \tag{3.3.2}$$

On considère l'ensemble E défini par:

$$E := \{v \in \mathbb{R}^n\ ;\ \forall i \in I(x)\ \ \langle \nabla g_i(x), v\rangle \leq 0 \text{ et } \forall j \in J\ \ \langle \nabla h_j(x), v\rangle = 0\}.$$

Proposition 3.2. *Si les contraintes sont de classe C^1, alors:*

$$T_K(x) \subset E.$$

Preuve. Soit $v \in T_K(x)$. Alors pour tout $k \in \mathbb{N}$ et tout $i \in I$ on a $g_i(x+s_k v_k) \leq 0$ et pour tout $j \in J$, $h_j(x+s_k v_k) = 0$. Si $i \in I(x)$ alors $g_i(x) = 0$ d'où

$$g_i(x + s_k v_k) = s_k \langle \nabla g_i(x), v_k \rangle + \|s_k v_k\| \varepsilon(s_k v_k) \leq 0,$$

et donc $\langle \nabla g_i(x), v \rangle \leq 0$. De plus, si $j \in J$, on a:

$$0 = h_j(x + s_k v_k) = s_k \langle \nabla h_j(x), v_k \rangle + \|s_k v_k\| \varepsilon(s_k v_k),$$

d'où $\langle \nabla h_j(x), v \rangle = 0$. □

L'inclusion inverse n'est pas toujours vraie, d'où la définition suivante.

Définition 3.5. *La contrainte K est qualifiée si et seulement si*

$$T_K(x) = E = \{v \in \mathbb{R}^n \ ; \ \forall i \in I(x) \ \ \langle \nabla g_i(x), v \rangle \leq 0 \text{ et } \forall j \in J \ \ \langle \nabla h_j(x), v \rangle = 0\}.$$

En pratique, pour vérifier que les contraintes sont qualifiées, on peut adopter la définition 3.2. Cependant la définition 3.5 permet de démontrer le théorème de Karush-Kuhn-Tucker à l'aide du lemme de Farkas. De plus, lorsque par exemple les contraintes d'égalité et d'inégalité sont convexes ou affines, cette définition permet directement de voir si les contraintes sont qualifiées.

3.3.3 Preuve de KKT par le lemme de Farkas

On suppose les données de classe C^1. Soit l'ensemble

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0, \ i \in I \ ; \ h_j(x) = 0 \ ; \ j \in J\},$$

et le problème de minimisation:

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

La théorème de Karush-Kuhn-Tucker s'énonce ainsi.

Théorème 3.7. *Si un point x est un minimum du problème et si les contraintes sont* **qualifiées***, alors il existe $\lambda_i \geq 0$, $i \in I$, il existe $\mu_j \in \mathbb{R}$ tels que*

$$\begin{array}{ll} (i) & \sum_i \lambda_{i=1}^l g_i(x) = 0, \\ (ii) & \nabla f(x) + \sum_{i=1}^l \lambda_i \nabla g_i(x) + \sum_{j=1}^m \mu_j \nabla h_j(x) = 0. \end{array}$$

Preuve. La condition d'Euler implique:

$$\forall v \in T_K(x), \ \langle -\nabla f(x), v \rangle \leq 0.$$

Comme la contrainte est qualifiée au point x on a $T_K(x) = E$ où

$$E := \{v \in \mathbb{R}^n \ ; \ \forall i \in I(x) \ \ \langle \nabla g_i(x), v \rangle \leq 0 \text{ et } \forall j \in J \ \ \langle \nabla h_j(x), v \rangle \leq 0 \text{ et } \langle -\nabla h_j(x), v \rangle \leq 0\}.$$

Par le lemme de Farkas: $\forall i \in I \ \exists \lambda_i \geq 0$, $\forall j \in J \ \exists (\nu_j, \rho_j) \in \mathbb{R}^2$ tels que:

$$-\nabla f(x) = \sum_{i \in I(x)} \lambda_i \nabla g_i(x) + \sum_{j=1}^m (\nu_j \nabla h_j(x) - \rho_j \nabla h_j(x)).$$

D'où la condition de stationarité en posant $\mu_j = \nu_j - \rho_j$. En prenant $\lambda_i = 0$ pour $i \notin I(x)$, la condition précédente s'écrit bien $\nabla f(x) + \sum_{i \in I} \lambda_i \nabla g_i(x) + \sum_{j \in J} \mu_j \nabla h_j(x) = 0$. De plus, si $i \in I(x)$, alors $g_i(x) = 0$. Si $i \notin I(x)$, alors on a pris $\lambda_i = 0$. D'où la condition de complémentarité. □

3.3.4 Cas des contraintes affines

Une fonction $\phi : \mathbb{R}^n \to \mathbb{R}$ est affine si et seulement si il existe $a \in \mathbb{R}^n$ et $b \in \mathbb{R}$ tels que $\phi(x) = \langle a, x\rangle + b$ pour tout $x \in \mathbb{R}^n$. Notons qu'alors $\phi(y) - \phi(x) = \langle \nabla\phi(x), y - x\rangle$ pour tout $x, y \in \mathbb{R}^n$.

Proposition 3.3. *Si les fonctions g_i et h_j sont affines, alors la contraintes K est qualifiée en tout point.*

Preuve. Soit $v \in \mathbb{R}^n$ tel que $\forall i \in I(x)$ $\langle \nabla g_i(x), v\rangle \leq 0$ et $\forall j \in J$, $\langle \nabla h_j(x), v\rangle = 0$. Montrons que $v \in T_K(x)$. Soit $s > 0$. Alors pour $i \in I(x)$ et $j \in J$, on a (en utilisant le caractère affine des fonctions):

$$g_i(x + sv) = g_i(x) + s\langle \nabla g_i(x), v\rangle \leq 0, \quad h_j(x + sv) = h_j(x) + s\langle \nabla h_i(x), v\rangle = 0$$

De plus, si $i \notin I(x)$, alors $g_i(x) < 0$, donc par continuité de g_i, il existe $\bar{s}_i > 0$ tel que pour tout $s \in [0, \bar{s}_i]$ on ait $g_i(x + sv) < 0$. Prenons $\bar{s} = \min_{i \notin I(x)} \bar{s}_i$. Alors $\bar{s} > 0$ et pour tout $s \in [0, \bar{s}]$ pour tout $i \in I$ on a : $g_i(x + sv) \leq 0$ et pour tout $j \in J$ on a $h_j(x + sv) = 0$. Prenons donc $s_k := \bar{s}/k$ et $v_k := v$. Il vient donc $x + s_k v_k \in K$ pour tout $k > 0$. Donc $v \in T_K(x)$. □

On se reporte à la section des exercices pour généraliser ce résultat au cas où g_i est supposée concave (uniquement pour les contraintes d'inégalité).

3.3.5 Contraintes convexes

On suppose dans cette section que $I := \{1, ..., l\}$ et que $g_i : \mathbb{R}^n \to \mathbb{R}$ est convexe pour tout $i \in I$. Soit l'ensemble de contraintes:

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0, \ \forall i \in I\},$$

et l'on considère le problème

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

Définition 3.6. *Soit $x_0 \in K$. On dit que la condition de* **Slater** *est vérifiée en x_0 si et seulement si:*

$$\forall 1 \leq i \leq l, \ g_i(x_0) < 0.$$

Proposition 3.4. *On suppose qu'il existe $x_0 \in K$ tel que la condition de Slater soit vérifiée au point x_0. Alors la contrainte K est qualifiée.*

Preuve. Soit $x \in \partial K$. Donc $g_i(x) = 0$. Soit $v = x_0 - x$. On a $v \neq 0$ car x_0 n'est pas sur la frontière de K ($g_i(x_0) < 0$ pour tout $i \in I$). Soit $i \in I(x)$. Par convexité:

$$g_i(x_0) \geq g_i(x) + \langle \nabla g_i(x), x_0 - x\rangle .$$

D'où $\langle \nabla g_i(x), v\rangle < 0$. D'où Mangasarian-Fromowitz. □

On peut donc appliquer le théorème de KKT. Soit $\mathcal{L}(x, \lambda) := f(x) + \sum_{i=1}^{l} \lambda_i g_i(x)$, $x \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}^l_+$. Soit x^* une solution du problème. Alors, il existe $\lambda^* \in \mathbb{R}^l_+$ tel que:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0.$$

Réciproquement on a le résultat suivant.

Théorème 3.8. *On suppose que f est convexe et que la contrainte est qualifiée en tout point. Alors tout point vérifiant les conditions nécessaires d'optimalité est un minimum du problème.*

Preuve. L'application $x \longmapsto \mathcal{L}(x, \lambda^*)$ est convexe et $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$. Ainsi x^* est un minimum de $x \longmapsto \mathcal{L}(x, \lambda^*)$. Il vient pour tout $y \in K$:

$$\mathcal{L}(y, \lambda^*) \geq \mathcal{L}(x^*, \lambda^*) = f(x^*).$$

De plus $\lambda_i^* \geq 0$ et $g_i(y) \leq 0$, ainsi $f(y) \geq \mathcal{L}(y, \lambda^*)$, d'où le résultat. □

3.3.6 Condition de qualification pour les contraintes d'inégalité

On suppose dans cette section que

$$K := \{x \in \mathbb{R}^n \; ; \; g_i(x) \leq 0, \; \forall i \in I\}.$$

Proposition 3.5. *Soit $x \in K$. La contrainte K est qualifiée au point x si il existe $\bar{v} \in \mathbb{R}^n$ tel que pour tout $i \in I(x)$ on ait $\langle \nabla g_i(\bar{x}), \bar{v}\rangle < 0$.*

Preuve. Soit $F := \{v \in \mathbb{R}^n \; ; \; \forall i \in I(x), \langle \nabla g_i(x), v\rangle < 0\}$. On sait que $T_K(x) \subset \bar{F}$ (adhérence). Montrons l'autre inclusion. Soit $v \in F$ et $i \in I(x)$ et $s > 0$. On a :

$$g_i(x + sv) = g_i(x) + s \langle \nabla g_i(x), v\rangle + \|sv\|\varepsilon(sv).$$

Comme $g_i(x) = 0$ et $\langle \nabla g_i(x), v\rangle < 0$ il existe $\bar{s}_i$ tel que pour tout $s \in [0, \bar{s}_i]$ on ait

$$\langle \nabla g_i(x), v\rangle + \|v\|\varepsilon(sv) < 0.$$

Ceci montre que pour $i \in I(x)$ on a $g_i(x + sv) < 0$. Soit $\bar{s} := \min \bar{s}_i$. On déduit que pour tout $s \in [0, \bar{s}]$ alors $g_i(x + sv) < 0$ pour tout $i \in I$. Donc $x + sv \in K$ pour $s \in [0, \bar{s}]$. On prend $s_k := \bar{s}/k$ et $v_k = v$. Donc $v \in T_K(x)$. Donc $F \subset T_K(x)$.

Il reste maintenant à montrer que si v est tel que pour tout $i \in I(x)$ on a $\langle \nabla g_i(\bar{x}), v\rangle \leq 0$ il existe une suite v_k de limite v telle que pour tout $i \in I(x)$ on ait $\langle \nabla g_i(\bar{x}), v_k\rangle < 0$. Par l'hypothèse il existe $\bar{v} \in \mathbb{R}^n$ tel que pour tout $i \in I(x)$ on ait $\langle \nabla g_i(x), \bar{v}\rangle < 0$. On obtient le résultat en prenant $v_k := v + 1/k\bar{v}$. Pour conclure on a donc $F \subset T_K(x) \subset \bar{F}$ et comme $T_K(x)$ est fermé on obtient bien $T_K(x) = \bar{F}$. □

En particulier, on a que si $I(x) = \emptyset$, alors la contrainte est qualifiée.

3.3.7 Cas des contraintes d'égalité

On suppose dans cette section que

$$K := \{x \in \mathbb{R}^n \; ; \; h_j(x) \leq 0, \; \forall j \in J\}.$$

Proposition 3.6. *La contrainte K est qualifiée en un point $x \in K$ si la famille $\{\nabla h_1(x), ..., \nabla h_m(x)\}$ est libre.*

Preuve. Soit $G := \{v \in \mathbb{R}^n \; ; \; \forall j \in J, \langle \nabla h_j(x), v\rangle = 0\}$. On sait que $T_K(x) \subset G$. Montrons l'autre inclusion.
Etape 1. Montrons qu'il existe un voisinage U de x et $\alpha > 0$ tels que pour tout $x' \in U$ on ait :

$$\min_{\|y\|=1} \| \sum_{j=1}^{m} y_j \nabla h_j(x')\| \geq \alpha > 0.$$

Si l'assertion est fausse, il existe une suite (x_n) de limite x et une suite y_n de norme 1 telle que

$$\lim \| \sum_{1 \leq j \leq m} y_j^m \nabla h_j(x_n)\| = 0.$$

Par extraction (compacité) et continuité il existe y de norme 1 telle que

$$\| \sum_{1 \leq j \leq m} y_j \nabla h_j(x)\| = 0.$$

D'où $\sum_{1 \leq j \leq m} y_j \nabla h_j(x) = 0$ et une contradiction par l'hypothèse d'indépendance.
Etape 2. Posons $\theta_h := \left(\sum_{1 \leq j \leq m} h_j(x + hv)^2\right)^{1/2}$. Montrons que $\lim_{h \to 0} \frac{\theta_h}{h} = 0$. Comme $h_j(x) = \langle \nabla h_j(x), v\rangle = 0$, on a :

$$h_j(x + hv) = h_j(x) + h \langle \nabla h_j(x), v\rangle + h\|v\|\varepsilon_j(hv) = h\|v\|\varepsilon_j(hv).$$

Donc

$$\frac{\theta_h}{h} = \|v\| \left(\sum_{1 \leq j \leq m} \varepsilon_j(hv)^2 \right)^{1/2} \to 0 \quad \text{si } h \to 0.$$

Etape 3. Soit $v \in \mathbb{R}^n$ tel que pour tout $j \in J$ on ait $\langle \nabla h_j(x), v\rangle = 0$. Montrons que $v \in T_K(x)$. Soit la fonction

$$\phi_h(w) := \left(\sum_{1\leq j\leq m} h_j(x+hv+w)^2\right)^{1/2} + \frac{1}{2h}\|w\|^2.$$

On a $\phi_h(0) = \theta_h$ et de plus pour tout $w \in \mathbb{R}^n$ tel que $\|w\| = (2h\theta_h)^{1/2}$ on a $\phi_h(w) \geq \theta_h$. Donc d'après une proposition de cours (minimisation sur un ouvert!) ϕ_h admet un minimum w_h dans la boule ouvert de centre 0 et de rayon $(2h\theta_h)^{1/2}$. Supposons que la quantité

$$\left(\sum_{1\leq j\leq m} h_j(x+hv+w_h)^2\right)^{1/2} \tag{3.3.3}$$

soit nulle pour au moins une suite $h_k \to 0$. Posons alors $v_k = v + w_k/h_k$. Le point $x + h_k v_k$ appartient donc à $T_K(x)$. De plus $v_k \to v$ car $\|w_k/h_k\| \leq \frac{(2h_k\theta_k)^{1/2}}{h_k} \to 0$. Donc $v \in T_K(x)$.

Etape 4. Il reste à montrer qu'effectivement la quantité (3.3.3) est nulle pour au moins une suite $h_k \to 0$. Si ce n'est pas le cas, alors on suppose que la quantité

$$\left(\sum_{1\leq j\leq m} h_j(x+hv+w_h)^2\right)^{1/2}$$

n'est pas nulle pour h suffisamment petit. Alors en écrivant les conditions d'optimalité on obtient (en omettant la dépendance en $x + hv + w_h$):

$$2\frac{\sum_{1\leq j\leq m} h_j \nabla h_j}{(\sum_{1\leq j\leq m} h_j^2)^{1/2}} + \frac{1}{h} w_h = 0.$$

Alors le vecteur y^h de coordonnées $y_j^h := \frac{h_j}{(\sum_{1\leq j\leq m} h_j^2)^{1/2}}$ est de norme 1. Donc par la première étape on déduit l'inégalité

$$2\alpha \leq \left\| 2\frac{\sum_{1\leq j\leq m} h_j \nabla h_j}{(\sum_{1\leq j\leq m} h_j^2)^{1/2}} \right\| = \frac{1}{h}\|w_h\| \leq \frac{1}{h}(2h\theta_h)^{1/2} \to 0$$

lorsque $h \to 0$. On donc une contradiction. □

3.3.8 Cas général (contraintes d'égalités et d'inégalités)

On suppose que

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0, \ i \in I \ ; \ h_j(x) = 0 \ ; \ j \in J\}.$$

Proposition 3.7. *Soit $x \in K$. La contrainte est qualifiée en x si les conditions suivantes sont satisfaites :*
(i) La famille $\{\nabla h_1(x), ..., \nabla h_m(x)\}$ est libre.
(ii) Il existe $\bar{v} \in \mathbb{R}^n$ tel que

$$\forall j \in J, \ \langle \nabla h_j(x), \bar{v}\rangle = 0 \text{ et } \forall i \in I(x), \ \langle \nabla g_i(x), \bar{v}\rangle < 0.$$

Preuve. On sait déjà que

$$T_K(x) \subset G := \{v \in \mathbb{R}^n \ ; \ \forall i \in I(x) \ \langle \nabla g_i(x), v\rangle \leq 0, \ \forall j \in J \ \langle \nabla h_j(x), v\rangle = 0\}.$$

Montrons l'inclusion inverse. Montrons tout d'abord que :

$$C := \{v \in \mathbb{R}^n \ ; \ \forall i \in I(x) \ \langle \nabla g_i(x), v\rangle < 0, \ \forall j \in J \ \langle \nabla h_j(x), v\rangle = 0\} \subset T_K(x).$$

Considérons $\tilde{K} := \{x \in \mathbb{R}^n \ ; \ \forall j \in J \ h_j(x) = 0\}$. D'après ce qui précède on sait que la contrainte $\tilde{K}$ est qualifiée au point x. Ainsi, $T_{\tilde{K}}(x) = \{v \in \mathbb{R}^n \ ; \ \forall j \in J, \ \langle \nabla h_j(x), v\rangle = 0\}$. Par conséquent pour tout $v \in C$, il existe $s_k \to 0$ et

$v_k \to v$ tel que $x + s_k v_k \in \tilde{K}$. Montrons que $x + s_k v_k \in K$ dès que k est suffisamment grand. Soit $i \in I(x)$. Alors $g_i(x) = 0$ et $\langle \nabla g_i(x), v \rangle < 0$. Donc dès que k est assez grand :

$$g_i(x + s_k v_k) = s_k \langle \nabla g_i(x), v_k \rangle + \|s_k v_k\| \varepsilon(s_k v_k) = s_k(\langle \nabla g_i(x), v_k \rangle + \|v_k\| \varepsilon(s_k v_k)) \leq 0.$$

Si $i \notin I(x)$ alors il existe k_i tel que pour tout $k \geq k_i$ on a $g_i(x + s_k v_k) \leq 0$. Donc si on choisit $\bar{k} := \max k_i$ on obtient

$$\forall k \geq \bar{k}, \; g_i(x + s_k v_k) \leq 0.$$

Comme on sait déjà que $x + s_k v_k \in \tilde{K}$, cad. pour tout $j \in J$, $h_j(x + s_k v_k) = 0$ on déduit que pour tout $k \geq \bar{k}$ $x + s_k v_k \in K$. Ceci implique bien que $v \in T_K(x)$.

Pour conclure il suffit de montrer que

$$\bar{C} = \{v \in \mathbb{R}^n \; ; \; \forall i \in I(x), \; \langle \nabla g_i(x), v \rangle \leq 0 \text{ et } \forall j \in J \; \langle \nabla h_j(x), v \rangle = 0\}.$$

En effet comme $C \subset T_K(x)$ et comme $T_K(x)$ est fermé, on aura le résultat désiré. Remarquons d'abord que :

$$\bar{C} \subset \{v \in \mathbb{R}^n \; ; \; \forall i \in I(x), \; \langle \nabla g_i(x), v \rangle \leq 0 \text{ et } \forall j \in J \; \langle \nabla h_j(x), v \rangle = 0\}.$$

Pour l'inclusion inverse, soit $v \in \mathbb{R}^n$ tel que pour tout $i \in I(x)$ alors $\langle \nabla g_i(x), v \rangle \leq 0$ et pour tout $j \in J$ $\langle \nabla h_j(x), v \rangle = 0$. Montrons que $v \in \bar{C}$. Soit $v_k := v + 1/k\bar{v}$ où $\bar{v}$ est donné par les hypothèses. Alors il vient pour tout $j \in J$:

$$\langle \nabla h_j(x), v_k \rangle = \langle \nabla h_j(x), v \rangle + \frac{1}{k} \langle \nabla h_j(x), \bar{v} \rangle = 0$$

et pour tout $i \; \in I(x)$ on a :

$$\langle \nabla g_i(x), v_k \rangle = \langle \nabla g_i(x), v \rangle + \frac{1}{k} \langle \nabla g_i(x), \bar{v} \rangle < 0.$$

On a donc montré que $v_k \in C$, $v_k \to v$, ainsi $v \in \bar{C}$ et on a bien $\bar{C} = T_K(x)$. □

3.4 Résumé des conditions de qualification

On appellera x^* un minimum du problème (P). Le tableau ci-dessous résume les diverses conditions de qualification obtenue précédemment.

Contraintes	Conditions de qualification
$h(x) = 0$	$\nabla h(x^*) \neq 0.$
$h_j(x) = 0, \; j \in J$	$\{\nabla h_1(x^*), ..., \nabla h_m(x^*)\}$ libre
$g(x) \leq 0$	$\nabla g(x^*) \neq 0$ si $g(x^*) = 0$
$g(x) \leq 0$ et g convexe	$\exists \; x_0 \in \mathbb{R}^n$ t.q. $g(x_0) < 0.$
$g_i(x) \leq 0, \; i \in I$	$\exists v \in \mathbb{R}^n$ t.q. $\forall i \in I(x^*) \quad \langle \nabla g_i(x^*), v \rangle < 0.$
$g_i(x) \leq 0, \; i \in I$ et g_i convexes	$\exists \; x_0 \in \mathbb{R}^n \; \forall i \in I(x_0) \; g(x_0) < 0.$
$g_i(x) \leq 0, \; i \in I \; h_j(x) = 0, \; j \in J$	$\{\nabla h_1(x^*), ..., \nabla h_m(x^*)\}$ libre et $\exists v \in \mathbb{R}^n$ t.q. $\forall j \in J \; \langle \nabla h_j(x), v \rangle = 0$ et $\forall i \in I(x^*) \quad \langle \nabla g_i(x^*), v \rangle < 0.$
$g_i(x) \leq 0, \; i \in I \; h_j(x) = 0, \; j \in J$ g_i et h_j affines	

3.5 Théorème de Karush-Kuhn-Tucker en dimension infinie

Le théorème de KKT se généralise en dimension quelconque (voir par exemple [7]). Toutefois, la preuve de ce résultat nécessite d'être adaptée à ce cadre (utilisation du théorème des fonctions implicites). On énoncera donc juste le résultat dans le cas des contraintes d'égalité et d'inégalité.

3.5.1 Cas des contraintes d'égalité

Soit H un espace de Hilbert sur $\mathbb{R}$. Soit $f : H \to \mathbb{R}$ une application. On s'intéresse au problème (3.1.1) avec contrainte d'égalité i.e.

$$\inf_{x\in K} f(x) \qquad \text{avec} \quad K := \{x \in H \ ; \ h_j(x) = 0, \ 1 \le j \le m\}, \tag{3.5.1}$$

où $h_j : H \to \mathbb{R}$ est de classe C^1, $1 \le j \le m$. On suppose f de classe C^1.

Théorème 3.9. *Soit x un minimum local du problème* (3.5.1). *On suppose que la famille $\{\nabla h_j(x)\}_{1\le j\le m}$ est libre. Alors il existe $\mu \in \mathbb{R}^m$ tel que:*

$$\nabla f(x) + \sum_{1\le j\le m} \mu_j \nabla h_j(x) = 0.$$

Remarque 3.4. *(i) Comme en dimension finie, on peut exprimer la condition d'optimalité d'ordre 1 en disant que le Lagrangien*

$$(x, \mu) \in H \times \mathbb{R}^m \mapsto \mathcal{L}(x, \mu) := f(x) + \sum_{1\le j\le m} \mu_j h_j(x).$$

est stationnaire en (x^*, μ^*) i.e. $\nabla_x \mathcal{L}(x^*, \mu^*) = 0$ où (x^*, μ^*) est donné par le théorème 3.9.
(ii) L'indépendance linéaire des gradients $\{\nabla h_j(x)\}_{1\le j\le m}$ correspond à la condition de qualification de Mangasarian-Fromowitz (avec contraintes d'égalité uniquement).

Exemple. Première valeur propre du laplacien. On pose $H = H_0^1(\Omega)$ où Ω est un ouvert borné régulier de $\mathbb{R}^n$. Pour $u \in H$ on pose:

$$f(u) = \int_\Omega |\nabla u(x)|^2 dx \quad \text{et} \quad h(u) = \int_\Omega u(x)^2 dx - 1.$$

Montrons que le problème

$$\alpha := \inf_{u\in H_0^1(\Omega),\ h(u)=0} f(u)$$

admet une solution dans H. Soit (u_n) une suite minimisante. Comme la suite $(\|u_n\|_{H_0^1(\Omega)})$ est bornée et en utilisant l'injection compacte de $H_0^1(\Omega)$ dans $L^2(\Omega)$, on peut (à sous-suite près) supposer qu'il existe $u \in H_0^1(\Omega)$ telle que (u_n) converge fortement vers u dans $L^2(\Omega)$, et (u_n) converge faiblement vers u dans $H_0^1(\Omega)$. D'où $h(u) = 0$. Par convergence faible: $\alpha \le f(u) \le \liminf_{n\to+\infty} \|u_n\|_{H_0^1(\Omega)} = \alpha$, ainsi $f(u) = \alpha$ et $h(u) = 1$, d'où le résultat. Montrons que la contrainte est qualifiée. On a : $\nabla h(u)(v) = 2\int_\Omega u(x)v(x)dx$ et donc comme u n'est pas nulle presque partout car $h(u) = \int_\Omega u^2(x)dx = 1$, on déduit que $\nabla h(u)$ est non-nul et donc la contrainte est qualifiée. En utilisant KKT il existe $\mu \in \mathbb{R}$ tel que

$$\forall v \in H \quad \int_\Omega \nabla u \cdot \nabla v = \mu \int_\Omega uv,$$

ce qui montre que u est une solution faible du problème:

$$\begin{cases} -\Delta u & = \mu u \quad \text{dans } \Omega, \\ \quad u & = 0 \quad \text{sur } \partial\Omega, \end{cases}$$

et que μ est la plus petite valeur propre car $f(u) = \lambda$ (faire $u = v$ dans la formulation faible).

3.5.2 Cas des contraintes d'inégalité

On s'intéresse maintenant au problème (3.1.1) avec contraintes d'inégalité i.e.

$$K = \{x \in H \ ; \ g_i(x) \le 0, \ 1 \le i \le l\},$$

où $g_i : H \to \mathbb{R}$ est de classe C^1, $1 \le i \le l$. On suppose f de classe C^1. Pour $x \in H$ on définit l'ensemble des indices des contraintes actives par:

$$I(x) = \{i \in \{1, ..., l\} \ ; \ g_i(x) = 0\}.$$

On prendra comme condition de qualification la condition de **Mangasarian-Fromowitz**: la contrainte est dite qualifiée au point x si et seulement si :

$$\exists v \in H, \ \forall i \in I(x) \quad \langle \nabla g_i(x), v\rangle < 0,$$

dans le cas où les applications g_i ne sont pas affines. Dans le cas où pour tout $1 \leq i \leq l$, $x \mapsto g_i(x)$ est affine, alors on supposera:

$$\exists v \in H,\ \forall i \in \{1, ..., l\} \quad \langle \nabla g_i(x), v \rangle = 0.$$

Remarque 3.5. *Si les fonctions g_i, $1 \leq i \leq l$ sont convexes, alors la condition précédente peut être remplacée par la condition de* **Slater***:*

$$\exists x_0 \in H,\ \forall i\{1, ..., l\} \quad g_i(x_0) < 0.$$

Théorème 3.10. *Soit x un minimum local de f sur K. Si les contraintes sont qualifiées au point x alors il existe $\lambda \in \mathbb{R}^l_+$ tel que:*

$$\begin{cases} \nabla f(x) + \sum_{1 \leq i \leq l} \lambda_i \nabla g_i(x) = 0, \\ \sum_{1 \leq i \leq l} \lambda_i g_i(x) = 0. \end{cases}$$

La première condition est la condition de stationnarité du lagrangien associé au problème d'optimisation et la seconde condition est la condition de complémentarité.

3.6 Conditions d'optimalité au second ordre

Soit $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq l$ et $h_j : \mathbb{R}^n \to \mathbb{R}$, $1 \leq j \leq m$ des fonctions de classe C^2. Soit K la contrainte fermée définie par :

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0, \ i \in I \ ; \ h_j(x) = 0 \ ; \ j \in J\},$$

avec $I = \{1, ..., l\}$ et $J = \{1, ..., m\}$ et soit le problème de minimisation :

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

Soit $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^l_+ \times \mathbb{R}^m \to \mathbb{R}$ le lagrangien associé à (P) défini par

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \sum_{i \in I} \lambda_i g_i(x) + \sum_{j \in J} \mu_j h_j(x).$$

Pour $x \in K$, on pose $I(x) := \{i \in I \ ; \ g_i(x) = 0\}$ et :

$$\Lambda(x) := \{(\lambda, \mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m \ ; \ \nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \text{ et } \lambda_i = 0 \ \forall i \notin I(x)\}.$$

On définit le **cône critique** par

$$C(x) := \{v \in T_K(x) \ ; \ \langle \nabla f(x), v \rangle = 0\}.$$

Théorème 3.11. *Si x est un minimum local de f sur K et si la contrainte est qualifiée au point x, alors:*

$$\forall v \in C(x) \quad \max_{(\lambda, \mu) \in \Lambda(x)} \left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda, \mu) v, v \right\rangle \geq 0.$$

Ce résultat sera démontré après avoir étudié la section dualité. Une version plus simple de ce résultat est la suivante.

Théorème 3.12. *Soit x un minimum local du problème et (λ, μ) les multiplicateurs associés. Si les vecteurs:*

$$\nabla g_i(x), \ i \in I(x), \quad \nabla h_j(x), \ 1 \leq j \leq m,$$

sont linéairement indépendants, alors $\nabla^2_x \mathcal{L}(x, \lambda, \mu)$ est positif dans le sous-ensemble:

$$M := \mathrm{Vect}(\nabla g_i(x), \ i \in I(x), \quad \nabla h_j(x), \ 1 \leq j \leq m)^\perp,$$

i.e. si d vérifie $\langle d, \nabla g_i(x) \rangle = 0$, $i \in I(x)$, $\langle d, \nabla h_j(x) \rangle = 0$, $1 \leq j \leq m$, alors

$$\left\langle \nabla^2_x \mathcal{L}(x, \lambda, \mu) d, d \right\rangle \geq 0.$$

Pour faire la preuve, commençons par le lemme suivant (parfois appelé lemme de Gordan).

Lemme 3.2. *Soit $(a_i)_{1 \leq i \leq l}$ et $(b_j)_{1 \leq j \leq m}$, $l + m$ vecteurs de $\mathbb{R}^n$. Considérons les deux assertions suivantes:*

$$(i) \quad \exists \ y \in \mathbb{R}^n, \text{ tel que } \forall 1 \leq i \leq l, \ \langle a_i, y \rangle = 0 \quad \text{et} \quad \forall 1 \leq j \leq m, \ \langle b_j, y \rangle < 0.$$

$$(ii) \quad \text{La famille } \{a_1, ..., a_l, b_1, ..., b_m\} \text{ est lineairement independante.}$$

Alors (ii) $\Rightarrow$ (i).

Preuve. Supposons (ii) et montrons (i). Par l'absurde, supposons que (i) soit fausse. On déduit que pour tout vecteur $y \in \mathbb{R}^n$, alors la condition $\forall \ 1 \leq i \leq l$, $\langle a_i, y \rangle = 0$ et $\forall \ 1 \leq j \leq m$, $\langle b_j, y \rangle \leq t$ implique $t \geq 0$. Posons alors $\tilde{a}_i := (a_i, 0) \in \mathbb{R}^{n+1}$ et $\tilde{b}_j = (b_j, -1) \in \mathbb{R}^{n+1}$ pour $1 \leq i \leq l$ et $1 \leq j \leq m$, et soit $w := (0_{\mathbb{R}^n}, 1)$. Enfin, posons $z = (y, t) \in \mathbb{R}^{n+1}$. La condition précédente se réécrit donc de la façon suivante:

$$\forall z \in \mathbb{R}^{n+1}, \ \forall \ 1 \leq i \leq l, \ \langle \tilde{a}_i, z \rangle = 0 \text{ et } \forall \ 1 \leq j \leq m, \ \left\langle \tilde{b}_j, z \right\rangle \leq 0 \ \Rightarrow \ \langle -w, z \rangle \leq 0.$$

Par le lemme de Farkas, on déduit que il existe $\alpha \in \mathbb{R}^l$ et $\beta \in \mathbb{R}^m_+$ tels que:

$$-w = \sum_{1 \leq i \leq l} \alpha_i \tilde{a}_i + \sum_{1 \leq j \leq m} \beta_j \tilde{b}_j,$$

ce qui montre que l'on a une relation de liaison non triviale entre les vecteurs $\{a_1, ..., a_l, b_1, ..., b_m\}$. Ceci contredit (ii). $\square$

Preuve. On applique le lemme précédent avec la famille $\nabla g_i(x),\ i \in I(x), \quad \nabla h_j(x),\ 1 \le j \le m$. On déduit qu'il existe un vecteur $y \in \mathbb{R}^n$ vérifiant la condition (i) ci-dessus qui montre donc que la contrainte est qualifiée au point x. Ainsi, on a:

$$T_K(x) = M.$$

Donc il existe une suite de points réalisables (x_k) telle que $x_k \to x$ telle que $\frac{x_k - x}{\|x_k - x\|} \to d$ et telle que $g_i(x_k) = 0$ pour $i \in I(x)$ et $h_j(x_k) = 0$. D'où en posant $d_k := x_k - x$:

$$0 \le f(x_k) - f(x) = \mathcal{L}(x_k, \lambda, \mu) - \mathcal{L}(x, \lambda, \mu) = \frac{1}{2}\left\langle \nabla_x^2 \mathcal{L}(x, \lambda, \mu) d_k, d_k \right\rangle + o(\|d_k\|^2),$$

par un développement de Taylor. D'où le résultat en passant à la limite. □

On énonce maintenant le théorème de conditions suffisantes au second ordre.

Théorème 3.13. *Soit $\bar{x} \in K$ et $(\lambda, \mu) \in \Lambda(\bar{x})$. On suppose que pour tout $d \in \mathbb{R}^n$ non nul vérifiant*

$$\langle d, \nabla g_i(\bar{x})\rangle \le 0\ \forall i \in I(\bar{x})\ ;\ \langle d, \nabla g_i(\bar{x})\rangle = 0\ \forall i \in I(\bar{x}) \text{ et } \lambda_i > 0\ ;\ \langle d, \nabla h_j(\bar{x})\rangle = 0\ \forall j \in J,$$

on a

$$\left\langle \nabla_x^2 \mathcal{L}(\bar{x}, \lambda, \mu) d, d \right\rangle > 0.$$

Alors $\bar{x}$ est un minimum local strict de (P) et il existe $c > 0$ et $\varepsilon > 0$ tel que:

$$\forall x \in \mathbb{R}^n,\ \|x - \bar{x}\| \le \varepsilon \implies f(x) \ge f(\bar{x}) + c\|x - \bar{x}\|^2.$$

Preuve. On raisonne par l'absurde : il existe $\varepsilon_k \downarrow 0^+$ et $x_k \to \bar{x}$ tels que $f(x_k) < f(\bar{x}) + \varepsilon_k \|x_k - \bar{x}\|^2$. On pose $d_k := \frac{x_k - \bar{x}}{\|x_k - \bar{x}\|}$. On peut supposer qu'il existe $d \in \mathcal{S}^{n-1}$ tel que $d_k \to d$ lorsque $k \to \infty$. Montrons que

$$\frac{1}{2}\left\langle \nabla_x^2 \mathcal{L}(\bar{x}, \lambda, \mu) d_k, d_k \right\rangle + o(\|d_k\|^2) < \varepsilon_k \|d_k\|^2. \tag{3.6.1}$$

En utilisant que les données sont C^2 et en faisant un développement limité à l'ordre 2 de $\mathcal{L}$ en $\bar{x}$ on a:

$$\mathcal{L}(x_k, \lambda, \mu) = \mathcal{L}(\bar{x}) + \langle \nabla \mathcal{L}(\bar{x}, \lambda, \mu), x_k - \bar{x}\rangle + \frac{1}{2}\left\langle \nabla_x^2 \mathcal{L}(\bar{x}, \lambda, \mu)(x_k - \bar{x}), x_k - \bar{x} \right\rangle + o(\|x_k - \bar{x}\|^2)$$

Comme $(\lambda, \mu) \in \Lambda(x)$ on a $\nabla \mathcal{L}(\bar{x}, \lambda, \mu) = 0$. De plus comme $x_k \in K$ pour tout k on a $\mathcal{L}(x_k, \lambda, \mu) \le f(x_k)$. On déduit donc en utilisant $\mathcal{L}(\bar{x}) = f(\bar{x})$:

$$\mathcal{L}(x_k, \lambda, \mu) = f(\bar{x}) + \frac{1}{2}\left\langle \nabla_x^2 \mathcal{L}(\bar{x}, \lambda, \mu)(x_k - \bar{x}), x_k - \bar{x} \right\rangle + o(\|x_k - \bar{x}\|^2) \le f(x_k) < f(\bar{x}) + \varepsilon_k \|x_k - \bar{x}\|^2.$$

d'où (3.6.1) en divisant par $\|x_k - \bar{x}\|$ et en passant à la limite. Ainsi d satisfait $\left\langle \nabla_x^2 \mathcal{L}(\bar{x}, \lambda, \mu) d, d \right\rangle \le 0$.

Montrons maintenant que le vecteur d vérifie:

$$\langle \nabla f(\bar{x}), d\rangle \le 0,\ \langle \nabla g_i(\bar{x}), d\rangle \le 0,\ \langle \nabla h_j(\bar{x}), d\rangle = 0.$$

On effectue un développement limité de $f(x_k) - f(\bar{x})$, $g_i(x_k) - g_i(\bar{x})$, $i \in I(\bar{x})$ et de $h_j(x_k) - h_j(\bar{x})$ au premier ordre:

$$\begin{cases} f(x_k) = f(\bar{x}) + \langle \nabla f(\bar{x}), x_k - \bar{x}\rangle + o(\|x_k - \bar{x}\|) < f(\bar{x}) + \varepsilon_k \|x_k - \bar{x}\|^2, \\ g_i(x_k) = \langle \nabla g_i(\bar{x}), x_k - \bar{x}\rangle + o(\|x_k - \bar{x}\|) \le 0, \\ h_j(x_k) = \langle \nabla h_j(\bar{x}), x_k - \bar{x}\rangle + o(\|x_k - \bar{x}\|) = 0, \end{cases}$$

d'où le résultat. Pour conclure, la condition du premier ordre s'écrit $\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) + \sum_j \mu_j \nabla h_j(\bar{x}) = 0$ et donc:

$$\langle \nabla f(\bar{x}), d\rangle + \sum_{i \in I(\bar{x})} \lambda_i \langle \nabla g_i(\bar{x}), d\rangle + \sum_j \mu_j \langle \nabla h_j(\bar{x}), d\rangle = 0$$

et donc on déduit que pour tout $i \in I(\bar{x})$ on a $\lambda_i \langle \nabla g_i(\bar{x}), d\rangle = 0$ (sinon le membre de gauche de l'inégalité précédente serait strictement négatif et donc on ne pourrait avoir $\nabla \mathcal{L}(\bar{x}, \lambda, \mu) = 0$). Lorsque $i \in I(\bar{x})$ et $\lambda_i > 0$ on a donc $\langle \nabla g_i(\bar{x}), d\rangle = 0$. Par conséquent on a bien trouvé un vecteur d non nul qui vérifie

$$\langle d, \nabla g_i(\bar{x})\rangle \le 0\ \forall i \in I(\bar{x})\ ;\ \langle d, \nabla g_i(\bar{x})\rangle = 0\ \forall i \in I(\bar{x}) \text{ et } \lambda_i > 0\ ;\ \langle d, \nabla h_j(\bar{x})\rangle = 0\ \forall j \in J,$$

et pour lequel $\left\langle \nabla_x^2 \mathcal{L}(\bar{x}, \lambda, \mu) d, d \right\rangle \le 0$. Ceci contredit donc l'hypothèse du théorème 2, d'où le résultat. □

3.7 Dualité

3.7.1 Généralités

Définition 3.7. *Soit X et Y deux ensembles non-vides. Soit $\ell : X \times Y \to \mathbb{R}$ $(x, y) \longmapsto \ell(x, y)$ une fonction de deux variables x et y. On dira qu'un point $(\bar{x}, \bar{y}) \in X \times Y$ est un point selle si et seulement si*

$$\forall (x, y) \in X \times Y, \quad \ell(\bar{x}, y) \leq \ell(\bar{x}, \bar{y}) \leq \ell(x, \bar{y}).$$

On a les propriétés suivantes.

Propriété 3.2. *(i) La valeur $\ell(\bar{x}, \bar{y})$ est constante pour tous les points selles $(\bar{x}, \bar{y}) \in X \times Y$ de ℓ.*
(ii) Si $(\bar{x}_1, \bar{y}_1)$ et $(\bar{x}_2, \bar{y}_2)$ sont deux points selles, alors il en est de même de $(\bar{x}_1, \bar{y}_2)$ et $(\bar{x}_2, \bar{y}_1)$.

Preuve. Montrons (i). On a $\ell(x_1, y) \leq \ell(x_1, y_1) \leq \ell(x, y_1)$ et $\ell(x_2, y) \leq \ell(x_2, y_2) \leq \ell(x, y_2)$. On déduit:

$$\ell(x_1, y_2) \leq \ell(x_1, y_1) \leq \ell(x_2, y_1) \leq \ell(x_2, y_2) \leq \ell(x_1, y_2),$$

d'où (i). Pour (ii), on déduit de ci-dessus que $\ell(x_1, y_2) = \ell(x_2, y_1) = \ell(x_1, y_1) = \ell(x_2, y_2)$. D'où: $\ell(x_1, y) \leq \ell(x_1, y_1) = \ell(x_1, y_2) = \ell(x_2, y_1) \leq \ell(x, y_2)$ ce qui montre que (x_1, y_2) est un point selle. On procède de même pour (x_2, y_1). $\square$

On définit deux points φ (à valeurs $+\infty$ éventuellement) et ψ (à valeurs $-\infty$ éventuellement) par:

$$\varphi(x) := \sup_{y \in Y} \ell(x, y), \quad \psi(y) := \inf_{x \in X} \ell(x, y).$$

Propriété 3.3. *(i) On a $\psi(y) \leq \ell(x, y) \leq \varphi(x)$ pour tout $(x, y) \in X \times Y$.*
(ii) L'application ℓ admet des points selles sur $X \times Y$ si et seulement si on a l'égalité $\min_{x \in X} \varphi(x) = \max_{y \in Y} \psi(y)$.

Preuve. Le point (i) est évident par la définition de φ et ψ. Supposons que ℓ ait un point selle (x_1, y_1). Alors $\ell(x_1, y) \leq \ell(x_1, y_1) \leq \ell(x, y_1)$ et donc $\varphi(x_1) \leq \ell(x_1, y_1) \leq \psi(y_1)$. Mais on a $\ell(x_1, y_1) \leq \varphi(x_1)$ et $\psi(y_1) \leq \ell(x_1, y_1)$, d'où $\varphi(x_1) = \ell(x_1, y_1) = \psi(y_1)$. Enfin, par (i): $\psi(y) \leq \ell(x, y) \leq \varphi(x)$ pour tout (x, y). Donc, $\max_y \psi(y) \leq \min_x \varphi(x)$. D'où l'égalité. La réciproque est similaire. $\square$

3.7.2 Points selles de Lagrangien

On s'intéresse maintenant aux points selles de Lagrangiens. Soit $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq l$ et $h_j : \mathbb{R}^n \to \mathbb{R}$, $1 \leq j \leq m$ des fonctions de classe C^1. Soit K la contrainte fermée définie par :

$$K := \{x \in \mathbb{R}^n \; ; \; g_i(x) \leq 0, \; i \in I \; ; \; h_j(x) = 0 \; ; \; j \in J\},$$

avec $I = \{1, ..., l\}$ et $J = \{1, ..., m\}$ et soit le problème de minimisation :

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

Soit $\mathcal{L} : \mathbb{R}^n \times (\mathbb{R}^l_+ \times \mathbb{R}^m) \to \mathbb{R}$ le lagrangien associé à (P) défini par

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \sum_{i \in I} \lambda_i g_i(x) + \sum_{j \in J} \mu_j h_j(x).$$

Proposition 3.8. *Un point selle de $\mathcal{L}$ sur $\mathbb{R}^n \times (\mathbb{R}^l_+ \times \mathbb{R}^m)$ est exactement un couple $(\bar{x}, (\bar{\lambda}, \bar{\mu}))$ tel que:*
(i) $\bar{x}$ minimise $\mathcal{L}(\cdot, \bar{\lambda}, \bar{\mu})$
(ii) $\bar{x} \in K$
(iii) $\bar{\lambda}_i g_i(\bar{x}) = 0$ pour tout $1 \leq i \leq l$.

Preuve. Supposons d'abord que $(\bar{x}, \bar{\lambda}, \bar{\mu})$ soit un point selle de $\mathcal{L}$ i.e.

$$\forall x \in \mathbb{R}^n, \; \forall (\lambda, \mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m, \; \mathcal{L}(\bar{x}, \lambda, \mu) \leq \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) \leq \mathcal{L}(x, \bar{\lambda}, \bar{\mu}). \tag{3.7.1}$$

On déduit donc (i). Pour montrer (ii), on a que $(\bar{\lambda}, \bar{\mu})$ maximise $\mathcal{L}(\bar{x}, \cdot, \cdot)$ sur $\mathbb{R}^l_+ \times \mathbb{R}^m$. Donc en utilisant la condition d'Euler[1]:

$$\nabla_{(\lambda,\mu)}\mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) \in N_{\mathbb{R}^l_+ \times \mathbb{R}^m}(\bar{\lambda}, \bar{\mu}) = \prod_{i=1}^{l} N_{\mathbb{R}_+}(\bar{\lambda}_i) \times \prod_{j=1}^{m} N_{\mathbb{R}}(\bar{\mu}) = \prod_{i=1}^{l} N_{\mathbb{R}_+}(\bar{\lambda}_i) \times \mathbb{R}^m.$$

On déduit que $h_j(x) = 0$ pour $1 \le j \le m$ et que $g_i(x) \le 0$ pour $1 \le i \le l$, ce qui montre (ii). Pour (iii), soit $\bar{\lambda}_i = 0$, soit sinon $\bar{\lambda}_i > 0$ et donc $N_{\mathbb{R}_+}(\bar{\lambda}_i) = \mathbb{R}$ et donc $g_i(\bar{x}) = 0$, d'où (iii).

Réciproquement, la condition (i) implique la deuxième inégalité dans (3.7.1). De plus, comme $\bar{x} \in K$ et $\lambda_i \ge 0$, on a:

$$\mathcal{L}(\bar{x}, \lambda, \mu) = f(\bar{x}) + \sum_{i=1}^{l} \lambda_i g_i(\bar{x}) + \sum_{j=1}^{m} \mu_j h_j(\bar{x}) \le f(\bar{x}),$$

et $f(\bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu})$ par (ii) et (iii). D'où le résultat. □

Corollaire 3.1. *Si $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point selle de $\mathcal{L}$ alors $\bar{x}$ vérifie les deux conditions de KKT:*

$$\nabla\mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0 \text{ et } \sum_{i=1}^{l} \lambda_i g_i(x) = 0.$$

et est une solution du problème (P).

Preuve. Supposons que $(\bar{x}, \bar{\lambda}, \bar{\mu})$ soit un point selle. La proposition 3.8 entraine que $\bar{x}$ vérifie les deux conditions de KKT. De plus, pour $x \in K$, on a:

$$f(x) \ge \mathcal{L}(x, \bar{\lambda}, \bar{\mu}) \ge \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}).$$

□

Pour $(\lambda, \mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m$ on pose

$$d(\lambda, \mu) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu).$$

Propriété 3.4. *Le problème*

$$\sup_{(\lambda,\mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m} d(\lambda, \mu) \tag{3.7.2}$$

est un problème de maximisation concave.

Preuve. On a pour $t \in [0, 1]$, pour $\lambda^1, \lambda^2 \in \mathbb{R}^l_+$ et $\mu^1, \mu^2 \in \mathbb{R}^m$:

$$\mathcal{L}(x, t\lambda^1 + (1-t)\lambda^2, t\mu^1 + (1-t)\lambda^2) = t\mathcal{L}(x, \lambda^1\mu^1) + (1-t)\mathcal{L}(x, \lambda^2, \lambda^2),$$

d'où en prenant la borne inférieure:

$$d(t\lambda^1 + (1-t)\lambda^2, t\mu^1 + (1-t)\lambda^2) \ge td(\lambda^1, \mu^1) + (1-t)d(\lambda^2, \mu^2).$$

□

Définition 3.8. *Le problème* (3.7.2) *est appelé problème dual du problème (P).*

Proposition 3.9. *On suppose que f et g_i, $i \in I$ sont convexes et que h_j, $j \in J$ est affine. Alors $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point selle de $\mathcal{L}$ sur $\mathbb{R}^n \times (\mathbb{R}^l_+ \times \mathbb{R}^m)$ si et seulement si $\bar{x}$ est une solution de (P) et $(\bar{\lambda}, \bar{\mu})$ est un multiplicateur de Lagrange.*

Preuve. La condition nécessaire est assurée par le corollaire 3.1. Supposons que $\bar{x}$ soit une solution de (P) et $(\bar{\lambda}, \bar{\mu})$ un multiplicateur de Lagrange. Le Lagrangien est maintenant convexe par rapport à x, aussi comme au point $(\bar{x}, \bar{\lambda}, \bar{\mu})$ on a $\nabla_x(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$ on déduit que $\bar{x}$ minimise $x \mapsto \mathcal{L}(x, \bar{\lambda}, \bar{\mu})$. Enfin, pour $x \in K$, $\mathcal{L}(\bar{x}, \lambda, \mu) \le f(\bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu})$ car $(\bar{\lambda}, \bar{\mu})$ est un multiplicateur de Lagrange. □

[1]Soit $C \subset \mathbb{R}^n$ un ensemble fermé convexe non-vide et $x \in C$. On appelle $N_C(x)$ le cône normal à C au point x défini par $N_C(x) := \{p \in \mathbb{R}^n\ ;\ \langle p, y - x\rangle \le 0\ \forall y \in C\}$. Etant donnée $f : C \to \mathbb{R}$ de classe C^1, la condition d'Euler d'optimalité d'ordre 1 s'écrit $\langle\nabla f(x), v\rangle \ge 0$ pour tout $v \in T_K(x)$ ou bien encore $-\nabla f(x) \in N_K(x)$, x étant supposé minimum local de f sur C

3.7.3 Problèmes convexes et dualité

Dans cette section, on s'intéresse au problème

$$\text{(P)} \qquad \inf_{x \in K} f(x) \quad K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0 \quad 1 \leq i \leq l\},$$

où $f, g_i : \mathbb{R}^n \to \mathbb{R}$ sont de classe C^1 et convexes.

Théorème 3.14. *On suppose que le problème est qualifié et qu'il admet au moins une solution. Alors:*

$$\min_{x \in K} f(x) = \sup_{\lambda \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda) = \inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda).$$

De plus, le problème $\sup_{\lambda \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda)$ *admet au moins une solution* λ^* *et le problème* $\inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*)$ *admet une solution* x^* *qui est solution du problème (P).*

Remarque 3.6. *(i) Le problème est qualifié si par exemple* $\exists x_0 \in K \ \forall 1 \leq i \leq l \ g_i(x_0) < 0$ *(Slater).*
(ii) On a toujours :

$$\sup_{\lambda \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda) \leq \inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda).$$

(iii) Le résultat est faux si le problème n'est pas convexe (voir section suivante).

Preuve. (du théorème de dualité). Comme (P) admet une solution et que la contrainte est qualifiée en tout point, il existe $\lambda^* \in \mathbb{R}^l_+$ tel que $\langle \lambda^*, g(x^*) \rangle = 0$ et $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$. Comme le lagrangien est convexe:

$$\forall x \in \mathbb{R}^n \ \mathcal{L}(x, \lambda^*) \geq \mathcal{L}(x^*, \lambda^*) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*).$$

D'où

$$\sup_{\lambda \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda) \geq \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*) = f(x^*).$$

Ainsi

$$\mathcal{L}(x^*, \lambda^*) = \min_{x \in K} f(x) \leq \sup_{\lambda \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda).$$

Par le lemme ci-dessous et ce qui précède, on a donc:

$$\inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda) = \inf_{x \in K} \sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda) = \inf_{x \in K} f(x) \leq \sup_{\lambda \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda).$$

Or l'autre inégalité est toujours vraie, d'où le résultat. $\square$

Lemme 3.3. *On a*

$$\sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda) = \begin{cases} f(x) & x \in K, \\ +\infty & x \notin K. \end{cases}$$

Preuve. Tout d'abord si $x \in K$, on a le résultat puisque:

$$f(x) = \mathcal{L}(x, 0) \leq \sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda) = f(x) + \sum_{1 \leq i \leq l} \lambda_i g_i(x) \leq f(x).$$

Supposons donc $x \notin K$. Alors il existe i_0 tel que $g_{i_0}(x) > 0$. Soit $\lambda^k := (0, ..., 0, k, 0, ..., 0)$ où k st à l'indice i_0. Alors $\mathcal{L}(x, \lambda^k) = f(x) + k g_{i_0}(x) \to +\infty$ lorsque $k \to +\infty$, donc $\sup_{\lambda \in \mathbb{R}^l_+} \mathcal{L}(x, \lambda) = +\infty$. $\square$

Définition 3.9. *Le problème dual du problème (P) est le problème (D) défini par:*

$$\text{(D)} \qquad \max_{\lambda \in \mathbb{R}^l_+} d(\lambda) \quad d(\lambda) := \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda).$$

3.7.4 Exemple de saut de dualité

Soit le problème
$$\min_{|x|\leq 1/2} \frac{x^4}{4} - \frac{x^2}{2}.$$
Il est clair que le minimum est atteint en $x = \pm 1/2$. Soit $\mathcal{L}(x, \lambda_1, \lambda_2) := \frac{x^4}{4} - \frac{x^2}{2} + \lambda_1(x - 1/2) + \lambda_2(-x - 1/2)$ avec $\lambda_i \geq 0$. Regardons $\min_{x\in\mathbb{R}} \mathcal{L}(x, \lambda_1, \lambda_2)$. Alors un minimum vérifie $\mathcal{L}'_x(x, \lambda_1, \lambda_2) = 0$ et $\mathcal{L}''_{x,x}(x, \lambda_1, \lambda_2) \geq 0$ i.e. $3x^2 - 1 \geq 0$, donc $|x| \geq 1/\sqrt{3}$. Ainsi, le minimum de $x \mapsto \mathcal{L}(x, \lambda_1, \lambda_2)$ n'est pas atteint dans $[-1/2, 1/2]$.

On peut également étudier l'exercice suivant.

Exercice 3.5. *On considère le problème d'optimisation*
$$\inf x_1x_2 \quad \text{t.q. } x_1 \geq 0,\ x_2 \geq 0,\ x_1^2 + x_2^2 \leq 1.$$
Soit $\mathcal{L}(x, \lambda_1, \lambda_2, \lambda_3) := x_1x_2 - \lambda_1x_1 - \lambda_2x_2 + \lambda_3(x_1^2 + x_2^2 - 1)$ *le Lagrangien associé au système,* $\lambda := (\lambda_1, \lambda_2, \lambda_3)$ *et*
$$d(\lambda) := \inf_{(x_1,x_2)\in\mathbb{R}^2} \mathcal{L}(x, \lambda_1, \lambda_2, \lambda_3) \tag{3.7.3}$$
1) Calculer la solution du problème initial.
2) On considère le problème $d(\lambda)$. *Montrer que la solution vérifie le système*
$$\begin{cases} 2\lambda_3x_1 + x_2 = \lambda_1, \\ x_1 + 2\lambda_3x_2 = \lambda_2, \end{cases}$$
puis l'égalité $\lambda_3(x_1^2 + x_2^2) + x_1x_2 = \frac{1}{2}(\lambda_1x_1 + \lambda_2x_2)$. *On notera* A *la matrice du système précédent. Déduire que*
$$d(\lambda) = \begin{cases} -\lambda_3 - \frac{1}{2}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}^T A^{-1}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} & \text{si } \lambda_3 > 1/2, \\ -\infty \quad \text{si } \lambda_3 \leq 1/2. \end{cases}$$
3) On s'intéresse maintenant au problème dual
$$\sup_{\lambda_1\geq 0,\lambda_2\geq 0,\lambda_3\geq 0} d(\lambda).$$
a) Montrer qu'un maximum vérifie nécessairement $\lambda_1 = \lambda_2 = 0$ *puis que* $\lambda_3 = 1/2$.
b) En déduire que $\sup_{\lambda_1\geq 0,\lambda_2\geq 0,\lambda_3\geq 0} d(\lambda) = -1/2$.

Correction: On considère le problème d'optimisation
$$\inf x_1x_2 \quad \text{t.q. } x_1 \geq 0,\ x_2 \geq 0,\ x_1^2 + x_2^2 \leq 1.$$
Montrons qu'il y a un saut de dualité pour ce problème (dont l'origine vient du fait que $x \longmapsto x_1x_2$ n'est pas convexe).
I) **Résolution du problème primal**.
On montre aisément que la solution de ce problème est $(0, 0)$ et donc si $K := \{x \in \mathbb{R}^2\ ;\ x_1 \geq 0,\ x_2 \geq 0,\ x_1^2 + x_2^2 \leq 1\}$ et $f(x_1, x_2) := x_1x_2$, alors $\inf_{x\in K} f(x) = 0$.
II) **Résolution du problème dual**.
Soit $L : \mathbb{R}^2 \times \mathbb{R}^3_+ \to \mathbb{R}$ le lagrangien du problème. On écrira dans la suite $x = (x_1, x_2)$ $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, et $\tilde{\lambda} = (\lambda_1, \lambda_2)$.
Le Lagrangien s'écrit :
$$L(x_1, x_2, \lambda_1, \lambda_2, \lambda_3) := x_1x_2 - \lambda_1x_1 - \lambda_2x_2 + \lambda_3(x_1^2 + x_2^2 - 1).$$
Le problème dual est:
$$\sup_{\lambda\in\mathbb{R}^3_+} \inf_{x\in\mathbb{R}^2} L(x, \lambda).$$
1) Caractérisons d'abord les cas d'existence d'un minimum pour le problème : $d(\lambda) := \inf_{x\in\mathbb{R}^2} L(x, \lambda)$
<u>Cas 1</u>: $\lambda_3 > 1/2$. On a
$$L(x_1, x_2, \lambda) = (\lambda_3 - 1/2)(x_1^2 + x_2^2) - \lambda_1x_1 - \lambda_2x_2 - \lambda_3 + 1/2(x_1 + x_2)^2 \geq (\lambda_3 - 1/2)(x_1^2 + x_2^2) - \lambda_1x_1 - \lambda_2x_2 - \lambda_3$$

La fonction $x \longmapsto (\lambda_3 - 1/2)(x_1^2 + x_2^2) - \lambda_1 x_1 - \lambda_2 x_2 - \lambda_3$ étant coercice, L est minorée par une fonction coercive et donc elle admet un minimum.
Cas 2 $\lambda_3 < 1/2$. Grâce à l'égalité $L(x_1, x_2) = (\lambda_3 - 1/2)(x_1^2 + x_2^2) - \lambda_1 x_1 - \lambda_2 x_2 - \lambda_3 + 1/2(x_1 + x_2)^2$ on déduit que pour $x_2 = -x_1$:

$$L(x_1, -x_1, \lambda) = 2(\lambda_3 - 1/2)x_1^2 - \lambda_1 x_1 + \lambda_2 x_1 - \lambda_3.$$

donc L n'est pas minorée et $\inf_{x\in\mathbb{R}^2} L(x_1, x_2) = -\infty$.
Cas 3: $\lambda_3 = 1/2$. Si $\lambda_1 \neq \lambda_2$ alors pour $x_1 = -x_2$ on déduit que

$$L(x_1, -x_1, \lambda) = (\lambda_2 - \lambda_1)x_1 - \lambda_3,$$

fonction qui n'est pas minorée. Ainsi dans ce cas $\inf_{x\in\mathbb{R}^2} L(x_1, x_2, \lambda) = -\infty$. Supposons maintenant $\lambda_1 = \lambda_2$. Alors on a :

$$L(x, \lambda) = -\lambda_1(x_1 + x_2) - \lambda_3 + 1/2(x_1 + x_2)^2 = \varphi(x_1 + x_2)$$

où $\varphi(y) := 1/2y^2 - \lambda_1 y - \lambda_3$ est une parabole dans $\mathbb{R}$. Il est clair que φ admet un minimum sur $\mathbb{R}$ atteint en λ_1. Ainsi, $L(\cdot, \lambda)$ admet un minimum lorsque $x_1 + x_2 = \lambda_1$ (non unique).

2) Calculons maintenant $d(\lambda)$ dans les cas 1 et 3.
cas 1 : $\lambda_3 > 1/2$. En écrivant que $\nabla_x L(x, \lambda) = 0$ on déduit que

$$\begin{cases} x_2 - \lambda_1 + 2\lambda_3 x_1 = 0, \\ x_1 - \lambda_2 + 2\lambda_3 x_2 = 0 \end{cases} \tag{3.7.4}$$

ce qui permet de voir que $x = A^{-1}\tilde{\lambda}$ où $A = \begin{pmatrix} 2\lambda_3 & 1 \\ 1 & 2\lambda_3 \end{pmatrix}$. On vérifie que A est symétrique définie positive et donc il en est de même de A^{-1}. De plus en combinant les deux équations de (3.7.4) on a $\frac{1}{2}(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_3(x_1^2 + x_2^2) + x_1 x_2$. Ainsi dans ce cas $d(\lambda) = -\lambda_3 - \tilde{\lambda}^T A^{-1}\tilde{\lambda}$.
cas 3. On a en remplaçant $x_1 + x_2$ par λ_1 : $d(\lambda) = -\frac{1}{2}\lambda_1^2 - \lambda_3$. Finalement le résultat est le suivant :

$$d(\lambda) = \begin{cases} -\lambda_3 - \tilde{\lambda}^T A^{-1}\tilde{\lambda} & \text{si} \quad \lambda_3 > 1/2 \\ -\frac{1}{2}\lambda_1^2 - \lambda_3 & \text{si} \quad \lambda_3 = 1/2 \text{ et } \lambda_1 = \lambda_2 \\ -\infty & \text{si} \quad \lambda_3 = 1/2 \text{ et } \lambda_1 \neq \lambda_2 \\ -\infty & \text{si} \quad \lambda_3 < 1/2 \end{cases}$$

3) Une fois cette expression obtenue, la résolution du problème dual est immédiate. On a $\tilde{\lambda}^T A^{-1}\tilde{\lambda} \geq 0$ pour tout $\tilde{\lambda} \in \mathbb{R}_+^2$. Ainsi dans le premier cas $d(\lambda) < -\lambda_3 < -1/2$. Ainsi, le maximum de $d(\lambda)$ est $-1/2$ atteint pour $\lambda_3 = 1/2$ et $\lambda_1 = \lambda_2 = 0$. D'où le saut de dualité :

$$\sup_{\lambda\in\mathbb{R}_+^3} \inf_{x\in\mathbb{R}^2} L(x, \lambda) = -\frac{1}{2} < \inf_{x\in K} \sup_{\lambda\in\mathbb{R}_+^3} L(x, \lambda) = \inf_{x\in K} f(x) = 0.$$

3.7.5 Preuve de la condition nécessaire au second ordre par dualité

Soit $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq l$ et $h_j : \mathbb{R}^n \to \mathbb{R}$, $1 \leq j \leq m$ des fonctions de classe C^2. Soit K la contrainte fermée définie par :

$$K := \{x \in \mathbb{R}^n \ ; \ g_i(x) \leq 0, \ i \in I \ ; \ h_j(x) = 0 \ ; \ j \in J\},$$

avec $I = \{1, ..., l\}$ et $J = \{1, ..., m\}$ et soit le problème de minimisation :

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

Soit $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}_+^l \times \mathbb{R}^m \to \mathbb{R}$ le lagrangien associé à (P) défini par

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \sum_{i\in I} \lambda_i g_i(x) + \sum_{j\in J} \mu_j h_j(x).$$

Pour $x \in K$, on pose $I(x) := \{i \in I \ ; \ g_i(x) = 0\}$ et :

$$\Lambda(x) := \{(\lambda, \mu) \in \mathbb{R}_+^l \times \mathbb{R}^m \ ; \ \nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \text{ et } \lambda_i = 0 \ \forall i \notin I(x)\}.$$

Le cône critique est défini par:

$$C(x) := \{v \in T_K(x) \; ; \; \langle \nabla f(x), v\rangle = 0\}.$$

Le but est maintenant de montrer la condition nécessaire d'optimalité d'ordre 2 (la preuve est technique et peut être sautée en première lecture).

Théorème 3.15. *Si x est un minimum local de f sur K et si la contrainte est qualifiée au point x, alors:*

$$\forall v \in C(x) \quad \max_{(\lambda,\mu)\in\Lambda(x)} \left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(x,\lambda,\mu)v, v \right\rangle \geq 0.$$

Preuve. La condition $\Lambda(x) \neq \emptyset$ équivaut à l'existence d'un multiplicateur de Lagrange $(\lambda,\mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m$ pour le problème (P). Comme x est un minimum local et que la contrainte est qualifiée au point x, on a bien $\Lambda(x) \neq \emptyset$. Comme la contrainte est qualifiée on a $T_K(x) = \{v \in \mathbb{R}^n \; ; \; \forall i \in I(x) \;\; \langle \nabla g_i(x), v\rangle \leq 0 \; ; \; \forall j \in J \;\; \langle \nabla h_j(x), v\rangle = 0\}$ d'où

$$C(x) = \{v \in \mathbb{R}^n \; ; \; \forall i \in I(x) \;\; \langle \nabla g_i(x), v\rangle \leq 0 \; ; \; \forall j \in J \;\; \langle \nabla h_j(x), v\rangle = 0 \; ; \text{ et } \langle \nabla f(x), v\rangle = 0\}.$$

On définit maintenant le cône tangent d'ordre 2 au point x par:

$$\forall x \in \mathbb{R}^n \; \forall v \in T_K(x), \;\; T^2_K(x,v) := \{w \in \mathbb{R}^n \; ; \; \exists t_k \downarrow 0^+ \; \exists w_k \to v \text{ t.q. } x + t_k v + t_k^2 w_k \in K\}.$$

En faisant un développement de f à l'ordre 2 en t_k au voisinage du point x on a:

$$f(x + t_k v + t_k^2 w_k) = f(x) + t_k \langle \nabla f(x), v\rangle + t_k^2 \langle \nabla f(x), w\rangle + \frac{t_k^2}{2} \langle Hess_f(x)v, v\rangle + o(t_k^2).$$

Comme $\langle \nabla f(x), v\rangle = 0$ et $\frac{1}{t_k^2}(f(x + t_k v + t_k^2 w_k) - f(x)) \geq 0$ on déduit que:

$$\forall v \in C(x) \; \forall w \in T^2_K(x,v) \;\; \langle \nabla f(x), w\rangle + \frac{1}{2} \langle Hess_f(x)v, v\rangle \geq 0. \tag{3.7.5}$$

On pose $I^2(x,v) := \{i \in I, \; g_i(x) = 0 \text{ et } \langle \nabla g_i(x), v\rangle = 0\}$ et soit H l'ensemble défini par:

$$H := \left\{ w \in \mathbb{R}^n \; ; \langle \nabla g_i(x), w\rangle + \frac{1}{2} \langle Hess_{g_i}(x)v, v\rangle \leq 0 \; \forall i \in I^2(x,v) \text{ et } \langle \nabla h_j(x), w\rangle + \frac{1}{2} \langle Hess_{h_j}(x)v, v\rangle = 0 \; \forall j \in J \right\}.$$

De même, en faisant un développement de g_i à l'ordre 2 en t_k au voisinage du point x, on a pour $i \in I^2(x,v)$:

$$g_i(x + t_k v + t_k^2 w_k) = g_i(x) + t_k \langle \nabla g_i(x), v\rangle + t_k^2 \langle \nabla g_i(x), w\rangle + \frac{t_k^2}{2} \langle Hess_{g_i}(x)v, v\rangle + o(t_k^2) \leq 0.$$

car $g_i(x) = 0$ et $\langle \nabla g_i(x), v\rangle = 0$. D'où $t_k^2 \langle \nabla g_i(x), w\rangle + \frac{t_k^2}{2} \langle Hess_{g_i}(x)v, v\rangle + o(t_k^2) \leq 0$ et donc

$$\langle \nabla g_i(x), w\rangle + \frac{1}{2} \langle Hess_{g_i}(x)v, v\rangle \leq 0$$

à la limite. On procède de même pour les contraintes d'égalité ce qui montre que $T^2_K(x,v) \subset H$. On admettra dans la suite que $T^2_K(x,v) = H$.

Soit $F(w) := \langle \nabla f(x), w\rangle + \frac{1}{2} \langle Hess_f(x)v, v\rangle$. On considère le problème d'optimisation:

$$\min_{w \in T^2_K(x,v)} F(w) \tag{3.7.6}$$

On admet que ce problème admet une solution. Le lagrangien $L(w,\lambda,\mu)$ associé à ce problème s'écrit :

$$L(w,\lambda,\mu) := F(w) + \sum_{i\in I(x)} \lambda_i(\langle \nabla g_i(x), w\rangle + \frac{1}{2} \langle Hess_{g_i}(x)v, v\rangle) + \sum_{j\in J} \mu_j(\langle \nabla h_j(x), w\rangle + \frac{1}{2} \langle Hess_{h_j}(x)v, v\rangle).$$

Comme L est linéaire, si sa dérivée est nulle, alors L est constante et la constante qui reste (partie sans w dans l'expression de L) s'écrit:

$$\frac{1}{2} \langle Hess_f(x)v, v\rangle + \frac{1}{2} \sum_{i\in I(x)} \lambda_i \langle Hess_{g_i}(x)v, v\rangle + \frac{1}{2} \sum_{j\in J} \mu_j \langle Hess_{h_j}(x)v, v\rangle = \frac{1}{2} \left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(x,\lambda,\mu)v, v \right\rangle.$$

Comme $(\lambda, \mu) \in \Lambda(x)$ on trouve que L est constant en dérivant L par rapport à w et en appliquant KKT:

$$\nabla_w L(w, \lambda, \mu) = \nabla f(x) + \sum_{i \in I(x)} \lambda_i \nabla g_i(x) + \sum_{j \in J} \mu_j \nabla h_j(x) = 0.$$

Maintenant le vecteur $u := \nabla f(x) + \sum_{i \in I(x)} \lambda_i \nabla g_i(x) + \sum_{j \in J} \mu_j \nabla h_j(x)$ est non nul. On a :

$$\mathcal{L}(tu, \lambda, \mu) = t\|u\|^2 + \frac{1}{2} \langle Hess_f(x)v, v\rangle + \frac{1}{2} \sum_{i \in I(x)} \lambda_i \langle Hess_{g_i}(x)v, v\rangle + \frac{1}{2} \sum_{j \in J} \mu_j \langle Hess_{h_j}(x)v, v\rangle$$

pour $t \neq 0$ et $(\lambda, \mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m$. Donc en faisant $t \to -\infty$ $\mathcal{L}(\cdot, \lambda, \mu)$ n'admet pas de minimum sur $\mathbb{R}^n$ si $(\lambda, \mu) \notin \Lambda(x)$. Finalement on a donc montré que:

$$\inf_{w \in \mathbb{R}^n} \mathcal{L}(w, \lambda, \mu) = \begin{cases} \frac{1}{2} \left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda, \mu)v, v \right\rangle, & \text{si } (\lambda, \mu) \in \Lambda(x) \\ -\infty, & \text{si } (\lambda, \mu) \notin \Lambda(x). \end{cases}$$

On notera dans la suite $\delta(\lambda, \mu) := \inf_{w \in \mathbb{R}^n} \mathcal{L}(w, \lambda, \mu)$, $(\lambda, \mu) \in \mathbb{R}^l \times \mathbb{R}^m$. Montrons que

$$\max_{\lambda \geq 0, \mu \in \mathbb{R}^m} \delta(\lambda, \mu) = \min_{w \in T^2_K(x,v)} F(w).$$

Le problème primal (2) est un problème linéaire en w sous contraintes linéaires d'inégalité et d'égalité. Ainsi, il s'agit d'un problème convexe (c.a.d. le coût et les contraintes sont convexes). Par le théorème de cours sur la dualité dans le cas convexe on sait qu'il n'y a pas de sauf de dualité (i.e. les problèmes primal et dual ont même valeur):

$$\min_{w \in T^2_K(x,v)} F(w) = \max_{(\lambda,\mu) \in \mathbb{R}^l_+ \times \mathbb{R}^m} \inf_{w \in \mathbb{R}^n} \mathcal{L}(w, \lambda, \mu) = \max_{(\lambda,\mu) \in \Lambda(x)} \left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda, \mu)v, v \right\rangle$$

Soit maintenant x un minimum local de f sur K et $v \in C(x)$. Alors, en utilisant la question 3), on déduit que $F(w) \geq 0$ pour tout $w \in T^2_K(x, v)$ et donc $\min_{w \in T^2_K(x,v)} F(w) \geq 0$ d'où le résultat par l'égalité ci-dessus. Par ce qui précède on a que $F(w) \geq 0$ pour tout $w \in T^2_K(x, v)$ et donc $\min_{w \in T^2_K(x,v)} F(w) \geq 0$ d'où le résultat par l'égalité ci-dessus.

□

Chapitre 4

Algorithmes numériques

4.1 Algorithmes de descente de gradient

Dans l'optique de minimiser une fonction objectif $f : \mathbb{R}^n \to \mathbb{R}$ et de résoudre un problème d'optimisation:

$$\min_{x \in C} f(x),$$

on étudie des algorithmes itératifs du type

$$x_{k+1} = \Phi(x_k, \nabla f(x_k)),$$

avec si possible $f(x_{k+1}) < f(x_k)$ pour tout $k \in \mathbb{N}$.

Définition 4.1. *Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ une application de classe C^1. Soit $x \in U$ tel que $\nabla f(x) \neq 0$. Une direction de descente pour f est un vecteur $d \in \mathbb{R}^n$ tel que $d \neq 0$ et:*

$$\exists \bar{t} > 0 \quad \forall t \in]0, \bar{t}[, \quad f(x + td) < f(x).$$

Remarque 4.1. *(i) Comme f est de classe C^1, un développement montre que:*

$$f(x + td) - f(x) = t \langle \nabla f(x), d \rangle + o(t),$$

d'où nécessairement $\langle \nabla f(x), d \rangle \leq 0$. De plus, si d vérifie $\langle \nabla f(x), d \rangle < 0$ alors d est une direction de descente. Il semble donc que $d := -\frac{\nabla f(x)}{\|\nabla f(x)\|}$ soit un bon choix de direction de descente à t fixé.
(ii) Le choix de direction d ci-dessus est motivé par la remarque suivante sur le système de type "gradient" (méthode de la plus grande pente). Considérons le problème de Cauchy:

$$\dot{x}(t) = -\nabla f(x(t)), \quad t > 0, \quad x(0) = x_0 \in \mathbb{R}^n.$$

On a $\frac{d}{dt} f(x(t)) = -\|\nabla f(x(t))\|^2$ et donc l'application $t \mapsto f(x(t))$ décroit strictement (sauf aux points critiques de f). La discrétisation en pas de temps régulier ($h := t_{k+1} - t_k$) de cette équation différentielle conduit donc à l'égalité

$$x_{x+1} - x_k = -h \nabla f(x_k), \quad k > 0.$$

On va considérer dans la suite des algorithmes du type:

$$x_{k+1} = x_k + t_k d_k, \quad x_0 \in \mathbb{R}^n.$$

Le réel t_k choisi à chaque étape s'appelle le pas de descente. Comme le montre la figure 4.1, le choix des paramètres (pas de descente, direction de descente) est crucial. La figure 4.1 est réalisée avec la fonction $f(x) = \frac{1}{2}x^2$, $x_0 = 2$, $s_k = 2 + \frac{3}{2^{k+1}}$ et $d_k = -\text{sign}(x_k)$. Ainsi, on a bien un algorithme de descente mais qui ne converge pas car $x_k = (-1)^k(1 + 1/2^k)$ pour tout $k \in \mathbb{N}$. On voit sur cet exemple que le choix du pas est crucial (i.e. ici il ne doit pas être trop grand). Si le pas est trop petit (par exemple $s_k = \frac{1}{2^k}$ dans l'exemple précédent), la méthode de descente peut ne pas converger vers un minimum local de la fonction à minimiser. On est donc amené à considérer différentes règles pour le choix du pas et de la direction de descente. Une règle (pas forcément bonne en pratique) est la recherche exacte.

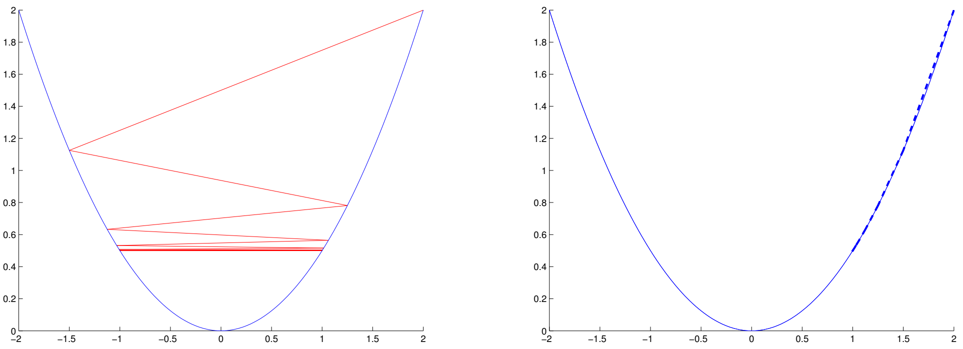


Figure 4.1: Figure de gauche : pas trop grand. Figure de droite : pas trop petit. Dans les deux cas, l'algorithme du gradient à pas variable ne converge pas.

Définition 4.2. *La méthode de descente de gradient avec recherche exacte consiste à choisir à chaque étape k le pas h_k par:*

$$h_k \in \arg\min_{t\in\mathbb{R}} f(x_k + td_k),$$

où $x_k \in \mathbb{R}^n$ est donné et $d_k \in \mathbb{R}^n$ est la direction de descente donnée.

4.1.1 Règle d'Armijo

Par rapport à l'exemple précédent, la règle d'Armijo permet d'éviter des pas trop grands. Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1 et (P) le problème de minimisation $\inf_{x\in\mathbb{R}^n} f(x)$. On se donne des paramètres $s > 0$, $\sigma \in]0,1[$, $\beta \in]0,1[$ et $\varepsilon > 0$. On considère l'algorithme de descente basé sur la règle *d'Armijo*:

- Soit $x_0 \in \mathbb{R}^n$ et $d_0 \in \mathbb{R}^n$ tel que $\|d_0\| = 1$ (initialisation). Pour $k = 1, ..., N_{iter}$ faire les étapes suivantes:
- Prendre $d_k \in \mathcal{S}^{n-1}$ tel que $-\frac{\langle \nabla f(x_k), d_k\rangle}{\|\nabla f(x_k)\|} \in [\varepsilon, 1]$ et tel que $\|d_k\| = 1$.
- pour $i = 0, 1, ...$ soit i_k le premier indice $i \geq 0$ pour lequel on a $f(x_k) - f(x_k + \beta^i s d_k) \geq -\sigma\beta^i s \langle \nabla f(x_k), d_k\rangle$. Si $i = 0$ on pose $\alpha_k = s$ et si $s \geq 1$ on pose $\alpha_k = s\beta^{i_k}$.
- Faire $x_{k+1} = x_k + \alpha_k d_k$ avec $\alpha_k = s\beta^{i_k}$

Théorème 4.1. *Si x^* est une valeur d'adhérence de la suite (x_k) générée, alors $\nabla f(x^*) = 0$.*

Preuve. Première étape. Montrons d'abord que l'étape 3 est possible. Sinon, pour tout i on a

$$f(x_k) - f(x_k + \beta^i s d_k) < -\sigma\beta^i s \langle \nabla f(x_k), d_k\rangle$$

Comme f est C^1, on a le développement:

$$f(x_k + td_k) = f(x_k) + t\langle \nabla f(x_k), d_k\rangle + o(t).$$

D'où il existe une suite (ε_i) de limite 0 telle que:

$$f(x_k + \beta^i s d_k) - f(x_k) = \beta^i s \langle \nabla f(x_k), d_k\rangle + \beta^i \varepsilon_i > \sigma\beta^i s \langle \nabla f(x_k), d_k\rangle.$$

et donc en faisant $i \to +\infty$ on obtient $0 \geq (\sigma - 1)\langle \nabla f(x_k), d_k\rangle$ et une contradiction.

Deuxième étape. On suppose maintenant que $\nabla f(x^*) \neq 0$. Quitte à extraire, on peut supposer qu'il existe $d^* \in \mathcal{S}^{n-1}$ tel que $d_k \to d^*$ et $x_k \to x^*$. Par l'étape 2 sur le choix de d_k, on obtient à la limite:

$$\frac{\langle \nabla f(x^*), d^*\rangle}{\|\nabla f(x^*)\|} < 0.$$

De plus, par l'inégalité de Cauchy-Schwarz on a :

$$f(x_k) - f(x_k + \beta^{i_k} s d_k) \geq -\sigma \beta^{i_k} s \langle \nabla f(x_k), d_k \rangle \geq s\sigma \beta^{i_k} \|\nabla f(x_k)\|,$$

ce qui permet de conclure que β^{i_k} tend vers 0 lorsque k tend vers l'infini (en effet, $\nabla f(x^*) \neq 0$ et par continuité $\|\nabla f(x_k)\| \to \|\nabla f(x^*)\|$). On remarque que par définition de la suite on a $\|x_{k+1} - x_k\| = s\beta^{i_k}$ et donc $\|x_{k+1} - x_k\| \to 0$.

Pour conclure, utilisons la définition de i_k. On a:

$$\begin{cases} f(x_k) - f(x_k + \beta^{i_k} s d_k) \geq -\sigma \beta^{i_k} s \langle \nabla f(x_k), d_k \rangle , \\ f(x_k) - f(x_k + \beta^{i_k - 1} s d_k) < -\sigma \beta^{i_k - 1} s \langle \nabla f(x_k), d_k \rangle \end{cases}$$

En utilisant le théorème des accroissements finis et la seconde égalité ci-dessus, il existe $z_k \in]x_k, x_k + \beta^{i_k - 1} s d_k[$ tel que:

$$-\frac{\alpha_k}{\beta} \langle \nabla f(z_k), d_k \rangle < -\sigma \frac{\alpha_k}{\beta} \langle \nabla f(x_k), d_k \rangle .$$

Ceci permet de conclure que $\langle \nabla f(x^*), d^* \rangle \geq 0$ (en passant à la limite) d'où une contradiction. $\square$

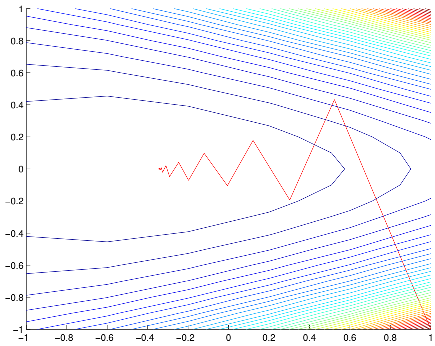


Figure 4.2: Règle d'Armijo avec $f(x_1, x_2) = e^{x_1 + 3x_2 - 0.1} + e^{x_1 - 3x_2 - 0.1} + e^{-x_1 - 0.1}$ et avec comme direction de descente $d_k = -\nabla f(x_k)$. Paramètres : $\alpha = 0.1$ et $\beta = 0.7$.

Exercice 4.1. *On remplace la règle d'Armijo par la règle suivante dite règle de minimisation limitée. On choisit t_k tel que:*

$$\exists s > 0 \quad \forall k \geq 0 \quad \forall\ 0 \leq t \leq s, \quad f(x_k + t_k d_k) < f(x_k + t d_k).$$

Montrer le même résultat de convergence que pour la règle d'Armijo.

4.1.2 Règle de Wolfe

Pour éviter que les pas soient trop petits, on utilise la règle de Wolfe. Soit $0 < c_1 < c_2 < 1$. Etant donné $x_k \in \mathbb{R}^n$ et une direction de descente $d_k \in \mathbb{R}^n$, on choisit $t_k > 0$ de sorte que l'algorithme de descente $x_{k+1} = x_k + t_k d_k$ s'écrive:

$$\begin{cases} f(x_k + t_k d_k) \leq f(x_k) + c_1 t_k \langle \nabla f(x_k), d_k \rangle , \\ \langle \nabla f(x_k + t_k d_k), d_k \rangle \geq c_2 \langle \nabla f(x_k), d_k \rangle . \end{cases}$$

La deuxième inégalité empêche au pas d'être trop petit car $c_2 < 1$ (faire un développement limité de cette expression en fonction de $t = t_k$ à k fixé) et la première au pas d'être trop grand. En pratique, on peut prendre $c_2 = 0.99$ et $c_1 = 0.0001$.

On peut montrer sous certaines hypothèses la convergence de l'algorithme de descente avec la règle de Wolfe. On suppose donné le réel $\varepsilon > 0$ et la suite (d_k) tels que:

$$\forall k \in \mathbb{N}, \ \cos\theta_k := \frac{-\langle \nabla f(x_k), d_k \rangle}{\|\nabla f(x_k)\|} \in (\varepsilon, 1], \quad \|d_k\| = 1. \tag{4.1.1}$$

Soit (x_k) la suite générée par l'algorithme de descente i.e. $x_{k+1} = x_k + t_k d_k$ où t_k vérifie la règle de Wolfe.

Lemme 4.1. *(Zoutendijk). Soit f une fonction minorée sur $M := \{x\ ;\ f(x) \leq f(x_0)\}$. On suppose f de classe C^1 et que:*

$$\exists L \geq 0\ \forall x, y \in M, \quad \|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|.$$

Alors $\sum \cos^2 \theta_k \|\nabla f(x_k)\|^2 < +\infty$.

Preuve. Par la règle de Wolfe on a : $(c_2 - 1)\langle \nabla f(x_k), d_k\rangle \leq \langle \nabla f(x_{k+1}) - \nabla f(x_k), d_k\rangle \leq t_k L$. On déduit que:

$$t_k \geq \frac{1 - c_2}{L}\langle -\nabla f(x_k), d_k\rangle.$$

Par la première condition de Wolfe, il vient donc:

$$f(x_k) - f(x_{k+1}) \geq \frac{c_1(1-c_2)}{L}|\langle \nabla f(x_k), d_k\rangle|^2 = \frac{c_1(1-c_2)}{L}\cos^2\theta_k\|\nabla f(x_k)\|^2.$$

D'où par somme télescopique: pour $q \in \mathbb{N}^*$:

$$f(x_0) - \inf_M f \geq f(x_0) - f(x_q) \geq \frac{c_1(1-c_2)}{L}\sum_{k=0}^{q}\cos^2\theta_k\|\nabla f(x_k)\|^2,$$

d'où le résultat. □

Corollaire 4.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1. Tout point x^* valeur d'adhérence de la suite (x_k) vérifie $\nabla f(x^*) = 0$.*

Preuve. Quite à extraire une sous-suite on suppose que la suite (x_k) converge vers x^*. Alors, la suite (x_k) est bornée et on peut appliquer le lemme précédent. □

4.1.3 Conditionnement pour la méthode de gradient à pas optimal

On considère l'algorithme du gradient à pas optimal pour la fonction:

$$x \mapsto f(x) := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle,$$

où A est une matrice symétrique définie positive de taille n, et $b \in \mathbb{R}^n$. On rappelle que la méthode du gradient à pas optimal consiste à chaque étape $x_{x+1} = x_k + \alpha_k d_k$ à choisir $\alpha \in \mathbb{R}$ de sorte à minimiser la fonction $\rho \mapsto \varphi(\rho) := f(x_k + \rho d_k)$.

Première étape. A chaque étape, la méthode à pas optimal s'écrit:

$$x_{k+1} = x_k - \alpha_k r_k, \quad \alpha_k = \frac{\|r_k\|^2}{\langle Ar_k, r_k\rangle}, \quad r_k = Ax_k - b.$$

Seconde étape. Calculons l'erreur $\|x_k - x^*\|_A$ où $\|x\|_A := \sqrt{\langle Ax, x\rangle}$. Soit $e_k := x_k - x^*$ où $Ax^* = b$.

$$\begin{aligned}
\|e_{k+1}\|_A^2 &= \langle A(e_k - \alpha_k r_k), e_k - \alpha_k r_k\rangle \\
&= \langle Ae_k, e_k\rangle - 2\frac{\langle Ar_k, e_k\rangle\langle r_k, r_k\rangle}{\langle Ar_k, r_k\rangle} + \langle Ar_k, r_k\rangle\frac{\langle r_k, r_k\rangle^2}{\langle Ar_k, r_k\rangle^2} \\
&= \langle Ae_k, e_k\rangle - 2\frac{\langle r_k, r_k\rangle^2}{\langle Ar_k, r_k\rangle} + \frac{\langle r_k, r_k\rangle^2}{\langle Ar_k, r_k\rangle} = \langle Ae_k, e_k\rangle - \frac{\langle r_k, r_k\rangle^2}{\langle Ar_k, r_k\rangle} \\
&= \langle Ae_k, e_k\rangle\left(1 - \frac{\|r_k\|^4}{\langle Ar_k, r_k\rangle\langle A^{-1}r_k, r_k\rangle}\right)
\end{aligned}$$

en utilisant que $\langle Ar_k, e_k\rangle = \langle r_k, Ae_k\rangle = \langle r_k, r_k\rangle$ et $Ae_k = r_k \iff e_k = A^{-1}r_k$.

Troisième étape. On admet l'inégalité de Kantorovitch: $\min_{x\in\mathbb{R}^n}\frac{\|x\|^4}{\langle Ax,x\rangle\langle A^{-1}x,x\rangle} = \frac{4\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2}$ avec λ_1, resp. λ_n la plus petite, resp. la plus grande valeur propre strictement positive de A. D'où:

$$\|e_{k+1}\|_A \leq \left(\frac{\gamma - 1}{\gamma + 1}\right)\|e_k\|_A.$$

Définition 4.3. *Pour une matrice symétrique définie positive A, le conditionnement est le réel $\gamma := \frac{\lambda_n}{\lambda_1}$ le rapport entre la plus grande valeur propre λ_n et la plus petite valeur propre λ_1.*

Ainsi, plus le conditionnement est grand, plus la méthode ci-dessus converge lentement:

Remarque 4.2. *L'analyse pour la fonction f quadratique ci-dessus est importante car étant donnée une fonction $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^2, si x^* est un point critique, on a au voisinage de 0:*

$$f(x) - f(x^*) = \langle H(x - x^*), (x - x^*)\rangle + o(\|x - x^*\|^2).$$

Ainsi, le comportement de f au voisinage de 0 est "proche" de la forme quadratique associée à H. Si on souhaite minimiser f, et si H est mal conditionné, il faudra alors raffiner la méthode pour améliorer la vitesse de convergence.

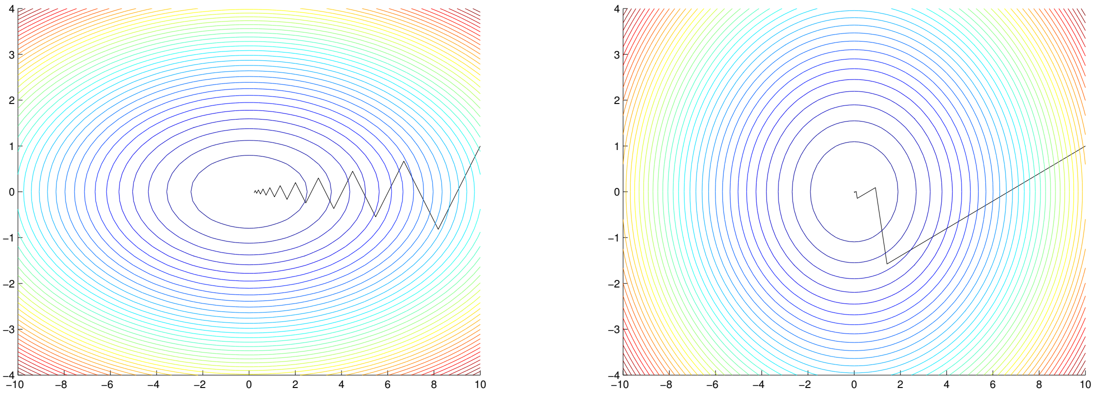


Figure 4.3: Exemples de minimisation exacte avec la fonctionnelle quadratique $f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2)$ avec $\gamma = 3$ (droite) et $\gamma = 10$ (gauche). Dans le second cas, la matrice est moins bien conditionnée et la méthode plus lente.

Exercice 4.2. *Démontrer l'inégalité de Kantorovitch. Pour cela, montrer que l'on peut supposer que A est diagonale et $\lambda_1\lambda_n = 1$. Montrer que l'inégalité équivaut à*

$$\left(\sum_{i=1}^n t_i\lambda_i\right)\left(\sum_{i=1}^n \frac{t_i}{\lambda_i}\right) \leq \frac{1}{4}\left(m + \frac{1}{m}\right)^2.$$

où $0 \leq t_i < 1$ et $\sum_{i=1}^n t_i = 1$. Utiliser ensuite que $ab \leq \frac{1}{2}(a+b)^2$ puis que la maximum de la fonction $\lambda_i \mapsto \lambda_i + \frac{1}{\lambda_i}$ sur $[m, \frac{1}{m}]$ est atteint aux extrémités.

4.1.4 Gradient projeté

Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1. On rappelle qu'une *direction de descente* est un vecteur $d \in \mathbb{R}^n$, $d \neq 0$ tel qu'il existe $\bar{t} > 0$ tel que $f(x + td) < f(x)$ pour tout $t \in]0, \bar{t}[$. On s'intéresse au problème

$$\min_{x \in C} f(x),$$

où $C \subset \mathbb{R}^n$ est un sous-ensemble convexe fermé non vide de $\mathbb{R}^n$. L'algorithme de gradient projeté fournit un algorithme permettant de résoudre numériquement ce problème de minimisation sous contraintes. On garde les mêmes notations que dans la règle d'Armijo. L'algorithme du gradient projeté s'écrit ainsi. Soit $\varepsilon > 0$.

- Choisir $x_0 \in C$. Pour $k = 1, ..., N_{iter}$, faire les étapes suivantes:
- Calculer $\bar{x}_k = \Pi_C(x_k - s\nabla f(x_k))$.

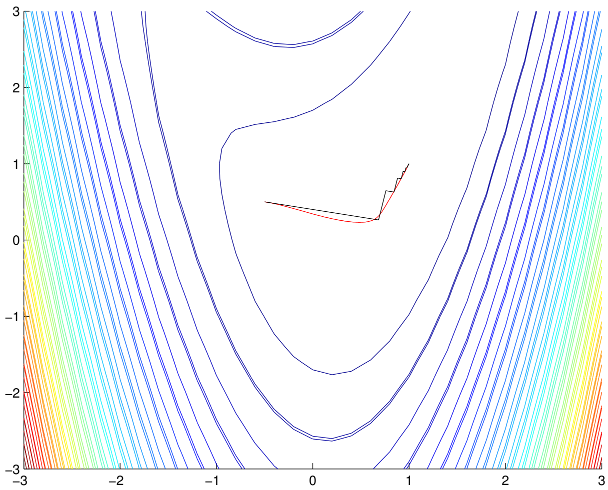


Figure 4.4: Comparaison entre descente de gradient à pas optimal (noir) et à pas fixe (rouge) pour la fonction banane $f(x_1, x_2) = (x_1 - 1)^2 + p(x_1 - x_2^2)^2$ avec $p = 1$ et le point initial $(-\frac{1}{2}, \frac{1}{2})$. Pour la descente à pas optimal : 40 itérations et pour la descente à pas fixe $\alpha = 0.01$

- Calculer le premier entier $m = m_k \in \mathbb{N}$ tel que $f(x_k) - f(x_k + \beta^m(\bar{x}_k - x_k)) \geq -\sigma\beta^m \langle \nabla f(x_k), \bar{x}_k - x_k\rangle$.
- Poser $x_{k+1} = x_k + \beta^m(\bar{x}_k - x_k)$.
- Si $\|x_{k+1} - x_k\| \leq \varepsilon$, alors arrêter l'algorithme.

On peut montrer le résultat suivant (voir preuve dans le chapitre exercice).

Théorème 4.2. *Soit x^* un point adhérent à la suite (x_k). Alors x^* vérifie $\langle \nabla f(x^*), x - x^*\rangle \geq 0$ pour tout $x \in C$.*

4.1.5 Algorithme d'Uzawa

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe C^1, $C \in M_{m,n}(\mathbb{R})$ et $d \in \mathbb{R}^n$. Soit le problème de minimisation

$$\inf_{Cx=d} f(x). \tag{P}$$

L'algorithme d'Uzawa correspond à l'algorithme du gradient appliqué au problème dual associé au problème (P).

- Considérer $\lambda_0 \in \mathbb{R}^m$ pour l'initialisation et soit $\rho > 0$.
- A l'étape k, on appelle x_k une solution de $\inf_{x\in\mathbb{R}^n} f(x) + \langle \lambda_k, Cx - d\rangle$.
- On pose $\lambda_{k+1} = \lambda_k + \rho(Cx_k - d)$.

Le but est de montrer le théorème suivant.

Théorème 4.3. *On suppose que f est de classe C^2 et elliptique de constante $\alpha > 0$ i.e. pour tout $x \in \mathbb{R}^n$*

$$\left\langle D^2 f(x)v, v\right\rangle \geq \alpha\|v\|^2, \quad \forall v \in \mathbb{R}^n.$$

Soit $\rho \in]0, \frac{2\alpha}{\|C\|^2}[$. Alors la suite (x_k) générée par l'algorithme converge vers l'unique minimum du problème (P).

Preuve. Comme f est coercive, le problème admet une solution x^* et les conditions d'optimalité donnent:

$$Cx^* = d, \quad \nabla f(x^*) + C^T\lambda^* = 0.$$

Les conditions d'optimalité donnent $\nabla f(x_k) + C^T\lambda_k = 0$. D'où:

$$\nabla f(x_k) - \nabla f(x^*) + C^T(\lambda_k - \lambda^*) = 0.$$

Par définition de λ_{k+1} on déduit: $\lambda_{k+1} - \lambda^* = \lambda_k - \lambda^* + \rho C(x_k - x^*)$. Ainsi, on déduit que:

$$\begin{aligned}
\|\lambda_{k+1} - \lambda^*\|^2 &= \|\lambda_k - \lambda^*\|^2 + 2\rho\,\langle \lambda_k - \lambda^*, C(x_k - x^*)\rangle + \rho^2\|C(x_k - x^*)\|^2 \\
&= \|\lambda_k - \lambda^*\|^2 + 2\rho\left\langle C^T(\lambda_k - \lambda^*), x_k - x^*\right\rangle + \rho^2\|C(x_k - x^*)\|^2 \\
&= \|\lambda_k - \lambda^*\|^2 - 2\rho\,\langle \nabla f(x_k) - \nabla f(x^*), x_k - x^*\rangle + \rho^2\|C(x_k - x^*)\|^2 \\
&\leq \|\lambda_k - \lambda^*\|^2 - 2\rho\alpha\|x_k - x^*\|^2 + \rho^2\|C\|\|x_k - x^*\|^2
\end{aligned}$$

L'hypothèse sur α implique donc la décroissance de la suite $\|\lambda_k - \lambda^*\|^2$ qui converge donc. Or:

$$\rho(2\alpha - \rho\|C\|^2)\|x_k - x^*\|^2 \leq \|\lambda_{k+1} - \lambda^*\|^2 - \|\lambda_k - \lambda^*\|^2.$$

D'où la convergence de x_k vers x^*. □

Remarque 4.3. *(i) Si C est de rang m, alors (λ_k) converge vers λ^*. En effet, la démonstration précédente entraine que (λ_k) est une suite bornée. Soit donc une sous-suite λ_{k_j} convergeant vers $\tilde{\lambda} \in \mathbb{R}^m$ lorsque $j \to +\infty$. A la limite, on obtient $\nabla f(x^*) + C^T\tilde{\lambda} = 0$. Or C est de rang m, donc surjective C^T est injective. Donc $\tilde{\lambda}$ est l'unique solution de l'équation $\nabla f(x^*) + C^T\lambda = 0$ (dont λ^* est déjà solution). Ainsi, $\lambda_{k_j} \to \lambda^*$. Par un résultat classique sur les sous-suites, on a bien $\lambda_k \to \lambda^*$.*
(ii) Dans la preuve ci-dessus, considérons la fonction $d(\lambda_0) := \min_{x\in\mathbb{R}^n} f(x) + \langle \lambda_0, Cx - d\rangle$. Par le théorème de Danskin, on peut montrer que cette fonction est dérivable en un point λ_0.Soit x_0 un minimum de cette fonction. On a pour $h > 0$ et $v \in \mathbb{R}^n$:

$$d(\lambda_0 + hv) \leq f(x_0) + \langle \lambda_0, Cx_0 - d\rangle + hv^T(Cx_0 - d) = d(\lambda_0) + hv^T(Cx_0 - d).$$

D'où $\langle \nabla d(\lambda_0), v\rangle \geq hv^T(Cx_0 - d)$. De même on a l'inégalité inverse en considérant $-h < 0$, d'où $\nabla d(\lambda_0) = Cx_0 - d$. Ainsi, l'algorithme d'Uzawa correpond à faire l'opposé de l'algorithme du gradient sur le multiplicateur λ_k

Avec les mêmes notations que précédemment, considérons maintenant le problème:

$$\min_{Cx\leq d} f(x).$$

Soit $P : \mathbb{R}^m \to \mathbb{R}^m$ la projection sur l'orthant positif $\mathbb{R}^m_+$ i.e. si $\lambda = (\lambda_i)_{1\leq i\leq m}$:

$$P(\lambda) = (\max(\lambda_i, 0))_{1\leq i\leq m}.$$

Par l'analyse du problème dual, on introduit l'algorithme d'Uzawa:

- Considérer $\lambda_0 \in \mathbb{R}^m_+$ pour l'initialisation et soit $\rho > 0$.
- A l'étape k, on appelle x_k une solution de $\inf_{x\in\mathbb{R}^n} f(x) + \langle \lambda_k, Cx - d\rangle$.
- On pose $\lambda_{k+1} = P(\lambda_k + \rho(Cx_k - d))$.

On a alors le même résultat de convergence :

Théorème 4.4. *On suppose que f est de classe C^2 et elliptique de constante $\alpha > 0$. Soit $\rho \in]0, \frac{2\alpha}{\|C\|^2}[$. Alors la suite (x_k) générée par l'algorithme converge vers l'unique minimum du problème* (P).

4.2 Méthode du gradient conjugué

Dans l'optique de résoudre un système linéaire $Hx + b = 0$, on considère le problème de minimisation:

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

où $x \in \mathbb{R}^n$, $H \in \mathcal{S}_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$. La méthode du gradient conjugué est une méthode itérative qui converge vers la solution exacte en au plus n itérations. De plus, elle s'étend dans un cadre plus général où f est quelconque (Gâteaux-différentiable). La méthode a des similarités avec le processus d'orthonormalisation de Gram-Schmidt.

Définition 4.4. *On dit que deux directions non nulles $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^n$ sont H-conjuguées si $\langle x, Hy \rangle = 0$.*

Lemme 4.2. *On considère une famille $d_0, ..., d_{n-1}$ de n vecteurs H-conjugués. Soit $x_0 \in \mathbb{R}^n$ donné et pour $0 \leq k \leq n-1$, on définit*

$$x_{k+1} = x_k + \lambda_k d_k,$$

où λ_k minimise f dans la direction d_k. Alors on a $\lambda_k = -\frac{\langle d_k, Hx_k + b \rangle}{\langle d_k, Hd_k \rangle}$.

Preuve. On a $f(x_k + td_k) = f(x_k) + t \langle d_k, Hx_k + b \rangle + \frac{1}{2} \langle d_k, Hd_k \rangle t^2$ d'où le résultat ($\frac{d}{dt} f(x_k + td_k) = 0$). □

Théorème 4.5. *Etant donné $x_0 \in \mathbb{R}^n$, la suite générée par:*

$$x_{k+1} = x_k + \lambda_k d_k \quad \lambda_k = -\frac{\langle d_k, Hx_k + b \rangle}{\langle d_k, Hd_k \rangle} \tag{4.2.1}$$

converge en au plus n itérations vers l'unique solution x^* du système linéaire $Hx^* + b = 0$.

Preuve. La famille $\{d_0, ..., d_{n-1}\}$ composée de n directions H-conjuguées forme une base de $\mathbb{R}^n$. Donc il existe un unique $(\beta_0, ..., \beta_{n-1}) \in \mathbb{R}^n$ tel que $x^* - x_0 = \sum_{0 \leq i \leq n-1} \beta_i d_i$ avec $\beta_i = \frac{\langle d_i, H(x^* - x_0) \rangle}{\langle d_i, Hd_i \rangle}$ (utiliser les directions H-conjuguées). Par l'algorithme on a $x_k - x_0 = \sum_{0 \leq i \leq k-1} \lambda_i d_i$ où $\lambda_i \in \mathbb{R}$. On déduit donc $\langle d_k, H(x_k - x_0) \rangle = 0$. Il vient alors:

$$\langle d_k, H(x^* - x_0) \rangle = \langle d_k, H(x_k - x_0) \rangle + \langle d_k, H(x^* - x_k) \rangle = \langle d_k, -b - Hx_k \rangle,$$

en utilisant $Hx^* + b = 0$. On déduit que $0 \leq k \leq n-1 \Rightarrow \beta_k = -\frac{\langle d_k, Hx_k + b \rangle}{\langle d_k, Hd_k \rangle} = \lambda_k$. Ainsi, au bout d'au plus n itérations on a $x_n = x^* = x_0 + \sum_{0 \leq i \leq n-1} \beta_i d_i = \sum_{0 \leq i \leq n-1} \lambda_i d_i$. □

Description de l'algorithme. A l'inverse d'une descente de gradient, l'algorithme utilise comme direction de descente les directions H-conjuguées.

- Etape 0 : choisir $x_0 \in \mathbb{R}^n$; poser $g_0 = Hx_0 + b$ et $d_0 = -g_0$.
- Pour $k \in \{0, ..., n-1\}$ faire si $g_k = 0$ STOP (on a convergé $x_k = x^*$ i.e. critère d'arrêt) ; sinon faire:

$$\begin{aligned}
\lambda_k &= -\tfrac{\langle d_k, g_k \rangle}{\langle d_k, Hd_k \rangle} \\
x_{k+1} &= x_k + \lambda_k d_k \\
g_{k+1} &= Hx_{k+1} + b \\
\beta_{k+1} &= \tfrac{\langle g_{k+1}, Hd_k \rangle}{\langle d_k, Hd_k \rangle} \\
d_{k+1} &= -g_{k+1} + \beta_{k+1} d_k
\end{aligned}$$

On peut vérifier par récurrence qu'au cours de l'algorithme les directions $(d_i)_{0 \leq i \leq n-1}$ sont bien H-conjuguées (vérifier d'abord sur d_2 et d_0).

Remarque 4.4. *(i) La consommation mémoire de cet algorithme est peu élevée : il faut stocker x_k, d_k, Hd_k, g_k et les scalaires λ_k, β_{k+1}.*
(ii) Pour des systèmes creux, l'algorithme est efficace (il suffit de savoir appliquer le produit de matrice Hx_k).
(iii) Si la matrice H a $r < n$ valeurs propres strictement positives, alors l'algorithme est encore plus rapide, il s'arrête en au plus r étapes.

D'une part, en notant que $\nabla f(x) = Hx + b$ et que $g_k = \nabla f(x_k)$ l'algorithme du gradient conjugué se réécrit:

- Etape 0 : choisir $x_0 \in \mathbb{R}^n$; poser $g_0 = \nabla f(x_0)$ et $d_0 = -g_0$.
- Pour $k \in \{0, ..., n-1\}$ faire si $g_k = 0$ STOP (on a convergé $x_k = x^*$ i.e. critère d'arrêt) ; sinon faire:

$$\begin{aligned}
\lambda_k &= -\tfrac{\langle d_k, \nabla f(x_k) \rangle}{\langle d_k, Hd_k \rangle} \\
x_{k+1} &= x_k + \lambda_k d_k \\
\beta_{k+1} &= \tfrac{\langle \nabla f(x_{k+1}), Hd_k \rangle}{\langle d_k, Hd_k \rangle} \\
d_{k+1} &= -\nabla f(x_{k+1}) + \beta_{k+1} d_k.
\end{aligned}$$

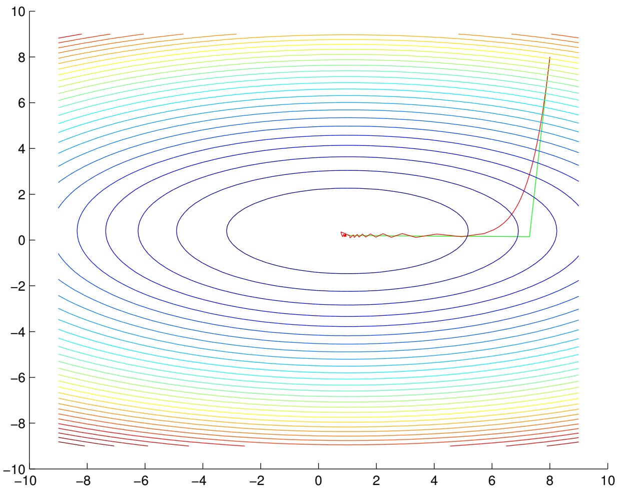


Figure 4.5: Comparaison entre la méthode du gradient conjugué en quelques itérations (vert) et la descente de gradient à pas optimal en 100 itérations (rouge) avec `A=[1,0;0,5]`, `B=[1;2]` (voir exercice de TP).

D'autre part on a le lemme suivant (formule de Fletcher-Reeves, 1964).

Lemme 4.3. *On a* $\beta_{k+1} = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}$, $0 \le k \le n-1$.

Preuve. Rappelons que $g_k = \nabla f(x_k)$. Notons les points suivants:

- La fonction $t \longmapsto f(x_k + td_k)$ est minimisée sur $\mathbb{R}$ par λ_k, aussi on a $\langle \nabla f(x_k + \lambda_k d_k), d_k \rangle = 0$ ce qui s'écrit $\langle \nabla f(x_{k+1}), d_k \rangle = 0$. On a donc $\langle d_{k-1}, \nabla f(x_k) \rangle = 0$ pour $1 \le k \le n$ ou encore $\langle d_{k-1}, g_k \rangle = 0$.
- On a $\langle d_k, \nabla f(x_k) \rangle = \langle \beta_k d_{k-1} - g_k, \nabla f(x_k) \rangle = - \langle g_k, \nabla f(x_k) \rangle = -\|\nabla f(x_k)\|^2$.
- On a $g_{k+1} - g_k = H(x_{k+1} - x_k) = \lambda_k H d_k$.

On déduit par ce qui précède que:

$$\beta_{k+1} = \frac{\langle g_{k+1}, Hd_k \rangle}{\langle d_k, Hd_k \rangle} = \frac{\langle g_{k+1}, g_{k+1} - g_k \rangle}{\langle d_k, g_{k+1} - g_k \rangle} = \frac{\|g_{k+1}\|^2 - \langle g_{k+1}, g_k \rangle}{\langle d_k, g_{k+1} \rangle - \langle d_k, g_k \rangle} = \frac{\|g_{k+1}\|^2 - \langle g_{k+1}, g_k \rangle}{\|g_k\|^2}.$$

Enfin par l'algorithme on peut écrire $g_k = \beta_k d_{k-1} - d_k$ d'où, en utilisant $\langle g_{k+1}, d_k \rangle = 0$ et $g_{k+1} = g_k + \lambda_k H d_k$:

$$\langle g_{k+1}, g_k \rangle = \langle g_{k+1}, \beta_k d_{k-1} - d_k \rangle = \beta_k \langle g_{k+1}, d_{k-1} \rangle - \langle g_{k+1}, d_k \rangle = \beta_k \langle g_{k+1}, d_{k-1} \rangle = \beta_k \langle g_k + \lambda_k H d_k, d_{k-1} \rangle = 0,$$

où dans la dernière inégalité on utilise que les directions d_k et d_{k-1} sont H-conjuguées. D'où le résultat. □

Ces deux remarques sont utiles en pratique pour adapter l'algorithme du gradient conjugué à la minimisation d'une fonction quelconque $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1 (i.e. pas seulement quadratique). On utilise alors l'algorithme suivant.

- Etape 0 : choisir $x_0 \in \mathbb{R}^n$; poser $g_0 = \nabla f(x_0)$ et $d_0 = -g_0$. Choisir $n_{iter} \in \mathbb{N}^*$.
- Pour $k \in \{0, ..., n_{iter}\}$ faire si $g_k = 0$ STOP (on a convergé $x_k = x^*$ i.e. critère d'arrêt) ; sinon faire:

$$\begin{array}{l} \lambda_k \in \arg\ \min_{t \in \mathbb{R}} f(x_k + td_k) \\ x_{k+1} = x_k + \lambda_k d_k \\ \beta_{k+1} = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2} \\ g_{k+1} = \nabla f(x_{k+1}) \\ d_{k+1} = -g_{k+1} + \beta_{k+1} d_k \end{array}$$

Remarque 4.5. *(i) Il existe d'autre choix de β_{k+1} dans la littérature dans le cas d'une fonction f quelconque : formule de Hestenes-Stiefel (1952), Polak-Ribière-Polyak (1969), Dai-Yuan (1999), formule de la descente conjuguée de Fletcher (1987) qui coïncident dans le cas quadratique.*

4.3 Algorithme de Newton

4.3.1 Méthode de Newton

Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ une application de classe C^1. Le but est de proposer une méthode numérique (méthode de Newton ou quasi-Newton) pour résoudre l'équation

$$f(x) = 0. \tag{4.3.1}$$

Remarque 4.6. *Soit $F : \mathbb{R}^n \to \mathbb{R}$ une application de classe C^2. Si x^* est un minimum local de F alors $\nabla F(x^*) = 0$. Ainsi, la méthode que l'on va proposer pour résoudre* (4.3.1) *pourra s'adapter pour trouver les points critiques de F (c.a.d. résoudre l'équation $\nabla F(x) = 0$) en posant $f = \nabla F$ qui est de classe C^1.*

Pour trouver un zéro de f on considère l'algorithme suivant appelé **Méthode de Newton**.

- Choisir $x_0 \in U$ et pour $k = 0, ..., N_{iter}$ faire $x_{k+1} = x_k - Df(x_k)^{-1}(f(x_k))$.

Définition 4.5. *Soit (x_k) une suite de $\mathbb{R}^n$ convergeant vers un point x^*. On dit que la suite x_k converge vers x^**

- *linéairement si il existe $r > 0$ tel que $\|x_{k+1} - x^*\| \leq r\|x_k - x^*\|$ pour tout $k \in \mathbb{N}$.*
- *sur-linéairement si $\frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|} \to 0$ quand $k \to +\infty$.*
- *de façon quadratique si $\exists K \geq 0 \quad \forall k \in \mathbb{N} \quad \|x_{k+1} - x^*\| \leq K\|x_k - x^*\|^2$.*

Théorème 4.6. *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ une application de classe C^1. Soit $x^* \in U$ tel que $f(x^*) = 0$ et tel que la différentielle $Df(x^*)$ soit inversible.*

1. *Il existe $\delta > 0$ tel que si $x_0 \in U$, $\|x_0 - x^*\| \leq \delta \Rightarrow$ la suite $(x_k)_{k\geq 0}$ converge vers x^*. De plus on a une convergence sur-linéaire:*
$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \to 0.$$
2. *Si il existe $L > 0$ tel que pour tout $x, y \in U$, $\|x - x^*\| \leq \delta$ et $\|y - x^*\| \leq \delta \Rightarrow \|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ alors la convergence est quadratique i.e.:*
$$\exists K \geq 0 \quad \forall k \in \mathbb{N} \quad \|x_{k+1} - x_k\| \leq K\|x_k - x^*\|^2.$$

Preuve. Rappelons que $f(x) = \int_0^1 Df(x^* + t(x - x^*))dt$. D'où comme $f(x^*) = 0$:

$$x_{k+1} - x^* = x_k - x^* - Df(x_k)^{-1}f(x_k) = Df(x_k)^{-1}\left[Df(x_k)(x_k - x^*) - \int_0^1 Df(x^* + t(x_k - x^*))(x_k - x^*)dt\right] \tag{4.3.2}$$

$$= Df(x_k)^{-1}\int_0^1 [Df(x_k) - Df(x^* + t(x_k - x^*))](x_k - x^*)dt. \tag{4.3.3}$$

D'où en prenant la norme de Frobenius (pour $A \in M_n(\mathbb{R})$, $\|A\| := \sqrt{\sum_{i,j} |a_{i,j}|^2}$) on déduit que

$$\|x_{k+1} - x^*\| \leq \|Df(x_k)^{-1}\|\|x_k - x^*\| \int_0^1 \|Df(x_k) - Df(x^* + t(x_k - x^*))\|dt.$$

Comme $Df(x^*)$ est inversible, on déduit qu'il existe $\delta > 0$ tel que pour tout $x \in U$, $\|x - x^*\| \leq \delta \Rightarrow \|Df(x)^{-1}\| \leq M$. Quitte à diminuer δ, on peut supposer que 2δ est un $\frac{1}{2M}$-module d'uniforme continuité pour Df i.e. $\|z - x^*\| \leq \delta$ et $\|y - x^*\| \leq \delta \Rightarrow \|y - z\| \leq 2\delta \Rightarrow \|Df(y) - Df(z)\| \leq \frac{1}{2M}$. D'où:

$$\|x_{k+1} - x^*\| \leq \frac{1}{2}\|x_k - x^*\|,$$

et donc la suite x_k d'une part vérifie $\|x_k - x^*\| \leq \delta$ pour tout k, d'autre part $x_k \to x^*$ car $\|x_k - x^*\| \leq \frac{1}{2^k}\|x_0 - x^*\|$ par réccurence. Pour (2) on a $\int_0^1 \|Df(x_k) - Df(x^* + t(x_k - x^*))\| \leq \frac{LM}{2}$ et on conclut comme précédemment. $\square$

Remarque 4.7. *La méthode de Newton suppose que le point initial x_0 est "proche" de la solution ce qui en pratique n'est pas évident. L'hypothèse fondamentale du théorème précédent est que $Df(x^*)$ est inversible. L'algorithme est couteux au sens où à chaque étape on doit inverser une matrice (i.e. $O(n^3)$ opérations). Les méthodes de quasi-Newton permettent d'approcher $Df(x_k)^{-1}$.*

4.3.2 Méthode de quasi-Newton

Le but est de résoudre une équation $f(x) = 0$ où $f : U \to \mathbb{R}^n$ mais on ne suppose pas (ou peu) connue $Df(x)^{-1}$. On considère l'algorithme de quasi-Newton **Méthode de quasi-Newton**.

- Choisir $x_0 \in U$ et pour $k = 0, ..., N_{iter}$ calculer x_k tel que $f(x_k) + M_k(x_{k+1} - x_k) = 0$ où $(M_k)_k$ est une suite de matrice dans $M_n(\mathbb{R})$.

Remarque 4.8. *(i) Toute la méthode repose sur le choix judicieux de la suite (M_k) que l'on va détailler ci-dessous.*
(ii) Si $M_k = Df(x_k)^{-1}$ on retrouve l'algorithme de Newton.

Proposition 4.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une application de classe C^1 et $(M_k)_k \in M_n(\mathbb{R})^{\mathbb{N}}$.*
(i) Si la suite $(M_k)_k$ vérifie

$$F(x_k) + M_k(x_{k+1} - x_k) = 0, \tag{4.3.4}$$

alors la suite (x_k) converge sur-linéairement vers x^* ssi $\|(M_k - Df(x^*))(x_{k+1} - x_k)\| = o(\|x_{k+1} - x_k\|)$ (critère de Deniis et Moré).
(ii) Si la suite $(M_k)_k$ vérifie

$$M_{k+1}(x_{k+1} - x_k) = F(x_{k+1}) - F(x_k), \tag{4.3.5}$$

alors la suite (x_k) converge vers x^* sur-linéairement ssi $\|(M_{k+1} - M_k)(x_{k+1} - x_k)\| = o(\|x_{k+1} - x_k\|)$.

Preuve. Le point (i) est proposé en exercice (critère de Dennis et Moré). Pour (ii) on a (comme f est C^1):

$$Df(x^*)(x_{k+1} - x_k) = f(x_{k+1}) - f(x_k) + o(\|x_{k+1} - x_k\|).$$

d'où

$$(M_{k+1} - M_k)(x_{k+1} - x_k) = (Df(x^*) - M_k)(x_{k+1} - x_k) + o(\|x_{k+1} - x_k\|).$$

Le point (ii) s'obtient donc en utilisant le résultat (i). □

En pratique (4.3.5) ne permet de déterminer M_{k+1} de façon unique car il y a n équations pour n^2 inconnues ($\frac{n(n+1)}{2}$ inconnues si $f = \nabla F$). On choisit alors M_{k+1} proche de M_k pour une certaine nombre:

$$\min_{M \in M_n(\mathbb{R})} \|M - M_k\| \quad \text{t.q.} \quad M(x_{k+1} - x_k) = F(x_{k+1}) - F(x_k) \quad \text{et} \quad M \in K, \tag{4.3.6}$$

où $K \subset M_n(\mathbb{R})$ est un ensemble convexe fermé prenant en compte des informations sur $f(x^*)$. Citons le résultat suivant de convergence.

Théorème 4.7. *Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 et $x^* \in \mathbb{R}^n$ tel que $f(x^*) = 0$ et tel que $Df(x^*)$ soit inversible. On suppose que $Df(x)$ est Lipschitzien et qu'il existe un voisinage V de x tel que:*

$$\forall x \in V \quad Df(x) \in K,$$

où K est un ensemble convexe fermé donné de $M_n(\mathbb{R})$. Alors, si (x_0, M_0) est suffisamment proche de $(x^*, Df(x^*))$ et si M_{k+1} est donné itérativement par (4.3.6) alors la suite (x_k) définie par récurrence par

$$x_{k+1} = x_k - M_k^{-1} f(x_k)$$

converge sur-linéairement vers x^*.

Exemple : formule de Broyden. Soit $B_0 \in M_n(\mathbb{R})$, $s \in \mathbb{R}^n$ non nul et $y \in \mathbb{R}^n$. La solution de

$$\min_{B \in M_n(\mathbb{R})} \frac{1}{2}\|B - B_0\|^2 \quad \text{t.q.} \quad Bs = y$$

est donnée par la perturbation de rang 1 de B_0 (voir exercice pour le démontrer):

$$B = B_0 + \frac{1}{\|s\|^2}(y - B_0 s)s^T$$

La méthode de quasi-Newton s'écrit alors

$$\begin{cases} & M_0 \in M_n(\mathbb{R}) \text{ et } \quad x_0 \in \mathbb{R}^n \quad \text{choisis} \\ M_{k+1} & = M_k + \frac{1}{\|x_{k+1} - x_k\|^2}(f(x_{k+1}) - f(x_k) - M_k(x_{k+1} - x_k))(x_{k+1} - x_k)^T \\ x_{k+1} & = x_k + M_{k+1}^{-1}(f(x_{k+1}) - f(x_k)) \end{cases}$$

Chapitre 5

Exercices

5.1 Exercices sur la convexité et les espaces de Hilbert

Exercice 5.1. *1) Montrer que* $A \in M_n(\mathbb{R}) \mapsto A^2$ *est Gâteaux-différentiable en tout point* $A \in M_n(\mathbb{R})$*.*
2) Soit $f : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ *définie par* $f(A) = e^{Tr(A)}A$*. Montrer que* f *est Gateaux-différentiable, Fréchet différentiable et de classe* C^1*.*

Exercice 5.2. *Pour* $u \in L^2([0,1],\mathbb{R})$*, on note* $\|u\|_2 := (\int_0^1 u^2)^{\frac{1}{2}}$*. Montrer l'inégalité de Poincaré en dimension 1:*

$$u \in H_0^1([0,1],\mathbb{R}) \;\Rightarrow\; \|u\|_2^2 \leq \|u'\|_2^2.$$

En déduire que $H_0^1([0,1],\mathbb{R})$ *peut-être muni du produit scalaire* $\langle u,v\rangle := \int_0^1 u'v'$*.*

Exercice 5.3. *1) Démontrer le théorème de projection sur un convexe:*

Théorème 5.1. *Soit* C *un ensemble convexe fermé non-vide d'un espace de Hilbert* H*. Alors pour tout* $x \in H$ *il existe un unique* $p(x) \in H$ *tel que:* $\|x - p(x)\| = \inf\{\|x-y\| \;;\; y \in C\}$*. Le point* $p(x)$ *est caractérisé par:*

$$\forall y \in C \;\; \langle x - p(x), y - p(x)\rangle \leq 0.$$

- Pour l'unicité, utiliser avec $x - y_1$ *et* $x - y_2$ *l'égalité du parallélogramme:*

$$\|\frac{u-v}{2}\|^2 + \|\frac{u+v}{2}\|^2 = \frac{1}{2}(\|u\|^2 + \|v\|^2).$$

- Pour l'existence, prendre une suite minimisante, et montrer que la suite est de Cauchy par l'identité ci-dessus.
- Démontrer la caractérisation variationnelle de la projection.
2) Démontrer le théorème de séparation.

Théorème 5.2. *Soit* H *un espace de Hilbert,* C *un convexe fermé non-vide, et* $x_0 \in H$ *tel que* $x_0 \notin C$*. Alors il existe* $p \in H$ *et* $\varepsilon > 0$ *tel que:*

$$\forall y \in C, \;\langle p, x_0\rangle \leq \langle p, y\rangle - \varepsilon. \tag{5.1.1}$$

Exercice 5.4. *Démontrer la propriété 2.8 i.e. étant donnée une fonction* $f : H \to \mathbb{R} \cup \{+\infty\}$ *convexe (où* H *est un espaces de hilbert), pour tout* $x \in Dom(f)$*, l'ensemble* $\partial f(x)$ *est convexe fermé.*

Exercice 5.5. Intersection de convexes en dimension infini.
Soit C_n *une suite décroissante de convexes fermés bornés non-vide d'un espace de Hilbert* H*.*
1) On définit une suite réelle (u_n) *par* $u_n := d(0, C_n)$*. Montrer que* (u_n) *est croissante majorée.*
2) On définit une suite (x_n) *de* C_n *vérifiant* $\|x_n\| \leq d(0, C_n) + \frac{1}{n}$ *pour tout* $n \in \mathbb{N}^*$*.*
a) Justifier l'existence d'une telle suite.
b) A l'aide de la règle du parallélogramme montrer que $n \geq m \Longrightarrow \|x_n - x_m\|^2 \leq 2(\|x_n\|^2 + \|x_m\|^2) - 4d(0, C_m)^2$*. (utiliser également la convexité de* C_m*).*
c) Déduire que la suite (x_m) *est une suite de Cauchy. On note* x *sa limite.*
3) Montrer que pour tout $m \in \mathbb{N}$*,* $n \geq m \Longrightarrow x_n \in C_m$*. En déduire que* $x \in C_m$ *et le résultat:*

$$\cap_{n\in\mathbb{N}} C_n \neq \emptyset.$$

Exercice 5.6. Existence en dimension infini.
Soit C un convexe fermé d'un espace de Hilbert H et $f : H \to \mathbb{R}$ une fonction convexe continue. On suppose que C est borné ou que $f(x) \to +\infty$ si $\|x\| \to +\infty$. Le but de l'exercice est de montrer que le problème

$$\inf_{x \in C} f(x)$$

admet une solution.
1) Soit $a \in \mathbb{R}$ et $E_a := \{x \in C \; ; \; f(x) \le a\}$. Montrer que si a est tel que $E_a \neq \emptyset$ alors E_a est borné.
2) Soit $m := \inf_{x \in C} f(x) \in \mathbb{R} \cup \{-\infty\}$. Soit $m_n \to m$ décroissante et $C_n := \{x \in C \; ; \; f(x) \le m_n\}$.
a) Montrer que C_n est un convexe fermé borné non vide.
b) En utilisant l'exercice 5.5, montrer qu'il existe $x_0 \in \cap_{n \in \mathbb{N}} C_n$.
c) En déduire que $f(x_0) = m$ et que m est fini.
3) Montrer que si f est strictement convexe alors le minimum est unique.
4) Etendre le résultat au cas où f est semi-continue[1] et quasi-convexe[2]

Exercice 5.7. Minorante affine pour une fonction convexe.
Soit $f : H \to \mathbb{R}$ une fonction convexe continue sur un ensemble convexe fermé C. Le but est de montrer que f admet une minorante affine: il existe $y \in H$ et $\gamma \in \mathbb{R}$ tels que:

$$\forall x \in K \quad f(x) \ge \langle x, y \rangle + \gamma.$$

1) On rappelle que l'épigraphe d'une fonction convexe $f : H \to \mathbb{R}$ $Epi(f) := \{(x,h) \in H \times \mathbb{R} \; ; \; h \ge f(x)\}$ est un ensemble fermé. Soit $x_0 \in K$ tel que $(x_0, h_0) \notin Epi(f)$. A l'aide du théorème de séparation, montrer qu'il existe $(y,z) \in H \times \mathbb{R}$ et $\alpha \in \mathbb{R}$ tels que:

$$\forall (x,h) \in K \times \mathbb{R}, \; f(x) \le h \implies \langle y, x \rangle + zh < \alpha < \langle y, x_0 \rangle + zh_0.$$

2) Montrer que $z < 0$ (supposer d'abord $z > 0$ et utiliser que $(x,h) \in Epi(f) \Rightarrow (x, h+t) \in Epi(f)$ pour $t > 0$; puis exclure le cas $z = 0$).
3) Montrer que $\forall x \in K$, $\langle y, x \rangle + zf(x) < \alpha$

Exercice 5.8. Existence d'un minimum pour une fonction fortement convexe.
1) Soit f une fonction fortement convexe de paramètre $\alpha > 0$. Montrer qu'il existe deux constantes $\gamma > 0$ et $\delta \in \mathbb{R}$ telles que:

$$f(x) \ge \gamma \|x\|^2 + \delta.$$

(appliquer la définition avec $t = \frac{1}{2}$; utiliser l'exercice (5.7) et l'inégalité de Cauchy-Swarz $\langle x, y \rangle \ge -\|x\|\|y\|$).
2) On pose $m := \inf_{x \in H} f(x)$ et soit (x_n) une suite minimisante. Montrer que (x_n) est une suite de Cauchy (utiliser l'inégalité de forte convexité et retrancher m à cette inégalité).

Exercice 5.9. *Soit H un espace de Hilbert et (x_n) une suite de H qui converge faiblement vers $x \in H$. Montrer que*

$$\|x\|_H \le \liminf_{n \to +\infty} \|x_n\|_H.$$

Exercice 5.10. *1) Soit X un espace de Hilbert et $\| \cdot \|$ la norme associé au produit scalaire. Montrer que $\| \cdot \|^2$ est fortement convexe.*
2) Si $A \in \mathcal{L}(X, X^*)$ vérifie : il existe $\alpha > 0$ tel que pour tout $x \in X$ on ait $\langle Ax, x \rangle \ge \alpha \|x\|^2$ alors $x \longmapsto \langle Ax, x \rangle$ est fortement convexe.

[1]On dit que $f : H \to \mathbb{R}$ est semi-continue inférieurement si pour tout $x_0 \in H$, $f(x_0) \le \liminf_{x \to x_0} f(x)$.
[2]On dit que $f : H \to \mathbb{R}$ est quasi-convexe si pour tout $\alpha \in \mathbb{R}$ tel que $S_\alpha := \{x \in H \; ; \; f(x) \le \alpha\}$ soit non-vide alors S_α est convexe.

5.2 Exercices sur l'optimisation sans contraintes

Exercice 5.11. *Soit E un espace métrique et $f : E \to \mathbb{R}$ une application continue. Si f a un sous-niveau non-vide compact, montrer alors que f admet un minimum global sur E.*

Exercice 5.12. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de la forme $f(x) := \sum_{1\le i\le n} f_i(x_i)$ où $x = (x_1, ..., x_n) \in \mathbb{R}^n$ et $f_i : \mathbb{R} \to \mathbb{R}$ est une fonction minorée et coercice, $1 \le i \le n$. Montrer alors que f est coercive sur $\mathbb{R}^n$.*

Exercice 5.13. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^2.*
a) Montrer que si $\bar{x}$ est un minimum local de f alors $D^2 f(\bar{x})$ est positive.
b) Montrer que si $\bar{x}$ est tel que $\nabla f(\bar{x}) = 0$ et si $D^2 f(\bar{x})$ est définie positive, alors $\bar{x}$ est un minimum local.

Exercice 5.14. *(*)* **Principe variationnel d'Ekeland en dimension finie**. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application semi-continue inférieurement et bornée inférieurement. Soit $\varepsilon > 0$, et $x \in \mathbb{R}^n$ tel que:*

$$f(x) \le \inf_{\mathbb{R}^n} f + \varepsilon.$$

1) On pose $g(z) = f(z) + \varepsilon\|z - x\|$. Montrer que g est semi-continue inférieurement, coervice et que l'ensemble K de ses minimiseurs est un sous-ensemble compact non-vide de $\mathbb{R}^n$.
2) Soit x_ε un point qui minimise f sur K. Montrer que $\|x_\varepsilon - x\| \le 1$ et que $f(x_\varepsilon) \le f(x)$.
3) Montrer que pour tout $z \in \mathbb{R}^n$, $z \neq x_\varepsilon$ on a $f(x_\varepsilon) < f(z) + \varepsilon\|z - x_\varepsilon\|$ (pour cela distinguer deux cas selon que $z \in K$ ou que $z \notin K$ où K est l'ensemble des minimum de g). Interpréter géométriquement ce résultat.

Exercice 5.15. *Soit U l'ensemble des $n-$ uplets distincts de $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ définie par:*

$$f(x) = \sum_{1\le i\le n} x_i^2 - \sum_{1\le i<j\le n} \ln(|x_i - x_j|).$$

1) Montrer que f est coercive. Etudier le comportement de f sur la frontière de U. En déduire que f atteint sa borne inférieure sur U.
2) Soit a un minimiseur de f sur U. Montrer que pour $1 \le k \le n$:

$$2a_k - \sum_{j\neq k} \frac{1}{a_k - a_j} = 0$$

2) Soit $H(t) = \Pi_{1\le k\le n}(t - a_k)$ et $Q_k(t) = H(t)/(t - a_k)$. Montrer que $H'(a_k) = Q_k(a_k)$ et $H''(a_k) = 2Q_k'(a_k)$.
3) En déduire que la condition d'optimalité s'écrit $H''(a_k) - 4a_k H'(a_k) = 0$ pour $1 \le k \le n$.
4) En déduire que $H''(t) - 4tH'(t) = -4nH(t)$ pour tout $t \in \mathbb{R}$, une façon de calculer les minima de f sur U, ainsi que le nombre de ces minima.

Exercice 5.16. *Soit a, b deux vecteurs de $\mathbb{R}^n$ non nuls et $f : \mathbb{R}^n \to \mathbb{R}$ la fonction définie par $f(x) := \frac{\langle x,a\rangle\langle x,b\rangle}{\|x\|^2}$ pour $x \in \mathbb{R}^n$ non nul.*
1) Montrer que la borne inférieure et la borne supérieure de f sur $\mathbb{R}^n\backslash\{0\}$ sont atteintes et calculer ces valeurs (indication: introduire les deux vecteurs $u := \frac{1}{2}(a/\|a\| + b/\|b\|)$ et $v := \frac{1}{2}(a/\|a\| - b/\|b\|)$.
2) Soit $M = ab^T + ba^T \in \mathcal{S}_n(\mathbb{R})$. Trouver la plus petite et la plus grande valeur propre de M.

Exercice 5.17. *(*)* **Théorème de Danskin**. *Soit $f : X \times Y \to \mathbb{R}$ une fonction continue où X est un ouvert de $\mathbb{R}^n$ et Y un compact d'un espace topologique F. On suppose que pour tout $(x, y) \in X \times Y$, $\nabla_x f(x, y)$ existe et $(x, y) \mapsto \nabla_x f(x, y)$ est continue sur $X \times Y$. Soit $\varphi : X \to \mathbb{R}$ la fonction définie par $\varphi(x) = \max_{y\in Y} f(x, y)$ et soit $Y(x)$ le sous-ensemble de Y défini par $Y(x) := \{y \in Y\ ;\ \varphi(x) = f(x, y)\}$.*
1)a) Soit $x_0 \in X$ et (x_k) une suite de X qui converge vers x_0, et (y_k) une suite telle que $\varphi(x_k) = f(x_k, y_k)$. Montrer que $\lim_{k\to+\infty} \varphi(x_k) \ge f(x_0, y)$ pour tout $y \in Y$.
b) En déduire que φ est continue en x_0.
2) Soit $h \neq 0$ fixé et $x_k = x_0 + t_k h$, où $(t_k)_{k\in\mathbb{N}}$ est une suite positive convergeant vers 0 et soit $y \in Y(x_0)$ et $(y_k)_{k\in\mathbb{N}} \in Y$ tels que $\varphi(x_k) = f(x_k, y_k)$.
a) Montrer qu'il existe $t_k' \in \mathbb{R}$ tel que $0 \le t_k' \le t_k$ et tel que $\frac{\varphi(x_k)-\varphi(x_0)}{t_k} \ge \langle\nabla f(x_0 + t_k' h, y), h\rangle$.
b) En déduire que: $\liminf_{k\to+\infty} \frac{\varphi(x_k)-\varphi(x_0)}{t_k} \ge \max_{y\in Y(x_0)} \langle\nabla_x f(x_0, y), h\rangle$.
3) Par un raisonnement similaire, établir que:

$$\limsup_{k\to+\infty} \frac{\varphi(x_k) - \varphi(x_0)}{t_k} \le \max_{y\in Y(x_0)} \langle\nabla_x f(x_0, y), h\rangle.$$

4) En déduire que $\varphi'(x_0, h)$ existe pour toute direction h ainsi qu'une valeur de $\varphi'(x_0, h)$.

5.3 Démonstration du lemme de Farkas

Exercice 5.18. *Le but de l'exercice est de démontrer la proposition suivante.*

Proposition 5.1. *Soit E un espace vectoriel normé, $q \in \mathbb{N}^*$ et $a_1, ..., a_q$ q vecteurs de E. Soit l'ensemble:*

$$K := \left\{ \sum_{i=1}^{q} \lambda_i a_i \ ; \ \lambda_i \geq 0, \ 1 \leq i \leq q \right\}.$$

Alors K est un cône convexe fermé de E.

1) On procède par récurrence. Montrer le résultat au rang $q = 1$.
2) On suppose le résultat vrai au rang $q \in \mathbb{N}^*$. Soit $a_1, ..., a_q, a_{q+1}$, $q+1$ vecteurs de E et l'ensemble K défini par:

$$K := \left\{ \sum_{i=1}^{q+1} \lambda_i a_i \ ; \ \lambda_i \geq 0, \ 1 \leq i \leq q+1 \right\}.$$

a) On suppose $-a_i \in K$ pour tout $1 \leq i \leq q+1$. Montrer le résultat.
b) On suppose qu'il existe $i \in \{1, ..., q+1\}$ tel que $-a_i \notin K$ et quitte à permuter les indices on suppose $-a_{q+1} \notin K$. Soit (y_n) une suite de $K^{\mathbb{N}}$ et $y \in E$ telle que $y_n \to y$ lorsque $n \to +\infty$. Montrer que pour tout $i \in \{1, ..., q+1\}$ et tout $n \in \mathbb{N}$ il existe $\lambda_{i,n} \in \mathbb{R}_+$ et $\mu_n \in \mathbb{R}_+$ tels que :

$$y_n = \sum_{i=1}^{q} \lambda_{i,n} a_i + \mu_n a_{q+1} = z_n + \mu_n a_{q+1}.$$

c) En raisonnant par l'absurde, montrer que la suite (μ_n) est bornée.
d) En déduire qu'à une sous-suite près, la suite (μ_n) converge et que $y \in K$.

Exercice 5.19. *On utilise maintenant l'exercice précédent pour montrer le* **lemme de Farkas**.

Lemme 5.1. *Soit H un espace de Hilbert et $(a, a_1, ..., a_q) \in H^{q+1}$. On a alors l'équivalence suivante:*
(i) $\forall x \in H$, $\langle a_i, x \rangle \leq 0$, $1 \leq i \leq q \Rightarrow \langle a, x \rangle \leq 0$.
(ii) $\exists (\lambda_1, ..., \lambda_q) \in \mathbb{R}_+^q$ $a = \sum_{i=1}^{q} \lambda_i a_i$.

1) Montrer que (ii) $\Rightarrow$ (i).
2) On souhaite montrer (i) $\Rightarrow$ (ii). Soit K l'ensemble défini par :

$$K := \left\{ \sum_{i=1}^{q+1} \lambda_i a_i \ ; \ \lambda_i \geq 0, \ 1 \leq i \leq q \right\}.$$

a) Montrer que l'ensemble K est un sous-ensemble convexe fermé non-vide de H.
b) On suppose $a \notin K$. Par le théorème de séparation établir qu'il existe $\varepsilon > 0$ et $q \in H$ un vecteur non-nul tels que:

$$\forall y \in K, \quad \langle q, y \rangle \leq \langle q, a \rangle - \varepsilon.$$

c) En utilisant que $ty \in K$ pour $t > 0$ et $y \in K$, montrer que $\langle q, y \rangle \leq 0$.
d) En déduire que $\sup_{y \in K} \langle q, y \rangle = 0$.
e) Conclure en montrant d'abord que $\langle q, a \rangle > 0$ et $\langle q, a_i \rangle \leq 0$ pour tout $1 \leq i \leq q$.

Exercice 5.20. *Démontrer le corollaire suivant (dans le cas d'inégalités et d'égalités).*

Corollaire 5.1. *Soit H un espace de Hilbert et $(a_1, ..., a_p, a_{p+1}, ..., a_{p+q}, a) \in H^{p+q+1}$. On a alors l'équivalence suivante:*
(i) $\forall x \in H$, $\langle a_i, x \rangle \leq 0$, $1 \leq i \leq p$ et $\langle a_{p+i}, x \rangle = 0$, $1 \leq i \leq q \Rightarrow \langle a, x \rangle \leq 0$.
(ii) $\exists (\lambda_1, ..., \lambda_q) \in \mathbb{R}_+^q$ et $\exists (\lambda_{p+1}, ..., \lambda_{p+q}) \in \mathbb{R}^q$ $a = \sum_{i=1}^{p+q} \lambda_i a_i$.

5.4 Exercices autour du théorème de KKT

Exercice 5.21. *Le but de cet exercice est de démontrer le théorème de KKT à l'aide du lemme de Farkas. Soit* $I := \{1, ..., l\}$ *et* $J := \{1, ..., m\}$ *où* l *et* m *sont deux entiers strictement positifs. Soit le problème de minimisation:*

$$\min_{x \in K} f(x), \quad \text{avec} \quad K := \{x \in \mathbb{R}^n \; ; \; g_i(x) \leq 0, \; i \in I \; ; \; h_j(x) = 0 \; ; \; j \in J\},$$

où $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq l$, $f_j : \mathbb{R}^n \to \mathbb{R}$, $1 \leq j \leq m$ *sont de classe* C^1 *sur* $\mathbb{R}^n$*. Soit* x *un minimum local du problème et* $I(x)$ *l'ensemble des contraintes actives. On suppose les contraintes qualifiées au point* x*. Soit*

$$E := \{v \in \mathbb{R}^n \; ; \; \forall i \in I(x) \;\; \langle \nabla g_i(x), v\rangle \leq 0 \text{ et } \forall j \in J \;\; \langle \nabla h_j(x), v\rangle \leq 0 \text{ et } \langle -\nabla h_j(x), v\rangle \leq 0\}.$$

1) Que peut-on dire des deux ensembles E *et* $T_K(x)$*?*
2) Appliquer la condition d'Euler au problème d'optimisation.
3) Par le lemme de Farkas, montrer que pour tout $i \in \{1, ..., l\}$ *et pour tout* $j \in \{1, ..., m\}$ *il existe* $\lambda_i \in \mathbb{R}_+$ *et* $(\nu_j, \rho_j) \in \mathbb{R}^2$ *tels que*

$$-\nabla f(x) = \sum_{i \in I(x)} \lambda_i \nabla g_i(x) + \sum_{j=1}^{m} (\nu_j \nabla h_j(x) - \rho_j \nabla h_j(x)).$$

En déduire la condition de stationarité du Lagrangien au point x*.*
4) Montrer que le vecteur $\lambda := (\lambda_i)_{1 \leq i \leq l}$ *peut être choisi de façon à satisfaire la condition de complémentarité.*

Exercice 5.22. *1) Soit* $A \in \mathcal{S}_n^{++}(\mathbb{R})$*. Vérifier que le problème d'optimisation*

$$\inf_{x \in \mathbb{R}^n} \langle Ax, x\rangle \quad \text{tel que} \quad 1 - \|x\|^2 = 0.$$

admet une solution, et résoudre ce problème en appliquant les conditions de KKT.
2) Montrer que le multiplicateur de Lagrange associé à la contrainte d'égalité est la plus petite valeur propre de A.
3) Etant donnés une famille orthonormée de vecteur $\{u_1, ..., u_{k-1}\}$ *avec* $k \geq 2$*, on considère le problème d'optimisation:*

$$\inf_{x \in \mathbb{R}^n} \frac{1}{2} \langle Ax, x\rangle \quad \text{tel que} \quad \frac{1}{2}(1 - \|x\|^2) = 0 \quad \text{et} \quad \langle x, u_i\rangle = 0, \qquad \forall i \in \{1, ..., k-1\}.$$

a) Ecrire les conditions de KKT pour ce problème (vérifier la qualification des contraintes).
b) Démontrer par récurrence sur l'entier $k \in \{2, ..., n-1\}$ *que les multiplicateurs de Lagrange* $(\mu_i^k)_{1 \leq i \leq k-1}$ *associées au* $k-1$ *contraintes* $\langle x, u_i\rangle = 0$ *ci-dessus sont nuls (utiliser l'hypothèse de récurrence pour montrer que* u_i*,* $1 \leq i \leq k-1$ *est vecteur propre pour* A*).*
c) Quel résultat d'algèbre linéaire obtient-on ainsi?

Exercice 5.23. *Pour* $b \in \mathbb{R}$*, soit le problème d'optimisation*

$$\min_{(x_1, x_2) \in \mathbb{R}^2} f(x_1, x_2) := (x_1 - 2)^2 + (x_2 - 1)^2 \quad \text{tel que} \quad x_1 + x_2 \leq b \quad \text{et} \quad -x_2 \leq 0.$$

1) Trouver en fonction de b la solution de ce problème notée $(x_1(b), x_2(b))$*.*
2) Pour chaque solution relier $\frac{d}{db} f(x_1(b), x_2(b))$ *aux multiplicateurs de Lagrange.*

Remarque: le résultat donne une interprétation des multiplicateurs de Lagrange : ils sont liés à la sensibilité du critère par rapport à des incertitudes (ici l'incertitude est paramétrée par b). Cette interprétation est fondamentale en économie.

Exercice 5.24. *Soit le problème*

$$\inf_{x \in \mathbb{R}^n} \frac{1}{2} \langle Ax, x\rangle - \langle b, x\rangle \quad \text{avec} \quad Bx = 0$$

où $A \in \mathcal{S}_n^{++}(\mathbb{R})$*,* $b \in \mathbb{R}^n$*, et* $B \in M_{m,n}(\mathbb{R})$*. Ecrire les conditions d'optimalité. Trouver la solution* x^* *et le multiplicateur de Lagrange associé en fonction des données.*

Exercice 5.25. *Résoudre dans* $\mathbb{R}^2$ *le problème*

$$\min_{x^2 + y^2 \leq 1} xy.$$

Exercice 5.26. *Résoudre dans* $\mathbb{R}^3$ *le problème*

$$\min_{\substack{x^2+y^2=1\\ y^2+z^2=4}} x+z.$$

Exercice 5.27. *Résoudre dans* $\mathbb{R}^3$ *le problème*

$$\min_{\substack{x^2+y^2+z^2\leq 1\\ x\geq 0}} x+y+z.$$

Exercice 5.28. *Résoudre dans* $\mathbb{R}^3$ *le problème*

$$\min_{\substack{x^2+y^2+z^2=1\\ x+y+z\leq 0}} x+2y+3z.$$

Exercice 5.29. *Parmi tous les triangles du plan de périmètre p fixé, lequel possède une aire maximale? (Indication: utiliser la formule de Héron* $\mathcal{A}=\sqrt{p(p-a)(p-b)(p-c)}$*, et remplacer A par* A^2*).*

Exercice 5.30. *Soit* P *l'hyperplan de* $\mathbb{R}^n$ *d'équation* $\langle c,x\rangle = d$ *où* $c\in\mathbb{R}^n$ *et* $d\in\mathbb{R}$. *Calculer la projection orthogonale d'un point* $y\in\mathbb{R}^n$ *sur* P.

Exercice 5.31. *Conditions suffisantes d'optimalité. Soit* $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$, $1\leq i\leq l$, $1\leq j\leq m$ *des fonctions de classe* C^1. *On considère le problème d'optimisation*

$$(P)\quad \min_{x\in\mathbb{R}^n} f(x),\ g_i(x)\leq 0,\ 1\leq i\leq l,\ h_j(x)=0,\ 1\leq j\leq m.$$

On suppose que x^* *est un point de Fritz-John du problème (P) et que* $Vect(\lambda_0\nabla f(x^*), \lambda_i\nabla g_i(x^*), \nabla h_j(x^*), 1\leq i\leq l, 1\leq j\leq m)=\mathbb{R}^n$. *On souhaite montrer que* x^* *est un minimum local de (P). On suppose par l'absurde qu'il existe une suite* x_k *de points réalisables tels que* $x_k\to x^*$ *et* $f(x_k)<f(x^*)$. *En écrivant* $x_k = x^* + t_k d_k$ *avec* $t_k>0$, $\|d_k\|=1$, *montrer que l'on obtient une contradiction et conclure.*

Cône tangent et qualification des contraintes

Exercice 5.32. *1) Calculer le cône tangent à un polygone* K *de* $\mathbb{R}^2$ *en tout point de* K.
2) Calculer le cône tangent à l'ensemble K *défini par*

$$K := \{(0,0\}\cup\{(t\cos(1/t), t\sin(1/t))\ ;\ t>0\}.$$

(On montrera que K *est fermé et on tâchera d'intuiter la solution).*

Exercice 5.33. *Soit* $K := \{(x,y)\in\mathbb{R}^2\ ;\ y\geq x^2\ : y\leq 0\}$. *Montrer que la contrainte* K *n'est pas qualifiée en* $(0,0)$. *Même question en* $(0,0)$ *avec* $K := \{(x,y)\in\mathbb{R}^2\ ;\ y\leq x^3\ ; y\geq 0\}$.

Exercice 5.34. *On suppose que les fonctions* g_i *sont concaves et de classe* C^1 *et que les fonctions* h_j *sont affines. Montrer qu'alors la contrainte* K *est qualifiée en tout point.*

Autres exercices d'application directe du théorème de Karush-Kuhn-Tucker

Exercice 5.35. *Démontrer l'inégalité de Cauchy-Schwarz dans* $\mathbb{R}^n$ *à l'aide des conditions de KKT. De même montrer l'inégalité d'Hölder.*

Exercice 5.36. *Résoudre le problème*

$$\min_{(x,y,z)\in\mathbb{R}^3\ xy=1} x^2+y^2+16z^2.$$

Exercice 5.37. *Soit n fonctions $f_i : \mathbb{R} \to \mathbb{R}$ de classe C^1 et le problème d'optimisation*

$$\min \sum_{1 \le i \le n} f_i(x_i)$$

sous les contraintes $x \ge 0$ et $\sum_{1 \le i \le n} x_i = 1$. On suppose l'existence d'une solution optimale x^* à ce problème. Montrer qu'il existe $\mu \in \mathbb{R}$ tel que si $x_i^* > 0$ alors $f'(x_i^*) = \mu$, et si $x_i^* = 0$, alors $f'(x_i^*) \ge \mu$.

Exercice 5.38. *Soit $C := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 3,\ -x^2 + 2y \le 0\}$. Résoudre le problème:*

$$\max_{(x,y) \in C} (x + 1)^2 + (y + 1)^2.$$

On montrera d'abord que $\lambda_0 > 0$. On discutera également en fonction des multiplicateurs (λ_1, λ_2) la solution du problème.

Exercice 5.39. *Points de Fritz-John. Soit C l'ensemble de contraintes défini par $C = \{(x, y) \in \mathbb{R}^2 | \quad (x-1)^3 + y \le 0,\ x \ge 0, y \ge 0\}$. On considère le problème*

$$\min_{(x,y) \in C} f(x, y) := -x.$$

Trouver le minimum global de f sur C. Quelles sont les contraintes actives? Ce point satisfait-il les conditions de Karush-Kuhn-Tucker? Interpréter géométriquement la condition de qualification.

Exercice 5.40. *Stricte complémentarité. Soit C l'ensemble des contraintes défini par $C = \{(x, y) \in \mathbb{R}^2 | \quad x^2 + y^2 \le 2, \quad x + y \ge 0\}$. On considère le problème*

$$\min_{(x,y) \in C} xy.$$

Montrer que tout minimum local satisfait les conditions de KKT. Calculer tous les points de KKT. Y-a-t-il stricte complémentarité? Calculer les minima locaux (utiliser le second ordre).

Exercice 5.41. *(Second ordre, et non positivité de la Hessienne). Soit le problème d'optimisation dans $\mathbb{R}^2$*

$$\min f(x, y) := -xy,$$

sous les contraintes $(x, y) \in K := \{(x, y) \in K \; ; \; x + y \le 2 \text{ et } x \ge 0 \; y \ge 0\}$.
1) Montrer que le point $(1, 1)$ est un minimum global de f sur K et calculer le multiplicateur associé.
2) Calculer $\frac{\partial^2 L}{\partial (x,y)^2}$

Exercice 5.42. *1) En utilisant les conditions nécessaires et suffisantes du second ordre, étudier le problème suivant:*

$$\max_{(x,y) \in C} (x + 1)^2 + (y + 1)^2,$$

où $C := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 2,\ y \le 1\}$.
2) Reprendre la question précédente avec $C := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 3,\ -x^2 + 2y \le 0\}$.

Exercice 5.43. *On considère le problème d'optimisation suivant:*

$$\min_{(x,y,z) \in C} x^2 + y^2 + z^2,$$

où $C = \{(x, y, z) \in \mathbb{R}^3 \mid \quad xyz \ge 8,\ x \ge 0,\ y \ge 0,\ z \ge 0\}$.
1) Trouver tous les points de KKT du problème.
2) Ecrire les conditions du second ordre pour trouver les minima et les maxima locaux du problème.

Problème de synthèse.

Exercice 5.44. *On considère les trois fonctions $f, h_1, h_2 : \mathbb{R}^n \to \mathbb{R}$ définies par:*

$$f(x) = \frac{1}{3}\sum_{i=1}^{n} x_i^3 \quad h_1(x) = \sum_{i=1}^{n} x_i \quad h_2(x) = \frac{1}{2}\left(\sum_{i=1}^{n} x_i^2 - n\right).$$

Soit le problème d'optimisation non-linéaire:

$$\inf_{x\in K} f(x), \quad \text{avec } K := \{x \in \mathbb{R}^n \ ; \ h_1(x) = 0 \text{ et } h_2(x) = 0\}.$$

1) Montrer qu'il existe une solution au problème et que le problème est qualifié.

2) **Etude des conditions du premier ordre**.

a) Ecrire les conditions de KKT. On appellera λ_1 et λ_2 les deux multiplicateurs de Lagrange. Montrer que $\lambda_1 = -1$.

b) On pose $x_- = -\lambda_2 - \sqrt{\lambda_2^2+4}$ et $\lambda_+ = -\lambda_2 + \sqrt{\lambda_2^2+4}$. Montrer qu'un point de KKT est caractérisé par les conditions suivantes :

$$\begin{cases} x = (\underbrace{x_+, \cdots, x_+}_{k \text{ fois}}, \underbrace{x_-, \cdots, x_-}_{n-k \text{ fois}}), \\ x_- = -\sqrt{\frac{k}{n-k}}, \quad x_+ = \sqrt{\frac{n-k}{k}}, \quad \lambda_1 = -1, \quad \lambda_2 = \frac{k}{n}\sqrt{\frac{k}{n-k}} - \frac{n-k}{n}\sqrt{\frac{n-k}{k}}. \end{cases}$$

c) Calculer la fonctionnelle objectif et montrer que le minimum global est atteint par le point de KKT tel que $k = n-1$.

d) De même, montrer que le maximum global est atteint pour $k = 1$.

3) **Etude du second ordre**. *Le but de cette question est d'étudier la nature des points de KKT (autre que le minimum et le maximum global). Soit $L : \mathbb{R}^n \times \mathbb{R}^2 \to \mathbb{R}$ le lagrangien associé au problème*

a) Calculer les gradients des contraintes aux points de KKT.

b) Montrer que $\nabla^2_{xx}L(x,\lambda) = 2\text{diag}(x) + \lambda_2 Id_{\mathbb{R}^n}$.

c) Ecrire les deux conditions d'orthogonalité d'un vecteur d par rapport aux gradients des contraintes en un point x vérifiant KKT. En déduire que d est orthogonal à $\nabla h_1(x)$ et $\nabla h_2(x)$ si et seulement si:

$$\sum_{i=1}^{k} d_i = \sum_{i=k+1}^{n} d_i = 0.$$

d) Expliciter deux constantes $c_1(n,k) > 0$ et $c_2(n,k) > 0$ telles que $\nabla^2_{xx}L(x,\lambda)$ vérifie:

$$\left\langle \nabla^2_{xx}L(x,\lambda)d, d\right\rangle = c_1(n,k)\sum_{1\le i\le k} d_i^2 - c_2(n,k)\sum_{k+1\le i\le n} d_i^2$$

e) En déduire que l'unique minimum local est exactement le point pour lequel $k = n-1$ et $d_n = 0$.

f) De même montrer que lorsque $k \neq 1$, les points de KKT ne sont pas des maxima locaux.

5.5 Conditions d'optimalité en dimension infini

Exercice 5.45. *Soit le problème d'optimisation:*

$$\inf_{x\in H_0^1(]0,1[)} J(x) := \int_0^1 [(|x'(t)|-1)^2 + x^2(t)]dt.$$

1) Montrer que si (x_n) est une suite bornée de $H_0^1(]0,1[)$ et $x \in H_0^1(]0,1[)$, alors il existe $C \geq 0$ tel que

$$\left|\int_0^1 ((|x_n'(t)|-1)^2 - (|x'(t)|-1)^2)dt\right| \leq C\|x_n - x\|_{H_0^1(]0,1[)}.$$

2) Déduire en utilisant l'inégalité de Poincaré que J est s.c.i.
3) Montrer que J est coercive.
4) On définit une suite de fonctions $x_n : [0,1] \to \mathbb{R}$ affines par morceaux par: $\dot{x}_n(t) = 1$ dans $[\frac{2k}{2n}, \frac{2k+1}{2n}]$, $k \in \{0,...,n-1\}$ et $\dot{x}_n(t) = -1$ dans $[\frac{2k+1}{2n}, \frac{2k+2}{2n}]$, $k \in \{0,...,n-1\}$ et $x_n\left(\frac{2k+1}{2n}\right) = \frac{1}{2n}$, $k \in \{0,...,n-1\}$.
a) Dessiner (x_n) et vérifier que $x_n \in H_0^1(]0,1[)$ et $J(x_n) \leq \frac{1}{4n^2}$.
b) Que vaut l'infimum de J sur $H_0^1(]0,1[)$? Montrer que cet infimum n'est pas atteint dans $H_0^1(]0,1[)$.

Exercice 5.46. *Soit Ω est un ouvert borné régulier de $\mathbb{R}^n$ et $f \in L^2(\Omega)$. On munit l'espace $H_0^1(\Omega)$ du produit scalaire:*

$$\langle u, v\rangle := \int_\Omega (\nabla u \cdot \nabla v + uv).$$

On pose $J(v) := \int_\Omega |\nabla v(x)|^2 dx - \int_\Omega f(x)v(x)dx$, $v \in H_0^1(\Omega)$. Montrer que $\nabla J(v)$ existe et est l'unique solution d'un problème aux limites que l'on précisera.

Exercice 5.47. *Soit Ω est un ouvert borné régulier de $\mathbb{R}^n$. Soit $J(v) := \int_\Omega |\nabla v(x)|^2 dx - \int_\Omega f(x)v(x)dx$ pour $v \in H_0^1(\Omega)$ où Ω est un ouvert borné de $\mathbb{R}^n$. Montrer que J est différentiable en tout point $u \in H_0^1(\Omega)$ et écrire l'équation d'Euler.*

Exercice 5.48. *Soit H un espace de Hilbert et K un convexe fermé non-vide, a une forme bilinéaire symétrique continue coercive sur H (i.e. $\exists\alpha > 0$ $\forall x \in H$, $a(x,x) \geq \alpha\|x\|^2$), et L une forme linéaire continue sur H. Montrer que $J(v) := \frac{1}{2}a(v,v) - L(v)$ admet un unique minimiseur dans K noté u qui est solution de l'inégalité variationnelle:*

$$\forall v \in K,\ a(u, v-u) \geq L(v-u).$$

Exercice 5.49. *On pose $H = H_0^1(\Omega)$ où Ω est un ouvert borné régulier de $\mathbb{R}^n$. Pour $u \in H$ on pose:*

$$J(u) = \int_\Omega |\nabla u(x)|^2 dx \quad \text{et} \quad h(u) = \int_\Omega u(x)^2 dx - 1.$$

Soit le problème de minimisation

$$\inf_{u\in H} J(u) \quad h(u) = 0.$$

1) Démontrer que le problème admet une solution (utiliser le Théorème de Rellich[3]).
2) Démontrer que J et h sont différentiables et calculer leur gradient respectif.
3) Appliquer KKT sur le problème

$$\inf_{u\in H_0^1(\Omega),\, h(u)=0} J(u).$$

4) Trouver l'expression du multiplicateur μ en fonction de la solution.
5) En utilisant que pour $v \in H_0^1(\Omega)$ on a $J(\frac{v}{\|v\|_{L^2(\Omega)}}) \geq J(u)$ déduire l'existence d'une constante de Poincaré que l'on exprimera à l'aide du multiplicateur.
6) Calculer la meilleure constante de Poincaré en dimension 1 sur $I = [0,1]$ en utilisant que le minimiseur vérifie

$$-u'' + \lambda u = 0, \quad u(0) = u(1) = 0.$$

et que donc λ est un multiple de π.

[3]Si Ω est un ouvert borné régulier de $\mathbb{R}^n$ alors l'injection canonique $H^1(\Omega)$ dans $L^2(\Omega)$ est compacte [9].

Exercice 5.50. *On garde les notations de l'exercice précédent. A l'aide de l'inégalité de Poincaré, montrer qu'étant donné $f \in L^2(\Omega)$, la fonctionnelle*

$$J(v) = \int_\Omega |\nabla v|^2 - \int_\Omega fv, \quad v \in H_0^1(\Omega)$$

est fortement convexe.

Exercice 5.51. *Soit Ω est un ouvert borné régulier de $\mathbb{R}^n$, $f \in L^2(\Omega)$, et $\varepsilon > 0$.*
1) Montrer que le problème

$$\inf_{v \in H_0^1(\Omega)} \int_\Omega \|\nabla v(x)\|^2 dx \quad \text{t.q.} \quad \|f - v\|_{L^2(\Omega)} \leq \varepsilon,$$

admet une unique solution u_ε.
2) Montrer que u_ε est une solution faible d'un problème que l'on précisera. Que peut-on dire si la contrainte n'est pas active?

5.6 Exercices autour de la dualité

Exercice 5.52. Dualité.
I) Résoudre par dualité le problème

$$\inf_{x^2+y^2\leq 1,\ y+z\leq 0} \frac{1}{2}((x-2)^2+y^2+z^2).$$

II) On considère le problème d'optimisation:

$$(\mathcal{P}) \qquad \inf_{Cx\leq d} \frac{1}{2}x^T Ax + b^T x$$

où A est une matrice de taille n symétrique définie positive, $b \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $C \in M_{ln}(\mathbb{R})$ et $d \in \mathbb{R}^l$. Montrer que le problème dual de $(\mathcal{P})$ est le problème:

$$\max_{\lambda\in\mathbb{R}^l_+} -\frac{1}{2}\lambda^T CA^{-1}C^T\lambda - (b^T A^{-1}C^T + d^T)\lambda.$$

Exercice 5.53. *Soit $A \in M_{m,n}(\mathbb{R})$ et le problème d'optimisation $\inf_{Ax\leq b} \langle c, x\rangle$ où $c \in \mathbb{R}^n$ et $b \in \mathbb{R}^m$. Montrer que le problème dual s'écrit:*

$$\sup_{A^T\lambda+c=0,\ \lambda\in\mathbb{R}^m_+} -\langle b, \lambda\rangle$$

De même dualiser le problème $\inf_{Ax=b\ x\geq 0} \langle c, x\rangle$ (on ne dualisera pas la contrainte "simple" $x \geq 0$).

Exercice 5.54. *On considère le problème de minimisation*

$$\inf_{Ax=b} \frac{1}{2}\langle Qx, x\rangle + \langle c, x\rangle,$$

où $Q \in \mathcal{S}_n^{++}(\mathbb{R})$ et $c \in \mathbb{R}^n$. Calculer la solution du problème en utilisant la dualité. Vérifier en résolvant directement le problème.

Exercice 5.55. *Soit $Q \in \mathcal{S}_n(\mathbb{R})$ et $\lambda_1(Q)$ sa plus grande valeur propre. Montrer que :*

$$\lambda_1(Q) = \max_{Tr(X)=1\ X\in\mathcal{S}_n^+(\mathbb{R})} Tr(QX)$$

(Indication: utiliser le théorème de dualité pour montrer que le dual de ce problème est $\max_{Q-yI_n\succeq 0} y$, puis qu'une matrice $M \in \mathcal{S}_n(\mathbb{R})$ vérifie $M \succeq 0$ si et seulement si pour tout $x \in \mathbb{R}^n$ $x^T Mx \geq 0$.)

5.7 Exercices autour des algorithmes numériques

Exercice 5.56. Autour de l'algorithme d'Uzawa.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe C^1, $C \in M_{m,n}(\mathbb{R})$ et $d \in \mathbb{R}^n$ et soit le problème de minimisation

$$\min_{Cx=d} f(x). \tag{P}$$

On rappelle que l'algorithme d'Uzawa correspond à l'algorithme du gradient appliqué au problème dual associé au problème (P)*. Il s'énonce comme suit. Soit $\varepsilon > 0$ fixé.*

- *Considérer $\lambda_0 \in \mathbb{R}^m$ pour l'initialisation et soit $\rho > 0$.*
- *A l'étape k, on appelle x_k une solution du problèmes sans contraintes:*

$$\min_{x\in\mathbb{R}^n} f(x) + \langle \lambda, Cx - d\rangle .$$

- *On pose $\lambda_{k+1} = \lambda_k + \rho(Cx_k - d)$.*
- *Si $\|x_{k+1} - x_k\| \leq \varepsilon$, alors l'algorithme s'arrête et on retourne x_{k+1}.*

Le but est de montrer le théorème suivant.

Théorème 5.3. *On suppose que f est de classe C^2 et elliptique de constante α i.e.:*

$$\exists \alpha > 0 \quad \forall x \in \mathbb{R}^n, \quad \left\langle D^2 f(x)v, v\right\rangle \geq \alpha\|v\|^2 \quad \forall v \in \mathbb{R}^n.$$

Soit $\rho \in]0, \frac{2\alpha}{\|C\|^2}[$[4]. Alors la suite (x_k) générée par l'algorithme converge vers l'unique minimum du problème (P)*.*

1) Démontrer que ce problème admet une unique solution.
2) Soit $x_0 \in \mathbb{R}^n$ un point qui réalise le minimum du problème

$$d(\lambda_0) := \min_{x\in\mathbb{R}^n} f(x) + \langle \lambda_0, Cx - d\rangle .$$

A l'aide du théorème de Danskin, montrer que $d(\lambda_0)$ est dérivable au point λ_0 et que $\nabla d(\lambda_0) = Cx_0 - d$ (montrer d'abord que $d(\lambda_0 + hv) - d(\lambda_0) \geq h\langle v, Cx_0 - d\rangle$, $h \in \mathbb{R}$, $v \in \mathbb{R}^m$).

3) a) En écrivant les conditions d'optimalité pour x^* montrer que $\lambda_{k+1} - \lambda^* = \lambda_k - \lambda^* + \rho C(x_k - x^*)$ (où λ^* est le multiplicateur de Lagrange associé à (P) *en x^*).*

b) En déduire que $\|\lambda_{k+1}-\lambda^*\|^2 \leq \|\lambda_k-\lambda^*\|^2-\rho(2\alpha-\rho\|C\|^2)\|x_k-x^*\|^2$ (utiliser que $\nabla f(x_k)-\nabla f(x^*) = -C^T(\lambda_k-\lambda^*)$ et l'ellipticité de f).

c) Déduire de b) que la suite $(\|\lambda_k - \lambda^*\|)$ converge.

d) Déduire à nouveau de b) que la suite (x_k) converge vers x^*.

Exercice 5.57. Descente de gradient dans le cas d'une fonction convexe.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe de classe C^1. On s'intéresse au problème

$$\min_{\mathbb{R}^n} f.$$

On considère une suite définie par:

$$x_{k+1} = x_k - \rho_k \frac{d_k}{\|d_k\|}, \; d_k = \nabla f(x_k),$$

où $\rho_k > 0$, $\rho_k \to 0$, et $\sum_{k\geq 0} \rho_k = +\infty$.
1) On suppose qu'il existe $\overline{x} \in \mathbb{R}^n$ tel que $f(\overline{x}) < \liminf_{k\to+\infty} f(x_k)$. Montrer que

$$\|x_{k+1} - \overline{x}\|^2 = \|x_k - \overline{x}\|^2 - 2\rho_k \left\langle \frac{d_k}{\|d_k\|}, x_k - \overline{x}\right\rangle + \rho_k^2.$$

2) Montrer qu'il existe $\alpha > 0$ et $k_0 \in \mathbb{N}$ tels que $f(x) \leq f(x_k)$ pour tout $k \geq k_0$ et $\|x - \overline{x}\| \leq \alpha$.
3) Déduire de la propriété de convexité (appliqué aux pointx x_k et $\overline{x} + \alpha\frac{d_k}{\|d_k\|}$) que $-2\rho_k \left\langle \frac{d_k}{\|d_k\|}, x_k - \overline{x}\right\rangle \leq -2\alpha\rho_k$, puis que $\|x_{k+1} - \overline{x}\|^2 \leq \|x_k - \overline{x}\|^2 - 2\alpha\rho_k + \rho_k^2$.
4) Conclure que $\liminf_{k\to+\infty} f(x_k) = \inf_{\mathbb{R}^n} f$ (on remarquera que $\rho_k \leq \alpha$ dès que $k \geq k_0$).

[4]Ici $\|C\|$ désigne la norme triple subordonnée de la matrice C.

Exercice 5.58. Convergence de l'algorithme du gradient projeté.
On suppose que la suite (x_k) générée par l'algorithme n'est pas un minimum local de f sur C (sinon on arrête l'algorithme de descente). Soit $d_k = \overline{x}_k - x_k$ et $\alpha_k = \beta^{m_k}$.
1) Montrer qu'un point $x^* \in C$ satisfait $\langle \nabla f(x^*), x - x^* \rangle \geq 0$ pour tout $x \in C$ si et seulement si x^* est la projection sur C de $x^* - s\nabla f(x^*)$.
2) a) Montrer que $x_k \in C$ pour tout k.
b) Montrer que $0 < \|d_k\|^2 \leq -s \langle \nabla f(x_k), d_k \rangle$ (utiliser que $x_k \in C$). En déduire que d_k est une direction de descente.
b) Soit (x_{k_i}) une suite extraite de la suite (x_k) qui converge vers un certain point $x^* \in C$. Montrer que

$$\lim_{i \to +\infty} \alpha_{k_i} \langle \nabla f(x_{k_i}), d_{k_i} \rangle = 0.$$

3) Montrer que $\langle \nabla f(x^*), d^* \rangle = 0$ où d^* est la limite de d_{k_i} (distinguer si la suite α_{k_i} converge vers 0 ou pas).
4) En déduire que $d^* = 0$ et par 1) que x^* est un point critique sur C de f.

Exercice 5.59. *Formule de Broyden. Soit $B_0 \in M_n(\mathbb{R})$ et $s, y \in \mathbb{R}^n$ avec $s \neq 0$. On rappelle qu'on utilise la norme de Frobenius sur les matrices. On considère le problème de minimisation:*

$$\min_{B \in M_n(\mathbb{R}),\ Bs=y} \frac{1}{2}\|B - B_0\|^2.$$

1) Etablir que l'unique solution de ce problème s'écrit $B = B_0 + \frac{(y-B_0 s)s^T)}{\|s\|^2}$ (écrire le Lagrangien du problème, et les conditions de Karusch-Kuhn-Tucker).
2) Etablir la formule PSB (Powell symmetric Broyden) si on rajoute la contrainte de symétrie dans les contraintes (i.e. $B = B^T$):

$$B = B_0 + \frac{(y - B_0 s)s^T) + s(y - B_0 s)T}{\|s\|^2} - \frac{\langle y - B_0 s, s \rangle}{\|s\|^4} ss^T$$

Exercice 5.60. *Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une fonction de classe C^1. Soit $\overline{x} \in \mathbb{R}^n$ tel que $f(\overline{x}) = 0$ et tel que $Df(\overline{x})$ soit inversible.*
1) Montrer que

$$f(x) = Df(\overline{x})(x - \overline{x}) + \int_0^1 [Df(\overline{x} + t(x - \overline{x})) - Df(\overline{x})](x - \overline{x})dt.$$

2) Soit $r > 0$ tel que si $y \in B(\overline{x}, r)$, alors $\|Df(y) - Df(\overline{x})\| \leq \frac{1}{2\|Df(\overline{x})\|^{-1}}$. Montrer que si $x \in B(\overline{x}, r)$, alors $f(x) \geq \frac{1}{2\|Df(\overline{x})\|^{-1}}\|x - \overline{x}\|$.

Exercice 5.61. *Méthodes de type quasi-Newton. Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une fonction de classe C^1. Soit $x^* \in \mathbb{R}^n$ tel que $f(x^*) = 0$ et tel que $Df(x^*)$ soit inversible. Au lieu de considérer la méthode de Newton usuelle $x_{k+1} = x_k - Df(x_k)^{-1}f(x_k)$, on s'intéresse à une suite x_k définie par $f(x_k) + M_k(x_{k+1} - x_k) = 0$ où la suite de matrice M_k est une approximation du jacobien de f au point x_k (on ne cherche pas d'expression de M_k dans cet exercice). On se propose de démontrer le critère de Dennis et Moré : la suite (x_k) converge sur-linéairement vers x^* (cad $\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \to 0$) ssi $\|(M_k - Df(x^*))(x_{k+1} - x_k)\| = o(\|x_{k+1} - x_k\|)$.*
1) a) Montrer que $f(x_{k+1}) = f(x_k) + Df(x^*)(x_{k+1} - x_k) + o(\|x_{k+1} - x_k\|)$ (utiliser le théorème des accroissements finis).
b) En déduire que $\|f(x_{k+1})\| = \|(M_k - Df(x^*))(x_{k+1} - x_k)\| + o(\|x_{k+1} - x_k\|)$.
2) On suppose le critère de Dennis et Moré vérifié. Montrer par l'exercice 3 que $\|x_{k+1} - x^*\| = o(\|x_{k+1} - x_k\|)$, puis conclure en utilisant l'inégalité triangulaire (dans le o).
3) On suppose la convergence sur-linéaire.
a) En utilisant l'inégalité des accroissements finis entre x_{k+1} et x^*, montrer que $\|(M_k - Df(x^*))(x_{k+1} - x_k)\| = o(\|x_k - x^*\|) + o(\|x_{k+1} - x_k\|)$.
b) Montrer que la convergence sur-linéaire implique $\|x_k - x^*\| = O(\|x_{k+1} - x_k\|)$, puis conclure.

5.8 TP sous matlab

Exercice 5.62. Choix du pas de descente. *1) Soit la fonction* $x \mapsto \frac{1}{2}x^2$. *On considère l'algorithme itératif:*

$$x_{k+1} = x_k - \alpha_k d_k, \quad x_0 = 2.$$

où $d_k := \frac{f'(x_k)}{|f'(x_k)|}$ *et* $\alpha_k := 2 + \frac{3}{2^{k+1}}$. *Représenter sous Matlab la suite* $(x_k)_{k\geq 0}$. *Vérifier que la suite* $(f(x_k))_{k\geq 0}$ *est décroissante. Interpréter en établissant une formule de récurrence pour la suite* $(x_k)_{k\ \geq 0}$[5].
2) Tracer et étudier de même la suite $(x_k)_{k\geq 0}$ *lorsque* $s_k := \frac{1}{2^{k+1}}$. *Que constate-on? En pratique, la règle d'Armijo (voir ci-dessous) permet d'éviter les pas trop grands et la règle de Wolfe permet d'éviter les pas trop petits.*

Exercice 5.63. Problème du conditionnement. *Le but est de représenter la suite générée par l'algorithme de descente de gradient (recherche exacte) pour une fonction quadratique. Soit* $q(x) = \frac{1}{2}\langle Qx, x\rangle - \langle b, x\rangle + a$ *où* Q *est une matrice symétrique définie positive de taille* n, $b \in \mathbb{R}^n$ *et* $a \in \mathbb{R}$.
1) Justifier que la méthode de descente de gradient avec recherche exacte (i.e. $\alpha_k \in \arg\min_{t\in\mathbb{R}} f(x_k + td_k)$ *avec ici* $d_k := -\nabla f(x_k)$*) s'écrit* $x_{k+1} = x_k - \alpha_k r_k$ *où* $\alpha_k = \|r_k\|^2 / \langle Qr_k, r_k\rangle$ *et* $r_k = Qx_k - b$.
2) Soit $f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2)$. *On considère le point initial* $(10, 1)$ *et on prend* $\gamma = 10$. *Représenter sur une même figure une vingtaine d'itérations de l'algorithme ainsi que les lignes de niveaux de la fonction* f. *Utiliser éventuellement* `meshgrid` *et* `contour` *pour représenter les lignes de niveau de* f *et la commande* `hold on`. *Que remarque-t-on si on prend* $\gamma = 2$*? Interpréter le résultat.*

Règle d'Armijo. Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1 et (P) le problème de minimisation $\inf_{x\in\mathbb{R}^n} f(x)$. On se donne des paramètres $s > 0$, $\sigma \in]0, 1[$, $\beta \in]0, 1[$ et $\varepsilon > 0$. On considère l'algorithme de descente basé sur la règle *d'Armijo*:

- Soit $x_0 \in \mathbb{R}^n$ et $d_0 \in \mathbb{R}^n$ tel que $\|d_0\| = 1$ (initialisation). Pour $k = 1, ..., N_{iter}$ faire les étapes suivantes:
- Prendre $d_k \in \mathcal{S}^{n-1}$ tel que $-\frac{\langle \nabla f(x_k), d_k\rangle}{\|\nabla f(x_k)\|} \in [\varepsilon, 1]$ et tel que $\|d_k\| = 1$.
- pour $i = 0, 1, ...$ soit i_k le premier indice $i \geq 0$ pour lequel on a $f(x_k) - f(x_k + \beta^i s d_k) \geq -\sigma\beta^i s \langle \nabla f(x_k), d_k\rangle$. Si $i = 0$ on pose $\alpha_k = s$ et si $s \geq 1$ on pose $\alpha_k = s\beta^{i_k}$.
- Faire $x_{k+1} = x_k + \alpha_k d_k$ avec $\alpha_k = s\beta^{i_k}$

Exercice 5.64. Programmation de la règle d'Armijo. *En utilisant la règle d'Armijo, on souhaite minimiser la fonction* $f(x_1, x_2) = e^{x_1+3x_2-0.1} + e^{x_1-3x_2-0.1} + e^{-x_1-0.1}$ *et avec comme direction de descente* $d_k = -\nabla f(x_k)$. *Soit* $\alpha = 0.1$ *et* $\beta = 0.7$. *L'algorithme s'écrit* $x_{k+1} = x_k + td_k$ *où pour chaque itération* k, t *est déterminé par la règle suivante: soit* $t = 1$ *; tant que* $f(x_k + td_k) > f(x_k) + \alpha t\nabla f(x_k) \cdot d_k$, *faire* $t \leftarrow \beta t$.
Programmer cette règle avec une dizaine d'itérations et en partant par exemple des points $(1, 1)$, $(-1, 1)$, $(1, -1)$. *On représentera sur une même figure les itérations de l'algorithme ainsi que les lignes de niveaux de la fonction* f *à l'aide des commandes* `[X,Y] = meshgrid(-1:0.4:1, -1:0.1:1)` *et* `contour(X,Y,Z,20)`.

Exercice 5.65. *On considère la fonction de Rosenbrock définie par* $f(x, y) = (x - 1)^2 + p(x^2 - y)^2$. *Cette fonction peut servir à tester l'efficacité de certains algorithmes de descente.*
1) Trouver le minimum de f *sur* $\mathbb{R}^2$. *Représenter* $z = f(x, y)$ *(à l'aide des commandes* `surface` *et* `view` *par exemple).*
2) Programmer la méthode du gradient à pas fixe avec $\alpha = 0.1$ *ou* $\alpha = 0.01$ *et* $p = 10$, *à partir du point* $(-1, 1)$. *Représenter sur une même figure les itérations et les lignes de niveau de* f. *Que constatez vous?*
3) Programmer la méthode du gradient à pas optimal (on utilisera les fonctions de Matlab `fminunc` *ou* `fminbnd` *à chaque itération afin de déterminer le pas de descente optimal). Représenter sur une même figure les itérations et les lignes de niveau de* f. *Que constatez vous?*

Exercice 5.66. Méthode du gradient conjugué. *On notera* $\langle x, y\rangle$ *ou* $x^T y$ *le produit scalaire dans* $\mathbb{R}^n$.
On souhaite minimiser la fonction quadratique $q(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$ *où* A *est symétrique définie positive de taille* n *et* $b \in \mathbb{R}^n$. *L'algorithme est le suivant: soit* x_0 *donné,* $g_0 := Ax_0 - b$, *et* $h_0 := g_0$. *Pour tout* $n \geq 1$, *on considère les itérations:*

$$\begin{cases} x_n = x_{n-1} - \alpha_n h_{n-1}, \ \alpha_n = \frac{h_{n-1} g_{n-1}}{h_{n-1}^T A h_{n-1}}, \\ g_n = Ax_n - b, \\ h_n = g_n + h_{n-1} \frac{g_n^T g_n}{g_{n-1}^T g_{n-1}}. \end{cases}$$

1) Implémenter l'algorithme avec `A=[1,0;0,5]`, `B=[1;2]`. *Tracer sur une même figure les lignes de niveau de* q *et les itérations de l'algorithme avec la méthode gradient conjugué et à pas optimal. Comparer les deux méthodes.*

[5] On montrera que $x_k = (-1)^k(1 + 1/2^k)$, $k \in \mathbb{N}$.

2) Question théorique. Essayer de généraliser l'algorithme ci-dessus dans le cas où la fonction $q(x)$ n'est plus quadratique (remplacer par le gradient de q).

Exercice 5.67. *Le but de l'exercice est de résoudre numériquement le problème aux limites*

$$\begin{cases} -u''(x) = f(x), \ x \in [0,1], \\ u(0) = u(1) = 0. \end{cases}$$

On se donne une subdivision de l'intervalle $[0,1]$ en N intervalles de longueur $h = 1/(N+1)$, et on considère les points $x_i = ih$ de l'intervalle $[0,1]$, $0 \leq i \leq N+1$. On cherche alors une solution approchée sous la forme du vecteur $u = (u_i)_{0 \leq i \leq N+1}$, où les u_i vérifient le schéma aux différences finies:

$$\begin{cases} \frac{1}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) = f(x_i), \ 1 \leq i \leq N, \\ u_0 = u_{N+1} = 0. \end{cases} \tag{5.8.1}$$

1) Justifier le schéma (5.8.1). Montrer que (5.8.1) est équivalent à $Au = f$ où A est une matrice de taille N, f un vecteur de $\mathbb{R}^N$ que l'on précisera.
2) Ecrire les itérations de l'algorithme de Newton (on rappelle qu'étant donné $f : \mathbb{R}^N \to \mathbb{R}^N$ de classe C^1, l'algorithme de Newton s'écrit $x_{k+1} = x_k - DF(x_k)^{-1}F(x_k)$ pourvu que $DF(x_k)$ soit inversible), et trouver une solution approchée du problème.
3) Représenter sur un même graphique la solution avec la méthode de Newton et la solution exacte de (5.8.1). On donnera pour chaque itération $\|F(x_k)\|$.

Exercice 5.68. *On s'intéresse à l'équation différentielle avec conditions aux limites:*

$$-y''(x) - \tanh(y(x)^2) + \frac{1}{2}y(x) = x^2, \ 0 < x < 1, \ y(0) = 0, \ y(1) = 0.$$

Le but de l'exercice est aussi de programmer une méthode de Newton basée sur les différences finies pour résoudre numériquement ce problème. Notations: soit $N \geq 1$, soit $h = \frac{1}{N+1}$, soit $x_i = ih$, $0 \leq i \leq N+1$ et $y_i = y(x_i)$, $0 \leq i \leq N$.
1) Justifier pourquoi on peut approcher $-y''(x_i)$ par $\frac{1}{h^2}[-y_{i+1} + 2y_i - y_{i-1}]$ et montrer que les points y_i sont reliées par les relations:

$$-\frac{1}{h^2}(-y_i + 2y_{i+1} - y_{i+2}) - \tanh(y_i^2) + \frac{1}{2}y_i - (ih)^2 = 0, \ 0 \leq i \leq N-1$$

avec la convention $y_0 = 0$ et $y_{N+1} = 0$. En déduire l'expression d'une fonction $F : \mathbb{R}^N \to \mathbb{R}^N$ dont on cherche un zéro.

2) Programmation de la méthode sous matlab à l'aide de la méthode de Newton.
a) Etant donné un point $y \in \mathbb{R}^N$, définir sous matlab la fonction F ainsi que la matrice DF. (N.B.: utiliser par ex. `gallery('tridiag',N,-1,2,-1)` *et* `ones(N,1)`*).*
b) Définir une fonction `y=Newton(N,Niter)` *qui définit la suite des itérées par la formule usuelle $y^{k+1} = y^k - DF^{-1}(y_k)F(y_k)$.*
c) Initialiser la méthode par le vecteur $y^0 = 0$ et fixer $N = 30$. Indiquer à chaque étape $\|F(y_k)\|$ (utiliser par exemple la commande `[i norm(F,1)]` *et tracer à chaque étape la solution par la commande* `plot(x,y)`*.*

3) Programmation de la méthode sous matlab à l'aide d'une méthode de type quasi-Newton et formule de Broyden.
a) Programme la fonction F et DF. N.B.: Utiliser par ex. `gallery('tridiag',N,-1,2,-1)` *et* `ones(N,1)`
b) Le but est de construire une suite de matrices M_k (méthode de quasi-Newton avec formule de Broyden.
c) Définir une fonction `y=Newton(N,Niter)` *qui définit la suite des itérées par la formule*

$$y^{k+1} = y^k - M_k^{-1}F(y_k),$$

Exercice 5.69. *1) Programmer l'algorithme d'Uzawa pour un problème d'optimisation général du type:*

$$\inf_{Cx \leq d} \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle + c$$

où $A \in M_n(\mathbb{R})$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$, $x \in \mathbb{R}^n$. Vérifier la convergence de l'algorithme pour le problème $\min_{Cx=d} \frac{1}{2}\|x - x_0\|^2$ où $x_0 \in \mathbb{R}^n$ et $Cx = d$ est l'équation d'un hyperplan (prendre par exemple $n = 4$, $C = [1,1,1,1]$ $d = 1$ et $x_0 \in \mathbb{R}^4$. Comparer avec la solution théorique $\frac{\|d - Cx_0\|}{\|C\|}$.
2) Etudier le problème $\inf_{u \in H_0^1([0,1])} \int_0^1 [\frac{1}{2}u'^2(x) - f(x)u(x)]dx$ sous la contrainte $u(x) \leq h(x)$ et $u(0) = u(1) = 0$ en discrétisant et en utilisant la méthode d'Uzawa.

Chapitre 6

Sujets d'examen

6.1 Session 2015-2016

6.1.1 Partiel (1h)

Exercice 6.1. *Soit $a > 0$, $b > 0$ et $c > 0$. On cherche le parallélépipède de volume* **maximal** *inclu dans l'ellipsoïde d'équation*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1.$$

1) Montrer l'existence d'une solution au problème (le volume d'un parallélépipède peut s'écrire $V(x, y, z) = 8xyz$).
2) Montrer qu'une solution $(\bar{x}, \bar{y}, \bar{z}) \in \mathbb{R}^3$ du problème vérifie $\bar{x} > 0$, $\bar{y} > 0$ et $\bar{z} > 0$.
3) Justifier que le problème est qualifié en tout point de C (où $C := \{(x, y, z) \in \mathbb{R}^3 \; ; \; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1\}$.
4) Ecrire les conditions d'optimalité KKT et en déduire l'unique solution $(\bar{x}, \bar{y}, \bar{z})$ du problème.

Exercice 6.2. *Soit $A \in M_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}^n$. Etant donné $x \in \mathbb{R}^n$, on notera $x = (x_1, ..., x_n)$ ses coordonnées dans la base canonique $(e_i)_{1 \leq i \leq n}$ de $\mathbb{R}^n$. L'écriture $x \geq 0$ équivaut à $x_i \geq 0$ pour $1 \leq i \leq n$. Dans cet exercice, on considère le problème de minimisation:*

$$\inf_{x \in C} \langle c, x \rangle \,, \tag{6.1.1}$$

où $C = \{x \in \mathbb{R}^n \; ; \; x \geq 0 \text{ et } Ax = b\}$. On suppose que $C \neq \emptyset$ et que $\alpha := \inf_{x \in C} \langle c, x \rangle > -\infty$. Le but de l'exercice est de montrer que le problème de minimisation (6.1.1) admet une solution.
1) a) Soit $B \in M_{m+1,n}(\mathbb{R})$ la matrice définie par $B = \begin{pmatrix} c \\ A \end{pmatrix}$ et K le cône convexe engendré par les n vecteurs $Be_i \in \mathbb{R}^{m+1}$, $1 \leq i \leq n$. Ecrire la définition de K.
b) Soit (x^k) une suite minimisante pour le problème (6.1.1). Ecrire les propriétés vérifiées par la suite (x^k).
c) Vérifier que $Bx^k \in K$ pour tout $k \in \mathbb{N}$.
2) On rappelle que K est **fermé**. *Montrer qu'il existe $(\lambda_1, ..., \lambda_n) \in \mathbb{R}^n_+$ tel que*

$$\sum_{1 \leq i \leq n} \lambda_i B e_i = \begin{pmatrix} \alpha \\ b \end{pmatrix}.$$

En déduire que λ est une solution du problème de minimisation.

Exercice 6.3. *(Il s'agit de deux questions* **indépendantes** *autour du lemme de Farkas).*
1) Démontrer la réciproque du lemme de Farkas.
2) Soit c, c_i, $1 \leq i \leq l$ et d_j, $1 \leq j \leq m$ des vecteurs de $\mathbb{R}^n$. On fait l'hypothèse suivante:

$$\forall v \in \mathbb{R}^n, \;\; \forall\, 1 \leq i \leq l, \; \langle c_i, v \rangle \geq 0 \quad \text{et} \quad \forall\, 1 \leq j \leq m, \; \langle d_j, v \rangle = 0 \quad \Longrightarrow \langle c, v \rangle \geq 0.$$

Montrer qu'il existe $\lambda_i \geq 0$ *pour* $1 \leq i \leq l$ *et* $\mu_j \in \mathbb{R}$ *pour* $1 \leq j \leq m$ *tels que:*

$$c = \sum_{1 \leq i \leq l} \lambda_i c_i + \sum_{1 \leq j \leq m} \mu_j d_j.$$

6.1.2 Examen final (2h)

Exercice 6.4. *On cherche à résoudre le problème*

$$\min_{(x,y,z) \in K} x^2 + y^2 + z^2 \quad \text{avec} \quad K = \{(x,y,z) \in \mathbb{R}^3 \; ; \; x+y+z = 1 \text{ et } x^2+y^2 \leq 1\}.$$

1) Montrer que ce problème admet une unique solution.

2) Montrer que la contrainte est qualifiée en tout point.

3) Trouver la solution du problème.

4) Question plus difficile : vérifier que l'on trouve la même solution en utilisant la méthode de dualité (calculer d'abord l'infimum du Lagrangien en fonction des multiplicateurs de Lagrange).

Exercice 6.5. *Soit* $h : [0,1] \times \mathbb{R} \to \mathbb{R}$ *une fonction continue minorée sur* $[0,1] \times \mathbb{R}$*,* a *et* b *deux réels, et* E *l'ensemble défini par:*

$$E = \{x \in H^1([0,1], \mathbb{R}) \; ; \; x(0) = a, \; x(1) = b\}.$$

On définit une fonctionnelle J *sur* $H^1([0,1], \mathbb{R})$ *par :*

$$J(x) = \int_0^1 \left(\frac{1}{2} x'(t)^2 + h(t, x(t)) \right) dt.$$

Le but de l'exercice est d'étudier le problème $\min_{x \in E} J(x)$ *: existence de solutions et conditions d'optimalité.*

1) a) Soit (x_n) *une suite minimisante de* E*. Montrer que* (x'_n) *est bornée dans* $L^2([0,1], \mathbb{R})$*.*

b) En déduire l'existence d'une fonction v *dans* $L^2([0,1], \mathbb{R})$ *et d'une sous-suite* (x'_{n_k}) *qui converge faiblement vers* v *dans* $L^2([0,1], \mathbb{R})$*.*

c) Démontrer que (x_{n_k}) *converge simplement sur* $[0,1]$ *vers une fonction* $\bar{x} \in E$ *que l'on précisera.*

d) On admet que (x_{n_k}) *converge uniformément sur* $[0,1]$ *vers* $\bar{x}$*. Montrer l'existence d'une solution au problème* $\min_{x \in E} J(x)$*.*

2) On suppose que h *de classe* C^1 *et on admet qu'un minimum du problème* $\bar{x} \in E$ *est de classe* C^2 *sur* $[0,1]$*. Donner une équation différentielle vérifiée par* $\bar{x}$*.*

3) Question difficile (bonus) : montrer à l'aide du lemme de Dubois-Raymond que $\bar{x}$ *est bien de classe* C^2*.*

Exercice 6.6. *Soit* p *et* n *deux entiers non nuls tels que* $p < n$*. On considère* A *une matrice symétrique définie positive de taille* n*,* $b \in \mathbb{R}^n$*,* $d \in \mathbb{R}^p$*, et* $C \in M_{p,n}(\mathbb{R})$ *une matrice à* p *lignes et* n *colonnes surjective. On rappelle que la notation* x^T *désigne le vecteur transposé de* x*. On s'intéresse au problème d'optimisation*

$$\min_{x \in \mathbb{R}^n, \; Cx = d} \frac{1}{2} x^T A x - b^T x.$$

1) Montrer que le problème admet une unique solution x^* *et qu'il existe* $\lambda^* \in \mathbb{R}^p$ *tel que*

$$Ax^* + C^T \lambda^* = b.$$

Le but de la suite de l'exercice est d'étudier l'algorithme suivant. Soit $\rho_1 > 0$ *et* $\rho_2 > 0$ *fixés. Pour* $k = 0$ *on choisit* $x_0 \in \mathbb{R}^n$ *et* $\lambda_0 \in \mathbb{R}^p$*. A l'étape* k *on effectue l'opération:*

$$\begin{cases} x_{k+1} = x_k - \rho_1 (Ax_k - b + C^T \lambda_k), \\ \lambda_{k+1} = \lambda_k + \rho_1 \rho_2 (Cx_{k+1} - d). \end{cases}$$

2) Dans toute la suite, on utilisera la notation I_m pour désigner la matrice identité de $\mathbb{R}^m$ où $m \in \mathbb{N}^*$. Montrer que

$$\begin{cases} x_{k+1} - x^* = (I_n - \rho_1 A)(x_k - x^*) - \rho_1 C^T(\lambda_k - \lambda^*), \\ \lambda_{k+1} - \lambda^* = \lambda_k - \lambda^* + \rho_1\rho_2 C(x_{k+1} - x^*). \end{cases}$$

3) a) En déduire que :

$$\begin{pmatrix} I_n & 0 \\ -\rho_1\rho_2 C & I_p \end{pmatrix} \begin{pmatrix} x_{k+1} - x^* \\ \lambda_{k+1} - \lambda^* \end{pmatrix} = \begin{pmatrix} I_n - \rho_1 A & -\rho_1 C^T \\ 0 & I_p \end{pmatrix} \begin{pmatrix} x_k - x^* \\ \lambda_k - \lambda^* \end{pmatrix}.$$

(dans la notation par bloc ci-dessus 0 désigne la matrice nulle de $M_{n,p}(\mathbb{R})$).

b) Déduire de l'égalité précédente l'égalité:

$$\begin{pmatrix} x_{k+1} - x^* \\ \lambda_{k+1} - \lambda^* \end{pmatrix} = J \begin{pmatrix} x_k - x^* \\ \lambda_k - \lambda^* \end{pmatrix} \quad \text{avec} \quad J = \begin{pmatrix} I_n - \rho_1 A & -\rho_1 C^T \\ \rho_1\rho_2 C(I_n - \rho_1 A) & -\rho_1^2\rho_2 CC^T + I_p \end{pmatrix} \in M_{n+p,n+p}(\mathbb{R}).$$

4) On note $|||\cdot|||$ la norme matricielle subordonnée à la norme euclidienne sur $\mathbb{R}^n$: étant donnée $B \in M_n(\mathbb{R})$,

$$|||B||| = \sup_{\|x\| \leq 1} \|Bx\|,$$

où $\|\cdot\|$ désigne la norme euclidienne sur $\mathbb{R}^n$.

a) Montrer que si ρ_1 et ρ_2 sont suffisamment petits alors $|||I_n - \rho_1 A||| < 1$ et $\left|\left|\left|-\rho_1^2\rho_2 CC^T + I_n\right|\right|\right| < 1$ (utiliser la diagonalisation en base orthonormale).

b) En déduire que $|||J||| < 1$ (question plus difficile).

c) Déduire de la question 4)b) la convergence de l'algorithme.

6.2 Session 2016-2017

6.2.1 Partiel

Exercice 6.7. *(10 points). Soit H un espace de Hilbert sur $\mathbb{R}$ muni d'un produit scalaire $\langle\cdot,\cdot\rangle$. On note $\|x\| := \sqrt{\langle x, x\rangle}$ la norme d'un vecteur $x \in H$. Soit $f : H \to \mathbb{R}$ une application de classe C^1, fortement convexe de paramètre $\alpha > 0$ et telle que l'application $x \mapsto \nabla f(x)$ soit L-Lipschitzienne sur H.*

1) Montrer que f admet un unique minimum x^* sur H.

On considère l'algorithme suivant. Soit $\rho > 0$. Un point $x_0 \in H$ étant donné, on définit une suite $(x_k)_{k\in\mathbb{N}} \in H^{\mathbb{N}}$ par:

$$x_{k+1} = x_k - \rho\nabla f(x_k), \quad k \in \mathbb{N}.$$

2) Soit $\psi : H \to \mathbb{R}$ l'application définie par $\psi(x) = x - \rho\nabla f(x)$. Montrer l'inégalité:

$$\forall x \in H,\ \forall y \in H, \quad \|\psi(x) - \psi(y)\|^2 \leq (\rho^2 L^2 - 2\alpha\rho + 1)\|x - y\|^2.$$

3) En déduire qu'il existe $\rho_0 > 0$ tel que pour tout $\rho \in]0, \rho_0]$ alors il existe $\lambda \in]0, 1[$ tel que:

$$\forall x \in H,\ \forall y \in H, \quad \|\psi(x) - \psi(y)\| \leq \lambda\|x - y\|.$$

4) Soit $\rho \in]0, \rho_0]$ et $\lambda \in]0, 1[$ comme dans la question 3). Démontrer que la suite $(x_k)_{k\in\mathbb{N}}$ définie à la question 1) converge fortement vers x^* (Indication: utiliser le caractère contractant de ψ).

Exercice 6.8. *(12 points). Soit $n \in \mathbb{N}^*$, soit Q une matrice symétrique de taille n, soit $c \in \mathbb{R}^n$, et soit $\Delta \in \mathbb{R}_+^*$. On notera I_n la matrice identité de taille n. Soit $f : \mathbb{R}^n \to \mathbb{R}$ l'application définie par*

$$f(x) = \frac{1}{2}\langle Qx, x\rangle + \langle c, x\rangle,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire usuel de $\mathbb{R}^n$. Soit K le sous-ensemble de $\mathbb{R}^n$ défini par $K = \{x \in \mathbb{R}^n \ ; \ \|x\| \leq \Delta\}$ où pour $x \in \mathbb{R}^n$, $\|x\| = \sqrt{\langle x, x \rangle}$. On considère dans cet exercice le problème d'optimisation:

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

1) Justifier l'existence d'un minimum global x^* au problème (P) et montrer que la contrainte est qualifiée en x^*.

2) En écrivant la contrainte K sous la forme $K = \{x \in \mathbb{R}^n \ ; \ \frac{1}{2}\left(\|x\|^2 - \Delta^2\right) \leq 0\}$, écrire les deux conditions de Karush-Kuhn-Tucker (KKT) au point x^* (on notera λ le multiplicateur associé à la contrainte d'inégalité).

3) En utilisant que f est quadratique et la question 2), montrer l'égalité:

$$\forall x \in \mathbb{R}^n, \quad f(x) - f(x^*) = \frac{1}{2} \langle (Q + \lambda I_n)(x - x^*), x - x^* \rangle - \frac{\lambda}{2}(\|x\|^2 - \|x^*\|^2).$$

4) On suppose que $\|x^*\| = \Delta$ (uniquement dans les questions 4a) et 4b)).

a) Montrer que si $x \in \mathbb{R}^n$ vérifie $\|x\| = \Delta$, alors $\langle (Q + \lambda I_n)d, d \rangle \geq 0$ où $d = x - x^*$.

b) En déduire que la matrice $Q + \lambda I_n$ est positive.

5) On suppose que $\|x^*\| < \Delta$ (uniquement dans cette question). Démontrer que la matrice Q est positive.

6) Réciproquement, on suppose que les deux conditions de KKT sont vérifiées en un certain point $(x^*, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+$ et que la matrice $Q + \lambda I_n$ est positive. Démontrer, en utilisant ce qui précède, que le point x^* est un minimum du problème (P) (Indication: distinguer selon que $\|x^*\| < \Delta$ ou $\|x^*\| = \Delta$).

Correction:

Exercice 1. 1) On sait que il existe $\alpha > 0$ et $\beta \in \mathbb{R}$ tels que $f(x) \geq \alpha\|x\|^2 + \beta$ ce qui entraine f coercive. Ainsi, toute suite minimisante (x_n) est bornée. Ainsi, à une sous-suite près (x_n) converge faiblement vers un certain point x^* qui vérifie (car f est continue) : $f(x^*) \leq \liminf f(x_n) = \inf f$. D'où le résultat. Notez que l'on peut aussi refaire la preuve du cours avec les suites de Cauchy. L'unicité provient du caractère strictement convexe.
2) En utilisant que $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \alpha \|x - y\|^2$ et que $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ pour tout x, y, on déduit le résultat en développant $\|x - y - \rho(\nabla f(x) - \nabla f(y))\|^2$.
3) On pose $\alpha := \rho^2 L^2 - 2\alpha\rho + 1$. Il suffit de garantir $\alpha \in]0, 1[$. La quantité $\rho^2 L^2 - 2\alpha\rho + 1$ tend vers 1 lorsque ρ tend vers 0, donc $\alpha > 0$ pour ρ suffisamment petit. De même, on a $\alpha < 1$ si et seulement si $\rho < 2\alpha/L$. D'où le résultat en prenant $\lambda := \sqrt{\alpha}$ pour ρ suffisamment petit.
4) C'est le théorème du point fixe dans un espace complet (en effet pour ρ suffisamment petit, ψ est contractante). Ainsi, (x_n) qui vérifie $x_{n+1} = \psi(x_n)$ converge vers l'unique point fixe de ψ. Un tel point fixe vérifie $\nabla f(x) = 0$. Or f admet un unique minimum global x^* qui vérifie $\nabla f(x^*) = 0$. Ainsi, (x_n) converge vers x^*.

Exercice 2.
1) Fonction continue sur un compact de de $\mathbb{R}^n$. Contrainte convexe et $g(0) < 0$ où $g(x) = \frac{1}{2}(\|x\|^2 - \Delta^2)$. Donc qualifiée.
2) KKT : $Qx^* + c + \lambda x^* = 0$ et $\lambda(\|x^*\|^2 - \Delta^2) = 0$ (pour la première relation, cela s'écrit aussi $\nabla f(x^*) + \lambda x^* = 0$).
3) a) f quadratique, donc Taylor à l'ordre 2 est exact et on remplace $\nabla f(x^*)$ en utilisant KKT:

$$f(x) - f(x^*) = \langle \nabla f(x^*), x - x^* \rangle + \frac{1}{2} D^2 f(x^*)(x - x^*, x - x^*) = -\lambda \langle x^*, x - x^* \rangle + \frac{1}{2} \langle Q(x - x^*), x - x^* \rangle.$$

Or on a $\frac{\lambda}{2}\|x - x^*\|^2 - \frac{\lambda}{2}\|x\|^2 + \frac{\lambda}{2}\|x^*\|^2 = -\lambda \langle x^*, x - x^* \rangle$ d'où l'égalité $f(x) - f(x^*) = \frac{1}{2} \langle (Q + \lambda I_n)(x - x^*), x - x^* \rangle - \frac{\lambda}{2}(\|x\|^2 - \|x^*\|^2)$.
b) N'oublions pas que $f(x) \geq f(x^*)$ car x^* est un minimum global de f sur K. Si $\|x\| = \Delta$ alors $f(x) - f(x^*) = \frac{1}{2} \langle (Q + \lambda I_n)(x - x^*), x - x^* \rangle \geq 0$.
c) La matrice $Q + \lambda I_n$ est symétrique, donc diagonalisable en base orthonormée. Soit $(\varepsilon_j)_{1 \leq j \leq n}$ une telle base. Alors, $\langle (Q + \lambda I_n)\varepsilon_j, \varepsilon_j \rangle = \frac{1}{\Delta^2} \langle (Q + \lambda I_n)\Delta\varepsilon_j, \Delta\varepsilon_j \rangle \geq 0$, donc $Q + \lambda I_n$ est bien positive.
4) Alors par complémentarité on a $\lambda = 0$ et on conclue directement que $f(x) - f(x^*) = \frac{1}{2} \langle Q(x - x^*), x - x^* \rangle \geq 0$ pour tout $x \in \mathbb{R}^n$, ce qui permet de conclure que Q est positive (en fait, si $\lambda = 0$, la contrainte n'est pas active, donc la théorie du second ordre sans contrainte montre directement que $D^2 f(x^*) = Q$ est positive.
5) On utilise $f(x) - f(x^*) = \frac{1}{2} \langle (Q + \lambda I_n)(x - x^*), x - x^* \rangle - \frac{\lambda}{2}(\|x\|^2 - \|x^*\|^2)$. Si $\|x^*\| < \Delta$ alors $\lambda = 0$ (par la condition de complémentarité), donc x^* est un minimum global sur $\mathbb{R}^n$. Si $\|x^*\| = \lambda$ alors $\lambda \geq 0$ et l'égalité ci-dessus montre que $f(x) - f(x^*) \geq 0$ pour tout $x \in \mathbb{R}^n$ tel que $\|x\| \leq \Delta$. Donc x^* est un minimum de f sur la boule fermée de centre 0, de rayon Δ (dans les deux cas).

6.3 Devoir Maison

6.3.1 Devoir Maison sur l'algorithme de Newton

Le but est d'estimer le nombre d'itérations dans l'algorithme de Newton pour avoir une précision donnée.

Notations. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe C^2, convexe. On note $\nabla f(x)$ le gradient de f en un point x et $\nabla^2 f(x)$ la hessienne de f en un point x. Le produit scalaire est noté $u^T v$ où $u, v \in \mathbb{R}^n$ et Id désigne la matrice identité de $\mathbb{R}^n$. On suppose que:

- La fonction f est strictement convexe sur $L_0 = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}$. De plus, on suppose qu'il existe $m > 0$ tel que pour tout $x \in L_0$, $\nabla^2 f(x) \geq mId$ (autrement dit la forme quadratique associée à $\nabla^2 f(x)$ vérifie pour tout $h \in \mathbb{R}^n$, $h^T \nabla^2 f(x) h \geq m\|h\|^2$).
- Il existe $L > 0$ tel que pour tout $x, y \in \mathbb{R}^n$, $\|\nabla^2 f(y) - \nabla^2 f(x)\| \leq L\|x - y\|$.
- Pour $x \in \mathbb{R}^n$, on note $\lambda(x) \in \mathbb{R}$ l'incrément de Newton défini par $\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{\frac{1}{2}}$.

Soit $x_0 \in L_0$. On considère l'algorithme de Newton suivant : $x_{k+1} = x_k + \alpha_k d_k$, $k \geq 0$, où $d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$ et α_k est déterminé par la règle d'Armijo (il ne s'agit pas exactement de l'algorithme de Newton "pur" vu en cours). On suppose $0 < \sigma < \frac{1}{2}$ dans la règle d'Armijo. On notera $\lambda_k = \lambda(x_k)$.

Partie I (Préliminaires).
1) Montrer que pour tout $x, y \in \mathbb{R}^n$ il existe $z \in [x, y]$ tel que $f(y) = f(x) + \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x)$. En déduire que pour que pour tout $x, y \in \mathbb{R}^n$, $f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{m}{2}\|y - x\|^2$.
2) Montrer que L_0 est borné. En déduire qu'il existe $M > 0$ tel que pour tout x dans L_0, $\nabla^2 f(x) \leq MId$. En déduire que pour tout $x, y \in \mathbb{R}^n$, on a : $f(y) \leq f(x) + \nabla f(x)^T(y - x) + \frac{M}{2}\|y - x\|^2$.
3)a) Montrer que $\lambda_k^2 = d_k^T \nabla^2 f(x_k) d_k \geq m\|d_k\|^2$
b) Montrer que $\lambda_k^2 \geq \frac{1}{M}\|\nabla f(x_k)\|^2$ (penser à utiliser 2) en utilisant l'inverse de la hessienne de f).
4) Soit $g : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1 qui vérifie pour tout $x, y \in \mathbb{R}^n$, $\|\nabla g(x) - \nabla g(y)\| \leq L\|x - y\|$. Montrer que pour tout $x, y \in \mathbb{R}^n$, on a: $\|g(x + y) - g(x) - \nabla g(x)^T y\| \leq \frac{L}{2}\|y\|^2$.
5) Montrer que f atteint son minimum sur $\mathbb{R}^n$. Dans la suite, on note f^* cette valeur et x^* le point tel que $f^* = f(x^*)$.

Partie II. On suppose qu'il existe $0 < \eta < \frac{m^2}{L}$ tel que à l'étape k, $\|\nabla f(x_k)\| \geq \eta$ (1) (intuitivement, ce cas couvre les étapes où l'algorithme est loin de la solution x^*). Soit $\gamma = \sigma\beta\eta^2 \frac{m}{M^2}$
1) Montrer en utilisant I2) que pour tout $\alpha > 0$, $f(x_k + \alpha d_k) \leq f(x_k) - \alpha\lambda_k^2 + \frac{\alpha^2 M}{2}\|d_k\|^2$.
2) Montrer en utilisant I3) puis en prenant $\alpha = \frac{m}{M}$ que $f(x_k + \alpha d_k) \leq f(x_k) - \frac{m}{2M}\lambda_k^2 < f(x_k) - \sigma\frac{m}{M}\lambda_k^2$.
3) Montrer que le pas dans la règle d'Armijo $\alpha_k \geq \frac{m}{M} \geq \beta\frac{m}{M}$ et montrer qu'une fois la recherche du pas terminé dans la règle d'Armijo, on a $f(x_k + \alpha_k d_k) < f(x_k) - \sigma\alpha_k\lambda_k^2 \leq f(x_k) - \sigma\beta\frac{m}{M}\lambda_k^2$.
4) En utilisant I3), déduire que $f(x_k + \alpha_k d_k) < f(x_k) - \frac{\sigma\beta m}{M}\eta^2$.
5) Montrer que le nombre d'itérations où l'hypothèse (1) est vérifiée est majoré par $\frac{1}{\gamma}(f(x_0) - f^*)$.

Partie III. On suppose qu'il existe $0 < \eta < \frac{m^2}{L}$ tel que à l'étape k, $\|\nabla f(x_k)\| < \eta$ (2) (intuitivement, ce cas couvre les étapes où la suite x_k est proche de x^*).
1) Montrer en utilisant I4) avec $x \longmapsto \nabla f(x)$ que $\|\nabla f(x_{k+1})\| \leq \frac{L}{2}\|\nabla^2 f(x_k)^{-1}\nabla f(x_k)\|^2$.
2) En déduire que $\|\nabla f(x_{k+1})\| \leq \frac{L}{2m^2}\|\nabla f(x_k)\|^2 \leq \eta$.
3) Montrer par récurrence que $\frac{L}{2m^2}\|\nabla f(x_{K+1})\| \leq \left[\frac{L}{2m^2}\|\nabla f(x_k)\|\right]^{2^{K-k}}$, $K \geq 1$.
4) Soit $\varepsilon > 0$. Donner une estimation (en fonction de ε) sur le nombre d'itérations nécessaires dans ce cas pour que $f(x) - f^* \leq \varepsilon$ (utiliser par exemple I2).
5) Conclusion. Donner une estimation totale sur le nombre d'itérations à effectuer pour que l'algorithme de Newton fournisse un point x tel que $f(x) \leq f^* + \varepsilon$.

Références

[1] G. Allaire, *Analyse numérique et optimisation*, Editions de l'ecole Polytechnique, 2005.

[2] D. Azé, J.-B. Hiriart-Urruty, *Analyse variationnelle et optimisation*, ellipses, 2010.

[3] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[4] M. Bergounioux, *Optimisation et contrôle des systèmes linéaires*, Dunod, 2001.

[5] D. M. Bertsekas, *Dynamic Prorgamming and Stochastic Control*, cours en ligne du MIT, 2011.

[6] F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer, 2000.

[7] F. Bonnans, *Optimisation Continue, Cours et Problèmes Corrigés*, Dunod, 2006.

[8] F. Bonnans, J.-C. Gilbert, C. Lemaréchal, C. Sagastizábal, *Numerical Optimization: Theoretical And Practical Aspects.*

[9] H. Brézis, *Analyse fonctionnelle Théorie et Applications*, Masson, 1987.

[10] P.G. Ciarlet, *Introduction à l'analyse matricielle et à l'optimisation*, 5ème édition, Dunod, 2007.

[11] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[12] O. Güler, *Foundations of Optimization*, Springer, 2010.

[13] J.-C. Culioli, *Introduction à l'optimisation*, Ellipses, 1994.

[14] J.-B. Hiriart-Urruty, *Les mathématiques du mieux faire, vol. 1, premiers pas en optimisation*, ellipses, 2010.

[15] J.-B. Hiriart-Urruty, *Optimisation et analyse convexe*, Mathématiques, PUF, 1998.

[16] J.-B. Hiriart-Urruty, *L'optimisation*, Que sais-je?, PUF, 1996.

[17] M. Minoux, *Programmation mathématique, théorie et algorithmes (tomes 1 et 2)*, Dunod 1983.

[18] T. Rockafellar, *Fundamentals of Optimization*, Lectures Note 2007, http://www.math.washington.edu/∼ rtr/mypage.html

[19] A. Ruszczynski, *Nonlinear Optimization*, Princeton University Press, 2006.