كلية العلوم و التقنيات فاس
Faculté des Sciences et Techniques de Fès

جامعة سيدي محمد بن عبد الله
Université Sidi Mohamed Ben Abdellah

Licence Mathématiques et Applications

(MA)

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du Diplôme de Licence Sciences et Techniques (LST)

Le sous-différentiel au sens de l'analyse convexe et le sous-différentiel au sens de Clarke

Réalisé par : **Assarrar Abdelghani**

Encadré par: **Pr. El HILALI ALAOUI Ahmed et Pr. KADRI Nasser**

Soutenu le 08/06/2018

Devant le jury composé de:

- **-Pr. EL HILALI ALAOUI Ahmed**
- **-Pr. EZZAKI Fatima**
- **-Pr. EL KHOMSSI Mohammed**
- **-Pr. HILALI Abdelmajid**
- **-Pr. KADRI Nasser**

Année Universitaire 2017 / 2018

FACULTE DES SCIENCES ET TECHNIQUES FES – SAISS

B.P. 2202 – Route d'Imouzzer – FES

Table des matières

REMERCIEMENTS

J'aimerais, avant tout remercier mes deux encadrant **Prof. M.El Hilali Alaoui Ahmed et Prof. M.El Kadri Nasser**, pour m'avoir accueilli et pour avoir mis à ma disposition tout ce qui a été nécessaire à la bonne réalisation de ce travail qui ne pourrait pas avoir été menée à terme sans leur confiance, patience et générosité, Ils m'ont encadré avec efficacité, clairvoyance et beaucoup de patience durant tout le temps que m'a exigé ce travail. Je leur exprime ma profonde reconnaissance pour leur soutien et encouragement face à la difficulté mais aussi pour leur disponibilité et le temps passé à la relecture et la correction de ce mémoire.

Je tiens également à remercier **Pr. Ezzaki Fatima, Pr. Hilali Abdelmajid ET Pr. El Khomssi Mohammed** pour avoir accepté d'évaluer ce travail et d'être membres du jury.

Je voudrais, enfin profiter de l'occasion pour exprimer mes remerciements à mes collègues et mes enseignants du départements de mathématiques de la Faculté des Sciences et Techniques FES pour leur encouragements, mes parents, mes frères et sœurs qui partagent avec moi les bonheurs et les tristesses de la vie .

1 Introduction

En mathématiques, et plus précisément en analyse convexe, le sous-différentiel est un concept permettant de décrire la variation locale d'une fonction convexe non nécessairement différentiable dans un sens classique ; celui au quel on attache aujourd'hui le nom de Fréchet. Au lieu d'être la pente de l'application linéaire tangente (c'est à dire la dérivée) au point considéré, qui n'existe pas nécessairement, le sous-différentiel d'une fonction convexe est l'ensemble des pentes de tous les minorants affines de la fonction , qui sont exactes en ce point ,c'est à dire qui en ce point la même valeur que la fonction convexe qu'elle minorent.

On sait que la notion de dérivée est fondamentale en analyse car elle permet d'approcher localement des fonctions par des modèles linéaires, plus simples à étudier. Ces modèles fournissent des renseignements sur les fonctions qu'ils approchent, si bien que de nombreuses questions d'analyse passent par l'étude des fonctions linéarisées (stabilité, inversibilité locale, etc). On rencontre beaucoup de fonctions convexes qui ne sont pas différentiables au sens classique, en particulier lorsque celles-ci résultent de constructions qui n'ont rien pour assurer la différentiabilité des fonctions qu'elles produisent. Il en est ainsi de la fonction duale (en) associée à un problème d'optimisation sous contraintes, pour en citer un exemple emblématique. Pour ces fonctions convexes non lisses, le sous-différentiel joue donc un rôle similaire à celui de la dérivée des fonctions plus régulières.

La notion de sous-différentiel connaît diverses extensions aux fonctions non nécessairement convexes, par exemple aux fonctions localement lipschitziennes qu'on va traiter dans le deuxième chapitre .

Dans la suite, nous nous intéressons essentiellement à la classe des fonctions convexes. Ces hypothèses requièrent un outil spécial : le sous-différentiel d'analyse convexe, noté $\partial f(x)$*.Les fonctions convexes apparaissent abondamment dans l'ingénierie et permettent de modéliser de nombreux phénomènes non linéaires (équations de la physique, traitement du signal, théorie des jeux et de l'économie, statistiques...). Elles ont des propriétés remarquables qui permettent d'analyser plus facilement leurs propriétés et aussi de les minimiser efficacement. De nom-*

breux problèmes non convexes irrésolvables peuvent être approchés par des problèmes convexes qui eux sont presque systématiquement minimisés en des temps rapides.

Mes remerciements s'adressent à l'auteur Aude Rondepierre du cours : Introduction à l'optimisation convexe non différentiable : INSA Toulouse .

Chapitre 1

2 Éléments d'analyse convexe

Dans ce chapitre, on va présenter quelques propriétés géométriques et topologiques remarquables des fonctions convexes, Dans toutes ces notes, on se place dans l'espace vectoriel $\mathbb{R}^n$, $n \in \mathbb{N}$. On le munit d'un produit scalaire $< .,. >$. La norme associée au produit scalaire est notée $\|.\|_2$. Elle est définie pour tout $x \in \mathbb{R}^n$ par $\|x\|_2 = \sqrt{< x, x >}$.

Définitions et propriétés basiques

Une différence importante à noter est qu'on autorise ici les fonctions à valoir $+\infty$ (mais pas$-\infty$). Ainsi les fonctions considérées dans cette partie seront de la forme :

$$f : \mathbb{R}^n \to R \cup \{+\infty\}$$

En analyse convexe, on considère très souvent des fonctions prenant des valeurs dans l'ensemble des réels auxquels on rajoute l'infini, i.e :

$$\mathbb{R} \cup \{+\infty\}$$

Un avantage est de pouvoir inclure directement les contraintes dans la fonctionnelle optimisée,car en effet :

$$\underset{x \in \mathbb{R}^n}{inf} f(x) = \underset{x \in dom(f)}{inf} f(x)$$

où la définition du dom(f) sera donné ci-dessous

2.1 Domaine d'une fonction

Définition 2.1. *Le domaine d'une fonction f est noté dom(f).Il est défini par :*

$$dom(f) = \{x \in \mathbb{R}^n, f(x) \prec +\infty\}$$

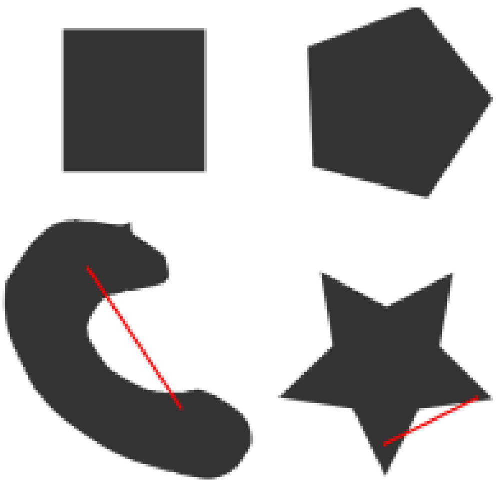
Dans toute la suite, on supposera (sans le préciser) que nos fonctions n'ont pas un domaine vide : $dom(f) \neq \emptyset$

2.2 Ensemble convexe

Définition 2.2. *Soit $X \subseteq \mathbb{R}^n$ un ensemble,Ce dernier est dit convexe si :*

$$\forall(x_1, x_2) \in X \times X, \forall\alpha \in [0, 1] \; \alpha x_1 + (1 - \alpha)x_2 \in X$$

Figure 2.1. *Les trois premiers dessins sont des exemples d'ensembles convexes en 2 dimensions, Les autres sont des exemples d'ensembles non convexes (notez qu'il existe des segments dont les extrémités appartiennent à l'ensemble, qui ne sont pas entièrement contenus dans les ensembles).*



La définition de la convexité reste identique pour les fonctions à valeur dans $\mathbb{R} \cup \{+\infty\}$.

2.3 Fonction convexe

Définition 2.3. *soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. f est dite convexe si :*

1. *dom(f) est un ensemble convexe.*
2. $\forall(x, y) \in dom(f) \times dom(f), \forall\lambda \in [0, 1], on\ a :$
   $f((1 - \lambda)x + \lambda y)) \preceq (1 - \lambda)f(x) + \lambda f(y)$

On dit que f est une fonction concave lorsque la proposition précédente a lieu avec $\succeq$ au lieu $\preceq$ et on dit que f est une fonction strictement convexe lorsque la proposition précédente a lieu avec $\prec$ au lieu de $\preceq$ (avec toutefois $x \neq y$ et $\lambda \in]0,1[$ afin d'éviter les cas automatiques d'égalité). De même, on dit que f est une fonction strictement concave lorsque la proposition précédente a lieu avec $\succ$ au lieu de $\succeq$ (avec toutefois $x \neq y$ et $\lambda \in]0,1[$ afin d'éviter les cas automatiques d'égalité).

Remarque 2.1. *On dit que $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est concave si $-f$ est convexe, et tout ce qu'on va dire s'applique aussi bien au fonctions concaves qu'aux fonctions convexes, à ce changement de signe près.*

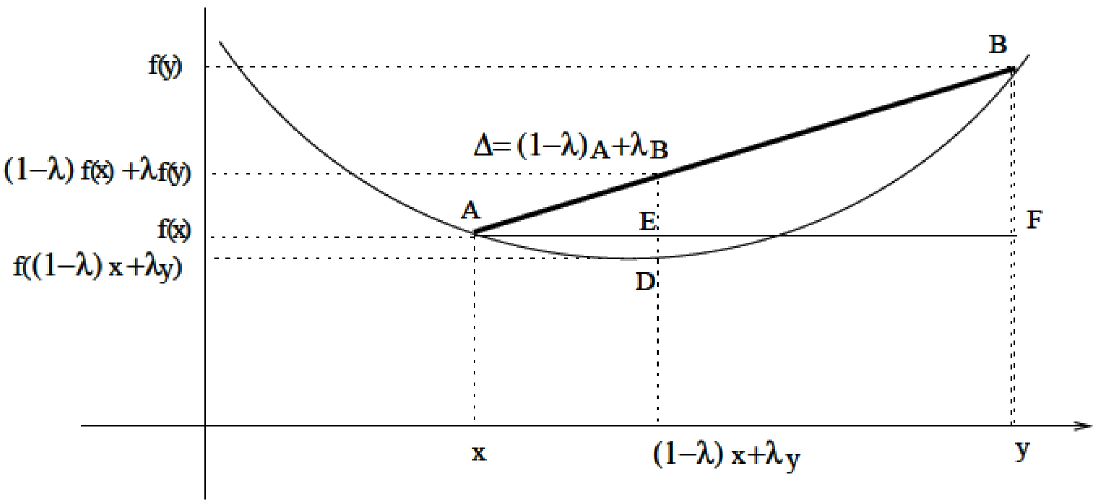


Figure 2.2. *la définition des fonctions convexes est naturelle, puisque l'interprétation graphique de la convexité d'une fonction est la suivante : pour tout couple de points A et B du graphe, le segment [AB] est situé $\ll au - dessus \gg$ du graphe. Notez qu'une fonction convexe peut ne pas être dérivable. On peut cependant montrer qu'elle est dérivable presque partout.*

Exemple 2.3.1.

1 *Toute norme $\|.\|$ sur $\mathbb{R}^n$ est convexe.*
2 *Toute forme linéaire (même discontinue) sur $\mathbb{R}^n$ est convexe.*
3 *Les sous-niveaux $L_f(\beta) := \{x \in dom(f), f(x) \preceq \beta\}$ d'une fonction convexe sont convexes .*
4 *La composition de fonctions convexes n'est pas nécessairement convexe, Par exemple :$f, g : \mathbb{R} \to \mathbb{R}$, $f(x) = exp(-x)$, $g(x) = x^2$ sont convexes alors que $fog(x) = exp(-x^2)$ n'est pas convexe .*

2.4 Combinaison convexe

Définition 2.4. *soient* $x_1, \ldots, x_m$ *m éléments de* $\mathbb{R}^n$ *.On dit que x est combinaison convexe de ces points s'il existe* $(\alpha_1, \ldots, \alpha_m)$ *tels que :*

* $\forall i \in \{1, \ldots, m\}, \alpha_i \in \mathbb{R}^+$
* $\sum_{i=1}^{m} \alpha_i = 1.$
* $x = \sum_{i=1}^{m} \alpha_i x_i.$

2.5 Enveloppe convexe

Définition 2.5. *soit* $X \subseteq \mathbb{R}^n$ *un ensemble.On appelle enveloppe convexe de X et on note* $conv(X)$ *l'ensemble convexe le plus petit contenant X. En d'autres termes c'est :*

$$conv(X) = \cap_{X_i \in \Omega} X_i$$

avec Ω *: la collection de tous les ensembles convexes contenant X .*

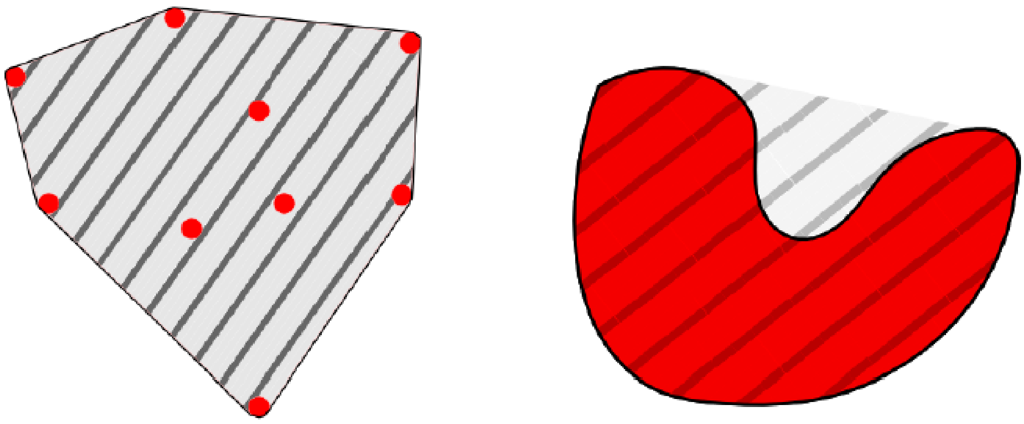


Figure 2.3. *Exemples d'enveloppes convexes. A gauche : enveloppe convexe d'un ensemble discret. A droite : enveloppe convexe d'un ensemble continu.*

La définition de la convexité permet d'obtenir un lemme souvent utile (notamment en probabilités) :

2.6 Inégalité de Jensen

Lemme 2.1. *soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, et soient $x_1, x_2, \ldots, x_m$ m points appartenant à $dom(f)$ et $\alpha_1, \ldots, \alpha_m$ des coefficients réels positifs tels que $\sum_{i=1}^{m} \alpha_i = 1$. alors :*

$$f(\sum_{i=1}^{m} \alpha_i x_i) \leqslant \sum_{i=1}^{m} \alpha_i f(x_i)$$

preuve. Par récurrence, on vérifie d'abord que pour $m = 2$, l'inégalité n'est rien d'autre que la définition de la convexité. Puis on suppose le résultat vrai au rang k et on le montre au rang k + 1 en réutilisant l'inégalité de la convexité.

Corollaire 2.1. *Soit $x \in \mathbb{R}^n$ une combinaison convexe de $x_1, \ldots, x_m$ alors*

$$f(x) \leqslant \max_{x \in \{1,\ldots,m\}} f(x_k)$$

Preuve. on a :
$x = \sum_{k=1}^{m} \alpha_k x_k$ avec $\alpha_k \in \mathbb{R}^+ \ \ \forall k \in \{1, ..., m\}$ et $\sum_{k=1}^{m} \alpha_k = 1$
or :

$$\begin{aligned} f(x) = f(\sum_{k=1}^{m} \alpha_k x_k) \leqslant \sum_{k=1}^{m} \alpha_k f(x_k) &\leqslant \max_{k \in \{1,\ldots,m\}} f(x_k). \sum_{k=1}^{m} \alpha_k. \\ &= \max_{k \in \{1,\ldots,m\}} f(x_k). \end{aligned}$$

2.7 Epigraphe d'une fonction

Définition 2.6. *soit f une fonction ,on appelle épigraphe de f et on le note epi(f) l'ensemble :*

$$epi(f) = \{(x, t) \in dom(f) \times \mathbb{R}, t \geqslant f(x)\}$$

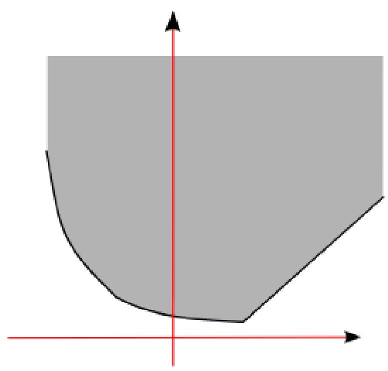


Figure 2.4. *L'épigraphe de la fonction est la zone grisée au-dessus du graphe de la fonction (en noir).*

Théorème 2.1. *f est convexe si et seulement si son épigraphe*

$$epi(f) = \{(x,t) \in dom(f) \times \mathbb{R}, t \geqslant f(x))\}$$

est convexe .

Preuve. soient (x_1, t_1) et (x_2, t_2) deux éléments de $epi(f)$
on a alors pour tout $\alpha \in [0,1]$:

$$\begin{aligned} \alpha t_1 + (1-\alpha)t_2 &\succeq \alpha f(x_1) + (1-\alpha)f(x_2) \\ &\succeq f(\alpha x_1 + (1-\alpha)x_2).(car\ f\ est\ convexe) \end{aligned}$$

Ainsi,$(\alpha x_1 + (1-\alpha)x_2, \alpha t_1 + (1-\alpha)t_2) \in epi(f)$.

Réciproquement, si $epi(f)$ est convexe, alors pour x_1, x_2 dans $dom(f)$, on a :

$$(x_1, f(x_1)) \in epi(f), (x_2, f(x_2)) \in epi(f).$$

Ainsi,$((1-\alpha)x_1 + \alpha x_2), (1-\alpha)f(x_1) + \alpha f(x_2)) \in epi(f)$, soit encore

$$f((1-\alpha)x_1 + \alpha x_2)) \preceq (1-\alpha)f(x_1) + \alpha f(x_2).$$

Chapitre 2

3 Le Sous différentiel au sens de l'analyse convexe

3.1 Introduction

L'analyse convexe est un des piliers des mathématiques appliquées. Elle intervient dans la modélisation et la résolution numérique de problèmes dans pratiquement tous les secteurs où la modélisation mathématique est pertinente : en ingénierie, en statistiques, en physique, en économie, en finance, dans les sciences de l'information, pour la simulation numérique et des données. L'objectif de cette partie est de fournir les fondements de l'analyse convexe, ses implications en générale et précisément dans l'optimisation en relation avec le calcul sous différentiel .

3.2 Gradient

Définition 3.1. *soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction , f est dit différentiable en x si et seulement si :*

$$\exists!\ x^* \in \mathbb{R}^n\ ,\ f(z) = f(x) + < x^*, z - x > + o(\|z - x\|)$$

en d'autres termes :

$$\lim_{z \to x} \frac{f(z) - f(x) - < x^*, z - x >}{\|z - x\|} = 0$$

si x^* existe , on l'appelle le gradient de f en x et on le note $\nabla f(x)$.

Théorème 3.1. ***Caractérisation différentielle de la convexité :*** *soit $f : \mathbb{R}^n \to \mathbb{R}$ un fonction différentiable . les propositions suivantes sont équivalentes :*

(a) f est convexe

(b) Ses hyperplans tangents sont des minorants :

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n,\ f(y) \succeq f(x) + < \nabla f(x), y - x >$$

(c) Le gradient de f est un opérateur monotone :

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n,\ < \nabla f(y) - \nabla f(x), y - x > \succeq 0$$

Note : un opérateur $F : \mathbb{R}^n \to \mathbb{R}^n$ est dit monotone si :

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n,\ < \nabla F(y) - \nabla F(x), y - x > \succeq 0$$

cette notion généralise la notion du gradient des fonctions convexes. elle apparait abondamment dans les E.D.P .

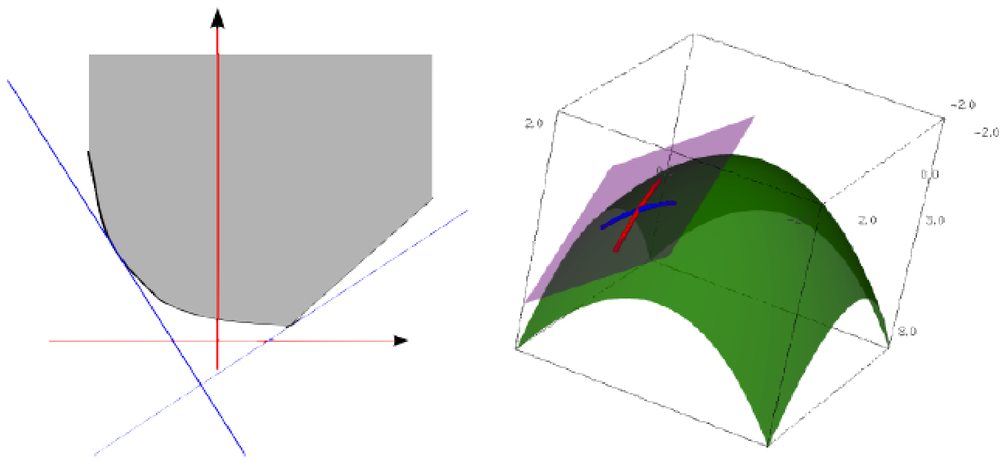


Figure 3.1. *Hyperplans tangents d'une fonction convexe (2D) et d'une fonction concave (3D). Notez que l'hyperplan tangent est un minorant pour la fonction convexe et un majorant pour la fonction concave.*

Preuve. commençons par la première équivalence $(a) \Leftrightarrow (b)$: on montre d'abord que $(a) \Rightarrow (b)$:
soit $\alpha \in [0,1]$ et $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$ on a :

$$\lim_{\alpha \to 0^+} \frac{f(x + \alpha(y-x)) - f(x)}{\alpha} = < \nabla f(x), y - x).$$

par définitions de la dérivée directionnelle .
et par convexité de f , on a de plus :

$$f(x + \alpha(y-x)) = f(\alpha y + (1-\alpha)x) \preceq \alpha f(y) + (1-\alpha) f(x)$$

donc

$$\frac{f(x + \alpha(y-x)) - f(x)}{\alpha} \preceq f(y) - f(x) \quad (\alpha \succ 0)$$

en passant a la limite quand $\alpha \to 0^+$, on obtient l'inégalité annoncée .
Montrons maintenant $(b) \Rightarrow (a)$
soient $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ et $z = \alpha x + (1-\alpha)y$.on peut écrire :

$$f(x) \succeq f(z) + < \nabla f(z), x - z >$$
$$f(y) \succeq f(z) + < \nabla f(z), y - z >$$

soit encore :

$$f(x) \succeq f(\alpha x + (1-\alpha)y) + < \nabla f(z), (1-\alpha)(x-y) >$$
$$f(y) \succeq f(\alpha x + (1-\alpha)y) + < \nabla f(z), \alpha(y-x) >$$

En multipliant la première inégalité par α, la seconde par $(1-\alpha)$ puis en additionnant les deux, on obtient l'inégalité de convexité.

3.3 Sous-gradient et sous différentiel

Définition 3.2. *Soit f une fonction convexe. Un vecteur $\eta \in \mathbb{R}^n$est appelé sous-gradient de f au point $x_0 \in dom(f)$ si :*

$$\forall x \in dom(f) \;\; f(x) - f(x_0) \succeq < \eta, x - x_0 >$$

L'ensemble de tous les sous-gradients en x_0 est appelé sous-différentiel de f. Il est noté $\partial f(x_0)$.

Interprétation géométrique

L'interprétation géométrique du sous-différentiel est la suivante. Il est formé par toutes les directions des hyperplans qui passent par le point $(x, f(x))$ et restent "sous" le graphe de la fonction f. Ces hyperplans sont appelés hyperplans support ou hyperplans d'appui au graphe de f en x.

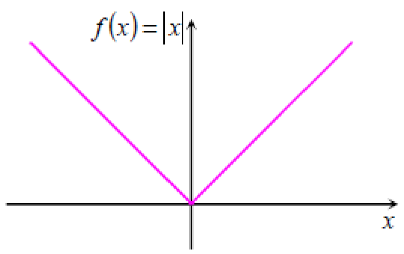
Exemple 3.3.1. *soit $f : x \in \mathbb{R} \to |x|$. Calculons les sous-gradients de f en tout point x de $\mathbb{R}$.*

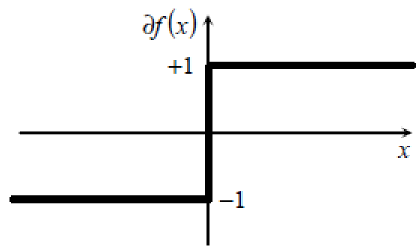
Analytiquement

on a : pour $x = 0$

$$\begin{aligned} \eta \in \partial f(0) &\Leftrightarrow f(x) - f(0) \succeq < \eta, x - 0 > \ \ \forall x \in \mathbb{R} \\ &\Leftrightarrow |x| \succeq \eta.x \\ &\Leftrightarrow \begin{cases} \ \ x \succeq \eta.x \ si \ x \succ 0 \ ou \\ -x \succeq \eta.x \ si \ x \prec 0 \end{cases} \\ &\Leftrightarrow \eta \preceq 1 \ ou \ \eta \succeq -1 \end{aligned}$$

ce qui donne en fin l'intervalle $\partial f(0) = [-1, 1]$ *.*





Géométriquement

Dans $\mathbb{R}^2$*, les hyperplans d'appui sont des droites et les sous-gradients associés leurs coefficients directeurs.*

$$\partial f(x) = \begin{cases} \{-1\} \ si \ x \prec 0 \\ \{+1\} \ si \ x \succ 0 \\ [-1, 1] \ \ si \ x = 0 \end{cases}$$

Lemme 3.1. *Soit f une fonction convexe le sous différentiel de f en* x_0 *noté :*

$$\partial f(x_0) = \{\eta \in \mathbb{R}^n, \forall x \in dom(f) \ \ f(x) - f(x_0) \succeq < \eta, x - x_0 >\}$$

est un ensemble convexe fermé.

Preuve. soient g_1 et g_2 deux éléments de $\partial f(x_0)$

on a $\forall x \in \mathbb{R}^n$:

$$f(x) - f(x_0) \succeq < g_1, x - x_0 >$$

$$f(x) - f(x_0) \succeq < g_2, x - x_0 >$$

soit $\alpha \in [0,1]$,en multipliant la première inégalité par α et la deuxième par $(1-\alpha)$ et en sommant , on voit que $\alpha g_1 + (1-\alpha) g_2 \in \partial f(x_0)$.

3.4 Opération sur les fonctions convexes

Définition 3.3. *Une fonction convexe est dite fermée ou semi-continue inférieurement (s.c.i.) si son épigraphe est une ensemble fermé.*
en d'autres terme :
on dit que f est semi continu en x_0 si :

$$\forall \epsilon \succ 0, \exists \eta \succ 0, \forall x \in \mathbb{R}^n \ tel\ que\ \|x - x_0\| \prec \eta \Rightarrow f(x) - f(x_0) \prec \epsilon$$

Soient $f_1 et f_2$ deux fonctions convexes s.c.i. et $\beta \succ 0$. Les fonctions suivantes sont convexes s.c.i. :

1 $f = \beta f_1$.
2 $f = (f_1 + f_2)\ avec\ dom(f) = dom(f_1) \cap dom(f_2.)$
3 $f = max\{f_1, f_2\}\ avec\ dom(f) = dom(f_1) \cap dom(f_2)$.

Théorème 3.2. *Soit ϕ, une fonction convexe s.c.i sur $\mathbb{R}^m$*
Soit $A : \mathbb{R}^n \to \mathbb{R}^m$ un opérateur linéaire et $b \in \mathbb{R}^m$. Alors la fonction
$x \to f(x) = \phi(Ax + b)$ est convexe s.c.i. et
$dom(f) = \{x \in \mathbb{R}^n, Ax + b \in dom(\phi)\}$.

Preuve. Soients $(x_1, x_2) \in dom(f) \times dom(f)$,
Alors pour tout $\alpha \in [0,1]$:

$$\begin{aligned} f(\alpha x_1 + (1-\alpha) x_2) &= \phi(A(\alpha x_1 + (1-\alpha) x_2) + b) \\ &= \phi(\alpha(Ax_1 + b) + (1-\alpha)(Ax_2 + b)) \quad (1 = \alpha + 1 - \alpha) \\ &\preceq \alpha \phi(Ax_1 + b) + (1-\alpha)\phi(Ax_2 + b) \quad (\phi\ est\ convexe) \\ &\preceq \alpha f(x_1) + (1-\alpha) f(x_2) \end{aligned}$$

La fermeture de l'épigraphe est due à la continuité de l'opérateur affine $x \to Ax + b$.

Théorème 3.3. *Soit f une fonction convexe s.c.i. et $x_0 \in int(dom(f))$. Alors $\partial f(x_0)$ est une ensemble non vide, convexe, borné.*

Preuve. Notons d'abord que le point $(f(x_0), x_0)$ appartient à la frontière de epi(f). D'après le théorème de séparation de Hahn-Banach, il existe donc un hyperplan d'appui à epi(f) au point $(f(x_0), x_0)$:

$$-\alpha\tau + < d, x > \preceq \ -\alpha f(x_0) + < d, x_0 > \tag{1}$$

pour tout $(\tau, x) \in \ epi(f)$. On peut choisir

$$\|d\|_2^2 + \alpha^2 = 1 \tag{2}$$

Puisque pour tout $\tau \succeq f(x_0)$ le point (τ, x_0) appartient à epi(f), on conclut que $\alpha \succeq 0$.
D'après le lemme suivant :

Lemme 3.2. *Soit f une fonction convexe et $x_0 \in int(dom(f))$. Alors f est majorée localement en x_0.*

il existe donc $\epsilon \succ 0 \ et \ M \succ 0$ tels que $B_2(x_0, \epsilon) \ \subset dom(f)$ et

$$f(x) - f(x_0) \preceq \ M\|x - x_0\|$$

pour tout $x \in B_2(x_0, \epsilon)$. Ainsi, d'après (1), pour tout x dans cette boule on a

$$\begin{aligned} < d, x - x_0 > \ &\preceq \alpha(f(x) - f(x_0)) \\ &\preceq \alpha M\|x - x_0\|_2 \end{aligned}$$

En choisissant $x = x_0 + \epsilon d$ on obtient $\|d\|_2^2 \preceq M\alpha\|d\|_2$. En utilisant la condition de normalisation (2), on obtient

$$\alpha \succeq \frac{1}{\sqrt{1 + M^2}}$$

Ainsi, en choisissant $g = d/\alpha$ on obtient pour tout $x \in dom(f)$

$$f(x) \succeq f(x_0) + < g, x - x_0 >$$

Finalement, si $g \in \partial f(x_0)$ et $g \neq 0$, alors en choisissant $x = x_0 + g/\|g\|_2$ on obtient

$$\epsilon\|g\|_2 =< g, x - x_0 > \preceq f(x) - f(x_0) \preceq M\|x - x_0\|_2 = M\epsilon$$

Ainsi $\partial f(x_0)$ est borné .

Remarque 3.1. *Le sous-différentiel peut ne pas exister sur le bord du domaine. Par exemple la fonction $f(x) = -\sqrt{x}$ sur $\mathbb{R}^+$ est convexe et fermée, mais le sous-différentiel n'existe pas en 0 car $\underset{x \to 0^+}{lim} f'(x) = -\infty$.*

Théorème 3.4. *soit f une fonction convexe et $x_0 \in int(dom(f))$. alors f est continue et localement Lipschitz en x_0.*

Preuve. D'après le lemme (3.2) ,

$$\exists M \prec +\infty \; et \; \epsilon \succ 0 \; tels \; que \; f(x) \preceq M, \quad \forall x \in B_2(x_0, \epsilon)$$

soit $y \in B_2(x_0, \epsilon)$ *et* $z \in \partial B_2(x_0, \epsilon)$ un point de la frontière tel que :

$$z = x_0 + \frac{1}{\alpha}(y - x_0) \; avec \; \alpha = \frac{1}{\epsilon}\|y - x_0\|_2$$

par construction $\alpha \preceq 1$ et $y = \alpha z + (1 - \alpha)x_0$
par convexité de f on a donc :

$$\begin{aligned} f(y) &\preceq (1 - \alpha)f(x_0) + \alpha f(z) \\ &\preceq f(x_0) + \alpha(M - f(x_0)) \\ &= f(x_0) + \frac{M - f(x_0)}{\epsilon}\|y_\alpha - x_0\|_2 \end{aligned}$$

on a donc $\forall y \in B_2(x_0, \epsilon)$:

$$|f(y) - f(x_0)| \preceq \frac{M - f(x_0)}{\epsilon}\|y_\alpha - x_0\|_2$$

3.5 Application à l'optimisation

Le sous-différentiel est central car il permet notamment de caractériser les minimiseurs d'une fonction. On considère le problème suivant :

$$Trouvrer \; x^* \; tel \; que \; f(x^*) = \underset{x \in \Re^n}{min} \; f(x)$$

ou f est une fonction convexe s.c.i .

Théorème 3.5. *x^* est solution du problème ci-dessus si et seulement si $0 \in \partial f(x^*)$.*

Preuve. si $0 \in \partial f(x^*)$, Alors :

$$\begin{aligned} f(x) &\succeq f(x^*) + < 0, x - x^* > \quad \forall x \in dom(f) \\ &\succeq f(x^*) \quad \forall x \in dom(f). \end{aligned}$$

Réciproquent, si :

$$f(x) \succeq f(x^*) \;\; \forall x \in dom(f)$$

alors : $0 \in \partial f(x^*)$ par définition du sous différentiel .

3.6 Règle de calcul du sous différentiel

Lemme 3.3. *soit f une fonction convexe s.c.i différentiable sur son domaine, alors :*

$$\forall x \in int(dom(f)) \;\; \partial f(x) = \{\nabla f(x)\}$$

Lemme 3.4. *soit f une fonction convexe s.c.i , avec dom(f) $\subseteq \mathbb{R}^n$, soit $A : \mathbb{R}^n \to \mathbb{R}^m$ un opérateur linéaire et $b \in \mathbb{R}^m$,la fonction $x \to \phi(x) = f(Ax+b)$ est convexe s.c.i et $\forall x \in int(dom(f))$ on a :*

$$\partial \phi(x) = A^T \partial f(Ax+b)$$

Lemme 3.5. *soit $f(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$ ou f_1 et f_2 sont deux fonctions convexes s.c.i sur $\mathbb{R}^n$ et $(\alpha_1, \alpha_2) \in \mathbb{R}^+ \times \mathbb{R}^+$ alors :*

$$\partial f(x) = \{\eta \in \mathbb{R}^n, \exists (x_1, x_2) \in \partial f_1(x) \times \partial f_2(x) \; et \; \eta = x_1 + x_2\}$$

Lemme 3.6. *soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe fermé , et $g : \mathbb{R} \cup \{+\infty\} \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe croissante , notons $h = gof$, alors $\forall x \in int(dom(f))$on a : ,*

$$\partial h(x) = \{\eta_1 \eta_2, \; \eta_1 \in \partial g(f(x)), \; \eta_2 \in \partial f(x)\}$$

Lemme 3.7. *soit $(f_i)_{i=1,\ldots,m}$ un ensemble de fonction convexes s.c.i , alors la fonction :*

$$f(x) = \max_{i=1,\ldots,m} f_i(x)$$

et aussi convexe s.c.i de $dom(f) = \cap_{i=1}^m dom(f_i)$ et $\forall x \in int(dom(f))$ on a :

$$\partial f(x) = conv(\partial f_i(x), \; i \in I(x))$$

où $I(x) = \{i \in \{1, ..., m\} \;\; f_i(x) = f(x)\}$

3.7 Transformé de Fenchel ou conjugué convexe

La fonction conjuguée, aussi appelée transformée de Fenchel ou transformée de Legendre- Fenchel est utilisée pour :

- convexifier une fonction (en calculant la bi-conjuguée, i.e. la conjuguée de la conjuguée).
- calculer le sous-différentiel d'une fonction convexe.
- calculer des problèmes dits "duaux", en optimisation. Ces problèmes apportent souvent beaucoup d'information sur les problème "primaux", i.e. ceux que l'on souhaite résoudre.
- passer de la mécanique lagrangienne à la mécanique hamiltonienne,...

Elle a été introduite par Mandelbrojt en 1939 puis précisée et améliorée par Fenchel en 1949. Cette transformée généralise la transformation de Legendre (1787).

Définition 3.4. ***Conjugué convexe :***
soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction , sa conjuguée est définie par

$$f^*(s) = \sup_{x \in \mathbb{R}^n} < s, x > -f(x)$$

L'application $f \to f^*$ est appelée transformation de Lengendre-Fenchel. La fonction f^* est appelée conjuguée convexe, transformée de Fenchel ou transformée de Legendre- Fenchel de f.

La motivation pour introduire cette transformation est la suivante. On peut définir la convexifiée fermée d'une fonction comme l'enveloppe supérieure de toutes les minorantes affines de f. Parmi toutes les minorantes affines $x \to < s, x > +\alpha$, on ne peut garder que celle qui est la plus haute, c'est-à-dire qui a le plus grand α. Il faut donc déterminer le plus grand α tel que $x \in \Re^n$:

$$< s, x > +\alpha \preceq f(x).$$

La plus petite valeur de $-\alpha$ est donnée par :

$$f^*(s) = \sup_{x \in \mathbb{R}^n} < s, x > -f(x)$$

Proposition 3.1. *Pour toute fonction f, f^* est convexe, fermée.*

Proposition 3.2. *Pour tout $(x, s) \in \mathbb{R}^n \times \mathbb{R}^n \quad f(x) + f^*(x) \succeq < x, s >$*

preuve. évidente d'après la définition.

Chapitre 3

4 Le sous-différentiel au sens de Clarke

4.1 Introduction

Le sous-différentiel de Clarke, dont il est principalement question ci-dessous, est une notion décrivant le comportement local d'une fonction en un point. Si la fonction est dérivable en ce point (il faut un peu plus que cela en réalité), ce sous-différentiel se confond avec la dérivée. Sinon c'est un ensemble d'approximations linéaires censées décrire toutes les possibilités de variation infinitésimale de la fonction. Ce sous différentiel est donc sujet à des variations brusques qui apparaissent lorsqu'on quitte un point de non-différentiabilité.

4.2 Définitions et propriétés basiques

Nous dénotons par $B(x, r)$ la boule ouverte de centre $x \in \mathbb{R}^n$ et de rayon $r \succ 0$, définie par

$$B(x, r) \equiv \{y \in \mathbb{R}^n \;\; tel\ que \;\; \|y - x\| \prec r\}$$

Commençons par la définition des fonctions Lipschitziennes.

On note par X un espace de Banach , on dit qu'une fonction $f : X \to \mathbb{R}$ est localement lipschitzienne de rang k, au voisinage de x si :

$$\forall \epsilon \succ 0, \forall (y, z) \in B(x, \epsilon) \;\; |f(y) - f(z)| \preceq k\|y - z\|$$

Une fonction localement Lipschitzienne en x présente, ainsi, un taux d'accroissement qui est borné dans un voisinage de x. D'autre part, une fonction localement Lipschitzienne en x n'est pas forcement différentiable en x.

4.3 La dérivée directionnelle généralisée

contrairement au cas convexe, l'hypothèse que la fonction $f : \mathbb{R}^n \to \mathbb{R}$ est localement Lipschitzienne en $x \in \mathbb{R}^n$ n'est pas suffisante pour l'existence des dérivées directionnelles de f. C'est pourquoi on doit généraliser le concept de dérivée directionnelle.

Définition 4.1. *la dérivée directionnelle généralisée d'une fonction f au point x à la direction v noté* $f^{\circ}(x; v)$ *est définie par :*

$$f^{\circ}(x; v) = \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{f(y + tv) - f(y)}{t}$$

On peut, ainsi, définir un sous-différentiel pour les fonctions localement Lipschitziennes d'une façon analogue au cas convexe :

Définition 4.2. *Pour une fonction* $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *localement Lipschitzienne en* $x \in \mathbb{R}^n$*, le sous-différentiel de Clarke de f en x, noté* $\partial_c f(x)$*, est l'ensemble non vide, compact et convexe de* $\mathbb{R}^n$ *dont la fonction support est :*
$d \in \mathbb{R}^n \rightarrow f^{\circ}(x; d)$*, c'est-à-dire :*

$$\partial_c f(x) \equiv \{s \in \mathbb{R}^n,\ f^{\circ}(x; d) \succeq < s, d > \ pour\ tout\ d \in \mathbb{R}^n\}$$

Les dérivées directionnelles généralisées peuvent être déterminées aussi à partir de $\partial_c f(x)$ *pour toute direction* $d \in \mathbb{R}^n$ *:*

$$f^{\circ}(x; d) = max\{< s, d > : s \in \partial_c f(x)\}$$

Les éléments de $\partial_c f(x)$ sont appelés sous-gradients de Clarke (ou gradients généralisés) de f en x.

4.4 Sous-différentiel de Clarke :Propriétés

Le sous-différentiel de Clarke généralise les notions de sous-différentiel d'une fonction convexe et de gradient d'une fonction différentiable :

Proposition 4.1.

1. *Si une fonction f est convexe, on a* $f^{\circ}(x; d) = f'(x, d)$ *pour tout* $d \in \mathbb{R}^n$*, et ainsi* $\partial f(x) = \partial_c f(x)$
2. *Si une fonction f localement Lipschitzienne en* $x \in \mathbb{R}^n$ *est différentiable en x, alors on a* $f'(x, d) =< \nabla f(x), d > \preceq f^{\circ}(x; d)$ *pour tout* $d \in \mathbb{R}^n$*, et donc* $\nabla f(x) \in \partial_c f(x)$
3. *Si une fonction f est continûment différentiable en x , on a l'égalité* $< \nabla f(x), d >= f^{\circ}(x; d)$ *pour tout* $d \in \mathbb{R}^n$*,de sorte que :*
   $\partial_c f(x) = \{\nabla f(x)\}$

toutes ces règles précédentes sont résumées dans un théorème très important :

Théorème 4.1. *si f est une fonction continument différentiable au voisinage de x on a :* $\partial_c f(x) = \{f'(x)\}$ *et si f est une fonction convexe s.c.i avec* $x \in int(dom(f))$ *alors :*

$$\partial_c f(x) = \partial f(x)$$

Proposition 4.2. *soient x un élément de X,et k un élément de* $\mathbb{R}$ *on :*

a- la fonction $v \to f^\circ(x;v)$ *est fini , positive homogène et sous additive sur X, elle satisfait :*

$$|f^\circ(x;v)| \preceq k\|v\| \; v \in X$$

b- pour tout $v \in X$*, la fonction* $(u;w) \to f^\circ(u;w)$ *est semi continu supérieurement à (x ;v),et la fonction* $w \to f^\circ(x;w)$ *est k-lipschitzienne sur X*

c- on a $f^\circ(x;-v) = (-f)^\circ(x;v); \; pour \; tout \; v \in X$

Preuve. compte tenu de l'état Lipschitz , la valeur absolu de la différence dans la définition de quotient $f^\circ(x;v)$ est borné par $k\|v\|$ pour tout y suffisamment proche de x et t suffisamment près de 0 , il s'ensuit que $|f^\circ(x;v)|$ admet la même limite supérieure donc le fait que :

$$f^\circ(x;\lambda v) = \lambda f^\circ(x;v) \; pour \; tout \; \lambda \succeq 0$$

est immédiate , passons maintenant a la sous additivité on a :

$$\begin{aligned} f^\circ(x;w+v) &= \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{f(y+tv+tw) - f(y)}{t} \\ &\preceq \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{f(y+tv+tw) - f(y+tw)}{t} + \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{f(y+tw) - f(y)}{t} \end{aligned}$$

la première limite supérieure dans cette expression est $f^\circ(x;v)$, étant donné que le terme $y + tw$ représente en fait seulement une variable muette convergeant vers x , on conclut que :

$$f^\circ(x;v+w) \preceq f^\circ(x;v) + f^\circ(x;w)$$

ce qui établit (a) ,
maintenant , soient $(x_i)_{i\in\mathbb{N}}$ et $(v_i)_{i\in\mathbb{N}}$ deux suites convergents respectivement vers x et v , par définition on a :

$$\forall i \in \mathbb{N}, \exists y_i \in X \; tel \; que \; t_i \succ 0 \; \|y_i - x_i\| + t_i \prec \frac{1}{i}$$

et

$$\begin{aligned} f^\circ(x_i; v_i) - \frac{1}{i} &\preceq \frac{f(y_i + t_i v_i) - f(y_i)}{t_i} \\ &= \frac{f(y_i + t_i v) - f(y_i)}{t_i} + \frac{f(y_i + t_i v_i) - f(y_i + t_i v)}{t_i} \end{aligned}$$

tenant compte de l'état lipschitzienne , on voit bien que le dernier terme est majoré par $k\|v_i - v\|$,en prenant la limite supérieure avec $i \to \infty$, on aura :

$$\underset{i \longrightarrow +\infty}{limsup} \; f^\circ(x_i; v_i) \preceq f^\circ(x; v)$$

ce qui établit les semi-continuité dans (b).
soient $x, v \in X$, tenant compte de l'état lipschitzienne on a :

$$f(y + tv) - f(y) \preceq f(y + tw) - f(y) + k\|v - w\|t$$

pour tout y près de x et pour tout $t \succ 0$ et en divisant par t ,on prend les limits supérieurs pour $y \to\to x$ et $t \to 0$ on trouve que :

$$f^\circ(x; v) \preceq f^\circ(x; w) + k\|v - w\|$$

puisque c'est également si on commute v et w on affirme (b) . pour démontrez (c) on calcule :

$$\begin{aligned} f^\circ(x; -v) &= \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \; \frac{f(y - tv) - f(y)}{t} \\ &= \underset{\substack{u \longrightarrow x \\ t \downarrow 0}}{limsup} \; \frac{(-f)(u + tv) - (-f)(u)}{t} \; pour \; u = y - tv \\ &= (-f)^\circ(x; v) \end{aligned}$$

ce qui termine la preuve.

Proposition 4.3. *soit f une fonction k-lipschitzienne au voisinage de x on a :*

$$\partial_c f(x) \subset B(0, k)$$

Preuve. d'apres la proposition (4.2),et la définition du gradient généralisé on a :

$$< \eta, v > \preceq f^{\circ}(\eta; v) \preceq k\|v\| \; pour \; tout \; \eta \in \mathbb{R}^n$$

donc , on aura :

$$< \eta, v > \preceq k\|v\| \quad \forall v \in X$$

ainsi $\partial_c f(x) \subset B(0, k)$

Exemple 4.4.1. *soit $f(x) = \|x\|$ une fonction lipschitzienne de rang k=1 , d'après la proposition(4.2) on a :*

$$f^{\circ}(0; v) \preceq \|v\| \; \forall v \in X$$

or on a : $f'(0; v) = \|v\|$ d'après la dérivée directionnelle usuelle.
et on an aussi :

$$f^{\circ}(0; v) \succeq f'(0; v) \quad \forall v \in X$$

ce qui donne :

$$f'(0; v) \preceq f^{\circ}(0; v) \preceq \|v\|$$

maintenant d'après la définition du sous différentiel généralisé de Clarke on a : $\eta \in \partial_c f(0)$ si et seulement si :

$$f^{\circ}(0; v) = \|v\| \succeq < \eta, v > \quad \forall v \in X$$

on conclut donc que :

$$\partial_c f(0) = B(0, 1)$$

Remarque 4.1. *Rarement qu'on calcul les dérivés à partir des différences de quotients , donc le calcul des gradients généralisés va être effectué à l'aide des règles de calcul .*

4.5 Règles de calcul du sous-différentiel de Clarke

En général, les règles de calcul du sous-différentiel de Clarke ne comprennent que des inclusions. On peut, toutefois, obtenir des égalités sous une condition suffisante plus forte que la continuité Lipschitzienne locale :

Proposition 4.4. *pour tout $\lambda \in \mathbb{R}$ on a :*

$$\partial_c(\lambda f)(x) = \lambda \partial_c f(x)$$

Preuve. notons que (λf) est lipschitzienne au voisinage de x :
pour $\lambda \succ 0$ on a :

$$(\lambda f)^{\circ}(x; .) = \lambda f^{\circ}(x; .) \ \ (d'apres\ la\ definition)$$

pour le cas négatif , il suffit de prendre $\lambda = -1$ on aura :

$$\begin{aligned}\eta \in \partial_c(-f)(x) &\Leftrightarrow (-f)^{\circ}(x; v) \succeq < \eta, v > \ \ \forall v \in X \\ &\Leftrightarrow f^{\circ}(x; -v) \succeq < \eta, v > \ \ \forall v \in X \\ &\Leftrightarrow -\eta \in \partial_c f(x)\end{aligned}$$

afin de continuer les règles de calcul du gradient généralisé , on va introduire chaque fois des propriétés fonctionnelles dont on aura besoin :

Définition 4.3. *Une fonction* $f : \mathbb{R}^n \to \mathbb{R}$ *localement Lipschitzienne est dite régulière si la dérivée directionnelle* $f'(x, d)$ *existe pour tout* $x, d \in \mathbb{R}^n$ *et si,de plus,on a* $f'(x, d) = f^{\circ}(x, d)$ *.*

Théorème 4.2. *soient f et g deux fonctions lipschitziennes au voisinage de x on a :*

$$\partial_c(f + g)(x) \subset \partial_c f(x) + \partial_c g(x)$$

on aura l'égalité dans le cas ou f et g sont deux fonctions régulières en x .

Preuve. on :

$$(f + g)^{\circ}(x; v) \preceq f^{\circ}(x; v) + g^{\circ}(x; v) \ \forall v \in X$$

d'après la définition de la dérivé directionnel généralisé on :

$$\begin{aligned}\underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{(f + g)(y + tv) - (f + g)(y)}{t} &\preceq \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{f(y + tv) - f(y)}{t} + \underset{\substack{y \longrightarrow x \\ t \downarrow 0}}{limsup} \frac{g(y + tv) - g(y)}{t} \\ &\preceq f^{\circ}(x; v) + g^{\circ}(x; v)\end{aligned}$$

lorsque f et g sont deux fonctions régulières en x , l'inégalité devient une égalité car on a :

$$f^{\circ}(x; v) + g^{\circ}(x; v) = f'(x; v) + g'(x; v) = (f + g)'(x; v) \preceq (f + g)^{\circ}(x; v)$$

ainsi on aura :

$$(f + g)^{\circ}(x; v) = f^{\circ}(x; v) + g^{\circ}(x; v)$$

Exemple 4.5.1. *un simple exemple pour le cas de l'inclusion stricte soient* $f(x) = \|x\|$ *et* $g(x) = -\|x\|$ *on a :*

$$\{0\} = \partial_c(f+g)(0) \underset{\neq}{\subset} \partial_c f(0) + \partial_c g(0) = 2B(0,1)$$

Proposition 4.5. *soient* $(f_i)_{i\in\{1,2,\ldots,n\}}$ *des fonctions lipschitziennes au voisinage de x et* λ_i *des scalaires pour tout* $i \in \{1,2,...,n\}$ *on pose* $f := \sum_{i=1}^{n} f_i\lambda_i$ *qui est lipschitzienne au voisinage de x , alors on a :*

$$\partial_c(\sum_{i=1}^{n} f_i\lambda_i)(x) \subset \sum_{i=1}^{n} \lambda_i\partial_c(f_i)(x)$$

Proposition 4.6. *soit f une fonction lipschitzienne au voisinage de x et g une fonction continument différentiable au voisinage de x on a :*

$$\partial_c(f+g)(x) = \partial_c f(x) + \{g'(x)\}$$

5 Conclusion :

Dans ce travail on a défini les notions du sous différentiel au sens de l'analyse convexe et le sous différentiel au sens de Clarke, et leurs applications à l'optimisation, ainsi que les différentes propriétés et règles de calcul.
il y a d'autres classes de fonction auxquelles on peut définir le sous différentiel avec un nouveau titre de propriété et règles de calcul

BIBLIOGRAPHIE

[1] Rockafellar. R. T., Convex Analysis. Princeton Mathematics Ser. 28. Princeton University Press, Princeton, New Jersey, 1970.

[2] Clarke. F.H., Optimization and Nonsmooth Analysis.Generalized gradients and applications. John Wiley AND Sons, New York,1983.

[3] THÈSE En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE Délivré par l'Institut Supérieur de l'aéronautique et de l'espace Spécialité : Systèmes automatiques Présentée et soutenue par Alberto Mota SIMÕES sous sujet :Synthèse de compensateurs structurés par l'optimisation non lisse.

[4] cous : Introduction à l'optimisation convexe non différentiable : INSA Toulouse