MASTER EN SCIENCES MATHÉMATIQUES

Epi-convergence et emboîtement épigraphique en optimisation

VERSTRAETEN, Sophie

Award date:
1996

Awarding institution:
Universite de Namur

Facultés Universitaires Notre-Dame de la Paix
Namur
Faculté des Sciences - Département de Mathématique
Année académique 1995-1996

EPI-CONVERGENCE ET EMBOITEMENT EPIGRAPHIQUE EN OPTIMISATION

Promoteur: J.J. Strodiot

Sophie Verstraeten

Je remercie spécialement Monsieur le Professeur Jean-Jacques Strodiot pour l'aide et les précieux conseils qu'il m'a apportés tout au long de l'élaboration de ce mémoire.

Je tiens également à remercier mes parents, Yves et ma soeur pour leur soutien durant ces quatre années d'étude.

Résumé

Tout d'abord, on introduit le concept d'épi-convergence. Après en avoir étudié le point de vue analytique et géométrique, on examine son implication en optimisation notamment dans la convergence des méthodes barrières et de pénalisations extérieures.

Ensuite, la notion d'emboîtement épigraphique d'une suite de fonctions est considérée comme relâchement de l'épi-convergence. On montre que, pour des algorithmes qui recherchent seulement des points stationnaires, la convergence peut être assurée en utilisant des approximations de la fonction objectif dont les dérivées directionnelles satisfont la propriété d'emboîtement épigraphique. On démontre que l'emboîtement épigraphique fournit un outil d'analyse de convergence pour plusieurs algorithmes de résolution d'inégalités variationnelles et de problèmes d'optimisation différentiable et non différentiable.

Abstract

First we introduce the concept of epi-convergence. We study the analytic and geometric point of view as well as its implication in optimization namely in the convergence of barrier methods and exterior penalization methods.

Then we see the epigraphical nesting of a sequence of fonctions as a relaxation of epi-convergence. We show that, for algorithms which only seek a stationary point, convergence can be assured by objective fonctions approximations whose directional derivatives satisfy an epigraphical nesting property. We proof that epigraphical nesting provides a unifying thead that ties together a number of different algorithms, including those for the solution of variational inequalities and smooth as well as non-smooth optimization.

Table des matières

Introduction

En programmation mathématique, on considère des problèmes du type

$$(\mathrm{P}) \qquad \min\{f(x) \ t.q. \ x \in \Re^n\}$$

où $f : \Re^n \to \Re \bigcup \{+\infty\}$.
La plupart des méthodes de résolution de ce problème consistent à remplacer la fonction objectif f par une suite d'approximations $\{f_k\}$ qui engendrent les problèmes:

$$(P_k) \qquad \min\{f_k(x) \ t.q. \ x \in \Re^n\}.$$

La question est de savoir sous quelles hypothèses les solutions des problèmes (P_k) permettent de résoudre le problème (P). Le but de cet exposé est de répondre à cette question via la notion d'épigraphe de fonctions.

Dans un premier temps, comme P.Kall [7], on retrace la découverte de la notion d'épi-convergence via son implication en optimisation. Le lien entre ce nouveau concept et les théories classiques de convergence de fonctions telles la convergence continue, uniforme et simple est établie.

Dans le second chapitre, après avoir introduit la notion de limites d'une suite d'ensembles (Attouch-Théra [1]), on aborde l'épi-convergence du point de vue géométrique en identifiant les fonctions et leurs épigraphes (Attouch-Wets [2]). Cela permet d'établir quelques nouveaux théorèmes de convergence.

On termine cette première partie en montrant la convergence de deux types de méthodes: les méthodes barrières et de pénalisations extérieures.

Dans la seconde partie, comme le suggère Higle et Sen [5], on remarque que l'épi-convergence est une condition suffisante mais non nécessaire à l'optimalité des points d'accumulation de la suite des solutions des problèmes (P_k). On introduit alors la notion d'emboîtement épigraphique des fonctions approximantes f_k. Or ce concept ne suffit pas à montrer la convergence de nombreux algorithmes d'optimisation notamment ceux qui recherchent des points stationnaires. On montre alors que la convergence de ces algorithmes est assurée si les dérivées directionnelles de f et des fonctions approximantes f_k satisfont la propriété d'emboîtement épigraphique (Higle-Sen [6]).

Ainsi, le Chapitre 5 traite de la minimisation sans containte d'une fonction non différentiable. On montre notamment que l'emboîtement épigraphique permet de prouver la convergence des méthodes de type Newton.

Dans le chapitre suivant, on démontre la convergence des méthodes de programmation quadratique successive pour des problèmes avec contraintes.

Enfin, le Chapitre 7 concerne la minimisation d'une fonction convexe non différentiable. Après avoir défini la méthode proximale ainsi que quelques unes de ses propriétés, on montre sa convergence via l'emboîtement épigraphique. On termine par une condition suffisante à l'emboîtement épigraphique: la notion de ∂−compatibilité.

Partie I

Épi-convergence

Motivation

On considère le problème d'optimisation suivant:

$$(P) \qquad \min\{f(x)\ t.q.\ x \in \Re^n\}$$

où f est une fonction de $\Re^n$ dans $\Re \bigcup \{+\infty\}$.

Dans la plupart des méthodes pour résoudre (P), on remplace la fonction objectif f par une suite de fonctions $\{f_k\}$ qui l'approxime. On est ainsi amené à rechercher, pour chaque entier k, une solution du problème:

$$(P_k) \qquad \min\{f_k(x)\ t.q.\ x \in \Re^n\}$$

supposé plus facile à résoudre que le problème initial (P).

Dans le but de s'assurer que les valeurs optimales f_k^* et les solutions x_k^* des problèmes (P_k) approximent bien la valeur optimale f^* et la solution x^* de (P), il nous faut imposer des conditions sur le lien entre la suite $\{f_k\}$ et la fonction f. De plus, étant donné que l'on est pas certain que les solutions x_k^* sont uniques, on ne peut s'attendre à ce que la suite $\{x_k^*\}$ converge. Tout ce qu'on peut espérer est que tout point d'accumulation de la suite $\{x_k^*\}$ soit solution de (P). L'hypothèse classique pour obtenir cette propriété est la convergence uniforme de la suite $\{f_k\}$ vers f sur tout sous-ensemble compact de $\Re^n$ ainsi qu' une certaine continuité de f.

Mais depuis les années soixante, on a introduit d'autres conditions comme celle de l'épi-convergence de la suite $\{f_k\}$. C'est ce concept qu'on va étudier dans cette première partie.

Chapitre 1

Épi-convergence et convergence de fonctions

Dans ce chapitre, on montrera comment la notion d'épi-convergence est liée à la théorie classique de la convergence de fonctions. En vue de cela, commençons par rappeler quelques concepts classiques de convergence.

1.1 Concepts classiques de convergence

On considère les fonctions $f : \Re^n \to \Re$ et $f_k : \Re^n \to \Re$, $k \in \mathrm{N}$. Dans ce chapitre et jusqu'à preuve du contraire, on se restreindra à des fonctions f_k et f à valeurs finies, ceci dans le but d'éviter des hypothèses techniques supplémentaires. On dira que

- la suite $\{f_k\}$ *converge uniformément* sur $\Re^n$ vers la fonction f si
  $\{f_k\}$ converge uniformément vers f sur tout sous-ensemble compact D de $\Re^n$, i.e., si pour tout ensemble compact D, on a la propriété:
  $\forall \epsilon > 0 \ \exists N(\epsilon) \ t.q. \ \forall x \in D, \forall k \geq N(\epsilon) \ |f_k(x) - f(x)| < \epsilon.$

- la suite $\{f_k\}$ *converge point par point* sur $\Re^n$ vers la fonction f si
  $\forall x \in \Re^n, \ \{f_k(x)\} \to f(x)$ i.e., si
  $\forall x \in \Re^n, \ \forall \epsilon > 0 \ \exists M \ t.q. \ \forall k \geq M \ |f_k(x) - f(x)| < \epsilon.$

- la suite $\{f_k\}$ *converge de façon continue* sur $\Re^n$ vers la fonction f si
  $\forall x \in \Re^n, \forall \{x_k\} \to x : \{f_k(x_k)\} \to f(x).$

Le théorème suivant exprime le lien entre ces trois types de convergence.

Théorème 1.1

Les trois assertions suivantes sont équivalentes:

- (U)*: f est continue et $\{f_k\} \to f$ uniformément sur $\Re^n$.*

- (C)*: $\{f_k\} \to f$ de façon continue sur $\Re^n$*

- (EQ)*: $\{f_k\} \to f$ point par point sur $\Re^n$ et les fonctions f_k sont presque équi-continues, i.e.,*
  $\forall x \in \Re^n, \forall \epsilon > 0, \exists$ une boule ouverte $B(x,\epsilon)$ de centre x et de rayon ϵ et un nombre $N(x,\epsilon)$ tels que
  $\forall y \in B(x,\epsilon), \forall k \geq N(x,\epsilon)$ $|f_k(y) - f_k(x)| < \epsilon$.

Remarque: l'assertion (EQ) implique que les fonctions f_k soient continues. Mais en général, la presque équi-continuité des fonctions f_k n'implique pas à elle seule leur continuité. En effet, il suffit de considérer la suite $\{f_k\}$ où $f_k : \Re \to \Re$ est définie par

$$f_k(x) = \begin{cases} \frac{x}{k} & \text{si } x \text{ est un rationnel} \\ 0 & \text{sinon.} \end{cases}$$

Ces fonctions ne sont évidemment pas continues, mais elles sont presque équi-continues car il suffit de prendre dans la définition $B(x,\epsilon) = \{y \ t.q. \ |y - x| < \frac{\epsilon}{2}\}$ et $N(x,\epsilon) > 2|x|/\epsilon$.

Voyons à présent la preuve du théorème.

Preuve

(i) **(U) $\Rightarrow$ (C)**
Soit $x \in \Re^n$ et $\{x_k\}$ une suite convergeant vers x. Montrons que la suite $\{f_k(x_k)\}$ converge vers $f(x)$.
Pour cela, considérons $\epsilon > 0$ arbitraire et un ensemble D compact contenant $\{x_k\}$ et x. Comme f est continue, la suite $\{f(x_k)\}$ converge vers $f(x)$ et donc

$$\exists k^1(\epsilon) \ t.q. \ \forall k \geq k^1(\epsilon), \ |f(x) - f(x_k)| < \frac{\epsilon}{2}.$$

D'autre part, comme $\{f_k\}$ converge uniformément vers f sur D , on a

$$\exists k^2(\epsilon) \ t.q. \ \forall k \geq k^2(\epsilon), \ \forall x \in D \ |f_k(x) - f(x)| < \frac{\epsilon}{2}.$$

Dès lors, pour $k > \max(k^1(\epsilon), k^2(\epsilon))$, on obtient

$$|f(x) - f_k(x_k)| \leq |f(x) - f(x_k)| + |f(x_k) - f_k(x_k)| < \epsilon.$$

(ii) **(C) $\Rightarrow$ (EQ)**

Si on choisit $x_k = x$ pour tout k dans la définition de (C), on obtient $\{f_k(x)\} \to f(x)$, i.e., la convergence point par point de la suite $\{f_k\}$.

Montrons à présent que les fonctions f_k sont presque équi-continues.
Supposons, par l'absurde, que (EQ) n'est pas vrai. Alors, il existe $x \in \Re^n$ et $\epsilon > 0$ tels que pour tout $k \in N_0$ et pour toute boule ouverte $B(x, 1/k)$ de centre x et de rayon $1/k$, on a

$$\exists y_k \in B(x, 1/k),\ \exists q_k > k \ t.q.\ |f_{q_k}(y_k) - f_{q_k}(x)| \geq \epsilon. \tag{1.1}$$

Sans perdre de généralité, on peut supposer que la suite $\{q_k\}$ est strictement croissante (si ce n'est pas le cas, on en considère une sous-suite). La suite $\{y_k\}$ convergeant vers x et la suite $\{f_k\}$ convergeant point par point vers f, on déduit de l'hypothèse que $\lim_{k\to\infty} f_{q_k}(y_k) = f(x)$ et $\lim_{k\to\infty} f_{q_k}(x) = f(x)$. Dès lors, pour k assez grand, on a

$$|f_{q_k}(y_k) - f(x)| \leq \frac{\epsilon}{2} \text{ et } |f(x) - f_{q_k}(x)| \leq \frac{\epsilon}{2}.$$

Donc $|f_{q_k}(y_k) - f_{q_k}(x)| \leq \epsilon$. Ce qui contredit l'inégalité (1.1).

(iii) **(EQ) $\Rightarrow$ (U)**

Soient $x \in D$ et $\epsilon > 0$. Par (EQ), il existe une boule $B(x, \epsilon)$ de centre x et de rayon ϵ et un nombre $N(x, \epsilon)$ tels que $\forall y \in B(x, \epsilon), \forall k \geq N(x, \epsilon)\ |f_k(y) - f_k(x)| < \epsilon$.
Appliquant alors l'hypothèse de convergence point par point, on obtient que

$$|f(y) - f(x)| \leq \epsilon \text{ pour tout } y \in B(x, \epsilon),$$

i.e., la fonction f est continue en x.

Soit D un compact de $\Re^n$ et soit $\epsilon > 0$.
Par hypothèse, à chaque point x de D, on peut associer une boule ouverte $B(x, \epsilon)$ de centre x et de rayon ϵ et un nombre $N(x, \epsilon)$ tels que

$$\forall y \in B(x, \epsilon),\ \forall k \geq N(x, \epsilon)\ |f_k(y) - f_k(x)| < \frac{\epsilon}{3}.$$

On a alors $D \subseteq \bigcup_{x \in D} B(x, \epsilon)$. Comme D est compact, il existe un sous-ensemble fini $\{x_1, \ldots x_m\}$ de D tel que $D \subset \bigcup_{i=1}^m B(x_i, \epsilon)$.

Ainsi si $y \in D$, il existe $j \in \{1, ..., m\}$ tels que $y \in B(x_j, \epsilon)$.
De plus, on a que

$$|f(y) - f_k(y)| \leq |f(y) - f(x_j)| + |f(x_j) - f_k(x_j)| + |f_k(x_j) - f_k(y)|.$$

Comme f est continue en x_j, on a

$$\forall y \in B(x_j, \epsilon)\ |f(y) - f(x_j)| \leq \frac{\epsilon}{3}.$$

La suite $\{f_k\}$ convergeant point par point vers f, on a $\{f_k(x_j)\} \to f(x_j)$ et il existe un entier M_j tel que

$$\forall k \geq M_j\ |f(x_j) - f_k(x_j)| < \frac{\epsilon}{3}.$$

Enfin, les fonctions f_k étant presque équi-continues, on a

$$\forall k \geq N(x_j, \epsilon)\ |f_k(x_j) - f_k(y)| < \frac{\epsilon}{3}.$$

Dès lors

$$|f(y) - f_k(y)| < \epsilon \text{ pour } k \geq \max_{1 \leq i \leq m} (N(x_i, \epsilon), M_i).$$

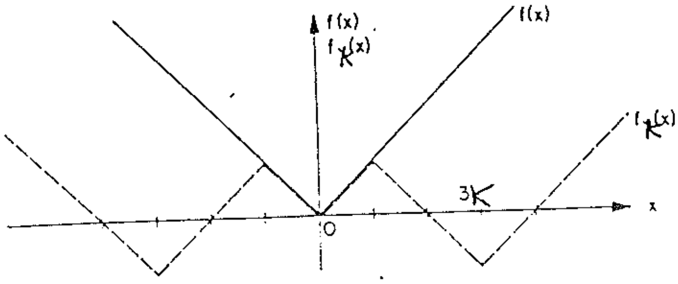
La suite $\{f_k\}$ converge donc bien uniformément vers la fonction f sur D. □

1.2 L'épi-convergence en optimisation

Remarquons d'abord que même sous les hypothèses de la convergence uniforme (U), il n'est pas certain que $\{x_k^*\} \to x^*$ ni même que $\{inf\ f_k\} \to inf\ f$, comme le montre l'exemple suivant: (Rappelons que, dans cette partie, x_k^* et x^* désignent respectivement un minimum de f_k et de f sur $\Re^n$)

Exemple:
Pour tout $x \in \Re$, définissons $f(x) = |x|$ et

$$f_k(x) = \begin{cases} |x| & \text{si } |x| \leq k \\ 2k - |x| & \text{si } k < |x| \leq 3k \\ |x| - 4k & \text{si } |x| > 3k \end{cases}$$



On a que $x^* = 0$, $inf\ f(x) = 0$, $x_k^* = \pm 3k$ et $inf\ f_k(x) = -k$.
On peut vérifier que (U) est satisfait. Cependant, la suite $\{x_k^*\}$ ne possède pas de point d'accumulation et la suite $\{inf\ f_k\}$ ne converge pas vers $inf\ f$.

Théorème 1.2

Supposons satisfaite une des assertions du théorème 1.1.
Alors

(a) $\limsup_{k\to\infty} inf\ f_k \leq inf\ f$,

(b) Si x_k^* *est solution de* (P_k) *et si* x^* *est un point d'accumulation de la suite* $\{x_k^*\}$, *alors* x^* *est solution de (P) et,* $\{x_{q_k}^*\}$ *étant la sous-suite de* $\{x_k^*\}$ *qui converge vers* x^*, *on a*

$$\lim_{k\to\infty}(inf\ f_{q_k}(x)) = inf\ f(x).$$

Preuve

Par le théorème 1.1, on peut supposer que l'assertion (C) est satisfaite.
Posons $f^* = inf\ f(x)$ et $f_k^* = inf\ f_k(x)$.

(a) Soit $x \in \Re^n$ et soit une suite $\{x_k\} \to x$. Par (C) et par la définition de f_k^*, on a que

$$f(x) = \lim_{k\to\infty} f_k(x_k) = \limsup_{k\to\infty} f_k(x_k) \geq \limsup_{k\to\infty}\ f_k^*.$$

Comme ceci est valable pour tout x dans $\Re^n$, en prenant l'infimum sur tous les x, on obtient

$$f^* \geq \limsup_{k\to\infty} f_k^*.$$

(b) Par (C) et par définition de $x_{q_k}^*$, on a

$$f(x^*) \leq \liminf_{k\to\infty} f_{q_k}(x_{q_k}^*) = \liminf_{k\to\infty} f_{q_k}^* \leq \limsup_{k\to\infty} f_{q_k}^*.$$

Montrons ensuite que $\limsup_{k\to\infty} f_{q_k}^* = \limsup_{k\to\infty} f_k^*$.
Puisque toute valeur d'adhérence de la sous-suite $\{f_{q_k}^*\}$ est aussi valeur d'adhérence de la suite $\{f_k^*\}$, on a immédiatement que

$$\limsup_{k\to\infty} f_{q_k}^* \leq \limsup_{k\to\infty} f_k^*.$$

Réciproquement comme $\{x^*_{q_k}\} \to x^*$, on a, par (C), que $\{f_{q_k}(x^*_{q_k})\} \to f(x^*)$. Donc, en utilisant la définition de f^*_k, on a que

$$\begin{aligned} \limsup_{k\to\infty} f^*_k &\leq \limsup_{k\to\infty} f_k(x^*_{q_k}) \\ &= \limsup_{k\to\infty} f_{q_k}(x^*_{q_k}) \\ &= \limsup_{k\to\infty} f^*_{q_k}. \end{aligned}$$

Comme $\limsup_{k\to\infty} f^*_k \leq f(x^*)$ (voir la preuve de (a)), on en déduit

$$f(x^*) \stackrel{def}{=} f^* = \liminf_{k\to\infty} f^*_{q_k} = \limsup_{k\to\infty} f^*_k.$$

□

On remarque que, pour montrer la partie (a), il aurait suffi de supposer que

$$\forall x \in \Re^n \ \exists \text{ une suite } \{z_k\} \to x \text{ telle que } \limsup_{k\to\infty} \ f_k(z_k) \leq f(x).$$

De même, pour obtenir la conclusion de (b), il aurait suffi de supposer (a) et la condition

$$\forall \{x_k\} \to x : f(x) \leq \liminf_{k\to\infty} f_k(x_k).$$

A la vue de ces remarques, on définit un nouveau concept de convergence qui implique les conclusions (a) et (b) du théorème 1.2.

Définition de l'épi-convergence

On dira que la suite $\{f_k\}$ épi-converge vers la fonction f si pour tout $x \in \Re^n$, on a les deux conditions suivantes:
(a) il existe une suite $\{y_k\} \to x$ telle que $\limsup_{k\to\infty} \ f_k(y_k) \leq f(x)$,
(b) pour toute suite $\{x_k\} \to x$, on a $f(x) \leq \liminf_{k\to\infty} f_k(x_k)$.
On notera $\{f_k\} \to_{epi} f$.

Remarquons que cette définition est encore valable si f_k et f sont des fonctions à valeurs dans $\Re \bigcup \{+\infty\}$. A partir de maintenant on supposera donc que c'est le cas.

Utilisant cette définition, le théorème précédent peut se généraliser immédiatement pour donner l'énoncé suivant:

Théorème 1.3

Si $\{f_k\} \to_{epi} f$, alors les deux assertions suivantes ont lieu:

(a) $\limsup_{k\to\infty} inf f_k \leq inf f$

(b) Si x_k^* est solution de (P_k) et si x^* est un point d'accumulation de la suite $\{x_k^*\}$, alors x^* est solution de (P) et, $\{x_{q_k}^*\}$ étant la sous-suite de $\{x_k^*\}$ qui converge vers x^*, on a

$$\lim_{k\to\infty}(inf f_{q_k}) = inf\ f.$$

□

Notons que la définition d'épi-convergence implique que pour tout $x \in \Re^n$, il existe une suite $\{y_k\}$ telle que $\lim_{k\to\infty} f_k(y_k) = f(x)$. Donc, l'épi-convergence impose une forme de convergence globale sur la suite des approximations $\{f_k\}$.

Enfin, remarquons que, dans le théorème 1.3, l'hypothèse

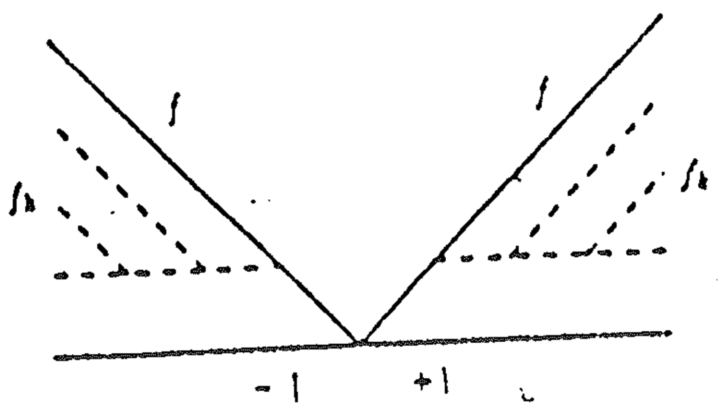
$$x_k^* \in \text{argmin}\{f_k(x)\ t.q.\ x \in \Re^n\}$$

(i.e. x_k^* est minimum global) ne peut être remplacée par un minimum local, comme le montre l'exemple suivant:

Exemple:

Soit à résoudre le problème: $\min\{f(x) = |x|\ t.q.\ x \in \Re^n\}$ et soient les fonctions

$$f_k(x) = \begin{cases} |x| & \text{si } |x| \leq 1, \\ 1 & \text{si } 1 \leq |x| \leq 1 + \frac{2}{k}, \\ |x| - \frac{2}{k} & \text{si } |x| \geq 1 + \frac{2}{k}. \end{cases}$$



La suite des fonctions $\{f_k\}$ épi-converge vers la fonction f. $x_k^* = 1 + \frac{1}{k}$ est un minimum local de f_k et $\lim_{k\to\infty} x_k^* = 1$. Cependant $x = 1$ n'est pas un minimum local de f. Donc, malgré l'épi-convergence de la suite $\{f_k\}$, une suite de minima locaux des fonctions f_k n'a pas pour point d'accumulation un minimum local de f.

1.3 Épi-convergence et convergence point par point

On sait que la convergence continue (C) implique la convergence point par point vers la même limite. Cependant l'épi-convergence à elle seule n'implique pas la convergence point par point, comme le montre l'exemple suivant:

Exemple:

Considérons la suite $\{f_k\}$ définie comme suit:

$$f_{2k}(x) = \begin{cases} 1/k & \text{si } x \text{ est rationnel,} \\ 1 - 1/k & \text{sinon.} \end{cases}$$

$$\begin{aligned} f_{2k+1}(x) &= 1 - f_{2k}(x) \\ &= \begin{cases} 1 - 1/k & \text{si } x \text{ est rationnel,} \\ 1/k & \text{sinon.} \end{cases} \end{aligned}$$

La suite $\{f_{2k}\}$ converge point par point vers la fonction $g(x) = \begin{cases} 0 & \text{si } x \text{ est rationnel,} \\ 1 & \text{sinon.} \end{cases}$

et la suite $\{f_{2k+1}\}$ converge point par point vers la fonction $h(x) = \begin{cases} 1 & \text{si } x \text{ est rationnel,} \\ 0 & \text{sinon.} \end{cases}$

Comme les fonctions $h(x)$ et $g(x)$ sont différentes pour tout $x \in \Re$, la suite $\{f_k\}$ ne converge pas point par point. Cependant elle épi-converge vers la fonction $f(x) = 0$.

En effet, pour chaque $x \in \Re^n$, on peut construire une suite $\{y_k\}$ dont les termes d'indice pair y_{2k} sont rationnels et vérifient $|y_{2k} - x| < 1/k$, et dont les termes d'indice impair y_{2k+1} sont irrationnels et vérifient $|y_{2k+1} - x| < 1/k$.

On a immédiatement que la suite $\{y_k\}$ converge vers x et que la suite $\{f_k(y_k)\}$ étant égale à $2/k$ converge vers 0. Ainsi on a

$$\limsup_{k \to \infty} f_k(y_k) \leq f(x) = 0.$$

D'autre part, il est évident que quelque soit la suite $\{x_k\}$ convergeant vers x, on a

$$0 = f(x) \leq \liminf_{k \to \infty} f_k(x_k).$$

Donc la suite $\{f_k\}$ est bien épi-convergente vers f. Réciproquement, on peut trouver une suite $\{f_k\}$ convergeant point par point mais qui n'épi-converge pas vers sa limite simple. Il suffit d'utiliser la proposition suivante.

Proposition 1.1

Si la suite $\{f_k\}$ épi-converge vers la fonction f, alors f est semi-continue inférieurement.

Preuve

Par l'absurde, supposons que la fonction f n'est pas semi-continue inférieurement en x. Alors il existe $\epsilon > 0$ tel que dans toute boule ouverte de centre x et de rayon $1/k$ $B(x, 1/k)$, on peut trouver un élément z_k satisfaisant

$$f(z_k) < f(x) - \epsilon.$$

Comme la suite $\{f_k\}$ épi-converge vers la fonction f, il existe une suite $\{\xi_k\}$ telle que, pour tout k,

$$\|\xi_k - z_k\| < 1/k \text{ et } f_k(\xi_k) < f(z_k) + 1/k.$$

Dès lors, pour tout k, on a

$$f_k(\xi_k) < f(z_k) + 1/k < f(x) - \epsilon + 1/k,$$

Et donc en passant à la limite inférieure,

$$\liminf_{k \to \infty} f_k(\xi_k) \leq f(x) - \epsilon.$$

Ceci contredit le fait que $f(x) \leq \liminf_{k \to \infty} f_k(\xi_k)$ pour toute suite $\{\xi_k\}$ convergeant vers x. D'où la thèse. □

Au Chapitre 3, on examinera une méthode d'optimisation utilisant une suite d'approximations $\{f_k\}$ qui converge point par point vers la fonction f. L'épi-convergence de cette suite étant importante pour conclure la convergence de la méthode, il sera intéressant de préciser quelles hypothèses sont à ajouter à la semi-continuité inférieure de la fonction f pour garantir l'épi-convergence. Avant de pouvoir répondre à cette question, examinons l'épi-convergence d'un point de vue géométrique.

Chapitre 2

Épi-convergence comme limite d'épigraphes

Dans ce chapitre, on verra que l'épi-convergence de fonctions peut être abordée de façon géométrique en identifiant les fonctions et leurs épigraphes.

2.1 Limite d'ensembles: définition et propriétés

Étant donné une suite de parties non vides $A_1, A_2, A_3, \ldots$ de $\Re^n$, on définit les limites inférieures et supérieures de la suite $\{A_n\}_{n\in N}$ par les formules

$$\text{Liminf} A_n \stackrel{def}{=} \{x \in \Re^n \ : \ \exists \{a_n\} \to x \text{ avec pour tout } n \in N, \ a_n \in A_n\},$$

$$\text{Limsup} A_n \stackrel{def}{=} \{x \in \Re^n \ : \ \exists\, n(1) < n(2) < n(3) < \ldots \text{et} \, \forall k \in N, \ a_k \in A_{n(k)} \text{ avec } \{a_k\} \to \ x \ \}.$$

En d'autres termes, LiminfA_n est l'ensemble formé par toutes les limites possibles de suites $\{a_n\}_{n\in N}$ avec $a_n \in A_n$ pour tout n, alors que LimsupA_n est formé par toutes les valeurs d'adhérence de telles suites.

On remarque qu'on a toujours l'inclusion Liminf $A_n \subseteq$ Limsup A_n.
Si l'égalité a lieu entre ces deux ensembles, on dira que *la suite* $\{A_n\}_{n\in N}$ *converge.* Dans ce cas l'ensemble LimA_n = LiminfA_n = Limsup A_n sera appelé *limite de la suite* $\{A_n\}$.

On peut aussi montrer que les ensembles LiminfA_n, LinsupA_n sont des parties fermées de $\Re^n$ [1]. La proposition suivante établit que dans le cas d'une suite $\{A_n\}_{n\in N}$ monotone, la limite est facilement déterminée.

Proposition 2.1

Si $\{A_n\}_{n\in N}$ est une suite de parties non vides de $\Re^n$ alors Lim A_n existe et vaut $cl(\bigcup_{n\in N} A_n)$ si $\{A_n\}$ est croissante et $cl(\bigcap_{n\in N} A_n)$ si $\{A_n\}$ est décroissante.

Preuve

Montrons, par exemple, le cas où la suite $\{A_n\}$ est croissante. La thèse revient à montrer que

$$\text{Limsup}A_n \subseteq cl(\bigcup_{n\in N} A_n) \subseteq \text{LininfA}_\text{n}.$$

Soit $x \in$ Limsup A_n. Montrons que $x \in cl(\bigcup_{n\in N} A_n)$.
Par définition de Limsup A_n, il existe une sous-suite $\{x_{n(k)}\}_{k\in N}$ convergeant vers x telle que pour tout k, $x_{n(k)} \in A_{n(k)}$.
Donc $x_{n(k)} \in \bigcup_{n\in N} A_n$ pour tout k et $x \in cl(\bigcup_{n\in N} A_n)$.

Montrons à présent la seconde inclusion.
Soit $x \in cl(\bigcup_{n\in N} A_n)$. Alors il existe une suite $\{x_k\}$ qui converge vers x avec $x_k \in \bigcup_{n\in N} A_n$ pour tout k.
Comme $\{A_n\}$ est croissante, pour tout $k \in N$, il existe $n_k \in N$ tel que

$$x_k \in A_n \quad \forall n \geq n_k.$$

On peut supposer que $\{n_k\}$ est strictement croissante.
Dès lors, la suite $\{y_j\}$ définie par

$$y_j = x_{n_k} \text{ si } n_k \leq j \leq n_{k+1}$$

converge vers x et comme $y_j \in A_j, \forall j \geq n_0$, on a que $x \in \text{Liminf}A_n$. □

2.2 Épi-convergence de suites de fonctions

On définit l'épigraphe d'une fonction f de $\Re^n$ dans $\Re\bigcup\{+\infty\}$ comme étant l'ensemble

$$epif \stackrel{def}{=} \{(x,a) \in \Re^n \times \Re \; t.q. \; f(x) \leq a\}.$$

Notons qu'une limite d'épigraphes est encore un épigraphe, plus précisément:

Proposition 2.2

Soit $\{f_k\}$ une suite de fonctions de $\Re^n$ dans $\Re\bigcup\{+\infty\}$. Alors

$$\text{Limsup}(epif_k) = epi(epi - liminf f_k);$$

$$\text{Liminf}(epif_k) = epi(epi - limsupf_k),$$

où

$$(epi - liminf f_k)(x) \stackrel{def}{=} inf\{\liminf f_{q_k}(x_{q_k})\ t.q.\ \{x_{q_k}\} \to x\};$$

et

$$(epi - limsupf_k)(x) \stackrel{def}{=} inf\{\limsup f_k(x_k)\ t.q.\ \{x_k\} \to x\}.$$

__Preuve__

Montrons par exemple la première égalité.
Montrons d'abord que Limsup $(epif_k) \subseteq epi(epi - liminf f_k)$

Soit un couple (x, a) appartenant au premier ensemble, montrons qu'il appartient également au second.
Par hypothèse, il existe une sous-suite $\{(x_{q_k}, a_{q_k})\}$ qui converge vers (x, a) telle que $f_{q_k}(x_{q_k}) \leq a_{q_k}$, pour tout k. Ainsi, en passant à la limite inférieure, on a

$$\liminf f_{q_k}(x_{q_k}) \leq \liminf a_{q_k}.$$

D'où

$$inf\{\liminf f_{q_k}(x_{q_k})\ t.q.\ \{x_{q_k}\} \to x\} \leq a,$$

i.e., $(epi - liminf f_k)(x) \leq a$. D'où la thèse.

Montrons ensuite que Limsup$(epif_k) \supseteq epi(epi - liminf f_k)$
Soit (x, a) tel que $(epi - liminf f_k)(x) \leq a$.
La thèse consiste à montrer qu'il existe une suite $\{(x_{q_k}, a_{q_k})\}$ qui converge vers (x, a) telle que $f_{q_k}(x_{q_k}) \leq a_{q_k}$. Distinguons deux cas:

1^{er} cas: $(epi - liminf f_k)(x) < a$
Par hypothèse, il existe une suite que l'on note $\{x_k\} \to x$ telle que

$$\liminf f_k(x_k) < a,$$

c'est-à-dire, il existe une sous-suite $\{x_{q_k}\} \to x$ telle que $\lim_{k\to\infty} f_{q_k}(x_{q_k}) < a$.
Donc la suite $\{(x_{q_k}, a)\}$ convient.

2^ecas: $(epi - liminf f_k)(x) = a$

On prend la suite $\{a_k\}$ qui converge de façon strictement décroissante vers a et telle que pour tout k, on a $(epi - liminf f_k)(x) < a_k$. Par le premier cas, on a donc que $(x, a_k) \in \text{Limsup}(epi f_k)$ pour tout k. De plus, comme $\text{Limsup}(epi f_k)$ est fermé, on a que $(x, a) \in \text{Limsup}(epi f_k)$. □

On a alors le résultat suivant.

Théorème 2.1

Les assertions suivantes sont équivalentes:

(1) $\{f_k\} \rightarrow_{epi} f$

(2) $\text{Lim}\ (epi f_k) = epi f$

__Preuve__

(1) $\Rightarrow$ (2)

On a toujours que $\text{Liminf}(epi f_k) \subseteq epi f \subseteq \text{Limsup}(epi f_k)$. Par conséquent, il suffit de montrer que $\text{Limsup}(epi f_k) \subseteq epi f \subseteq \text{Liminf}(epi f_k)$.

Soit un x arbitraire.

Tout d'abord, montrons que $\text{Limsup}(epi f_k) \subseteq epi f$.

Soit $(x, a) \in epi(epi - liminf f_k)$. Alors il existe une suite $\{x_{q_k}\} \rightarrow x$ telle que

$$\liminf f_{q_k}(x_{q_k}) \leq a.$$

Or, par définition de l'épi-convergence, on a $\liminf f_k(x_k) \geq f(x)$ pour toute suite $\{x_k\}$ convergeant vers x. Donc $f(x) \leq a$ et $(x, a) \in epi f$.

Montrons ensuite que $epi f \subseteq \text{Liminf}(epi f_k)$. Pour cela, soit $(x, a) \in epi f$.

Par définition de l'épi-convergence, il existe une suite $\{x_k\} \rightarrow x$ telle que $\text{limsup} f_k(x_k) \leq f(x)$. On a donc que $\text{limsup}\ f_k(x_k) \leq a$ c'est-à-dire, par la proposition 2.2,

$$(x, a) \in \text{Liminf}(epi f_k).$$

D'où la première implication est vérifiée.

(2) $\Rightarrow$ (1)

Soit un x arbitraire.

(a) Montrons d'abord qu'il existe une suite $\{y_k\} \to x$ telle que

$$\limsup_{k\to\infty} \ f_k(y_k) \leq f(x).$$

Par hypothèse, on a que $epif \subseteq \text{Liminf}(epif_k)$. Or, le couple $(x, f(x)) \in epif$ car $f(x) \leq f(x)$.
Donc, $(x, f(x)) \in \text{Liminf}(epif_k)$, c'est-à-dire, par la proposition 2.2,

$$inf\{\limsup f_k(x_k) : \{x_k\} \to x\} \leq f(x).$$

Donc, il existe une suite $\{x_k\}$ qui converge vers x telle que $\limsup f_k(x_k) \leq f(x)$.

(b) Montrons ensuite que pour toute suite $\{x_k\} \to x$, on a $f(x) \leq \liminf_{k\to\infty} \ f_k(x_k)$.
Par hypothèse, on a $\text{Limsup}(epif_k) \subseteq epif$.
Or le couple $(x, epi - liminf f_k(x)) \in \text{Limsup}(epif_k)$, car par la proprosition précédente, cela signifie que

$$(epi - liminf f_k)(x) \leq (epi - liminf f_k)(x).$$

On a donc que

$$(x, (epi - liminf f_k)(x)) \in epif,$$

i.e.,

$$f(x) \leq (epi - liminf f_k)(x).$$

Par conséquent,

$$f(x) \leq inf\{liminf f_k(x_k) : \{x_k\} \to x\}.$$

Donc, pour toute suite $\{x_k\} \to x$, on a $f(x) \leq \liminf_k \ f_k(x_k)$.

□

2.3 Théorèmes de convergence

Montrons à présent quelques théorèmes de convergence. Mais tout d'abord, précisons quelques nouvelles notations:
$A \stackrel{def}{=} \{x \in \Re^n \ t.q. \ f(x) = inf f\}$;
$A_k \stackrel{def}{=} \{x \in \Re^n \ t.q. \ f_k(x) = inf f_k\}$;
$\epsilon - A_k \stackrel{def}{=} \{x \in \Re^n \ t.q. \ f_k(x) - \epsilon \leq inf f_k\}$ où $\epsilon > 0$.
(Rappelons que $inf f$ est noté f^* et $inf f_k, f_k^*$.)

Théorème 2.2

Si $\{f_k\} \rightarrow_{epi} f$, *alors*

$$\text{Liminf}(\epsilon - A_k) \subset \text{Limsup}(\epsilon - A_k) \subset \epsilon - A. \tag{2.1}$$

Preuve

Pour vérifier (2.1), il suffit de vérifier la seconde inclusion puisque la première est toujours vraie.
Soit $x \in \text{Limsup}(\epsilon - A_k)$, montrons que $x \in \epsilon - A$.
Par la définition de Limsup, il existe un ensemble d'indices $K \subset \mathrm{N}$ tel que la suite $\{x_k\}_{k\in K}$ converge vers x et tel que l'on ait

$$f_k(x_k) \leq f_k^* + \epsilon.$$

De plus, par la définition d'épi-convergence et de la fonction $epi - liminf f_k$, on a

$$f(x) \leq (epi - liminf f_k)(x) \leq \liminf_{k\in K} f_k(x_k) \leq \limsup_{k\in K} f_k^* + \epsilon.$$

D'autre part, par la définition d'épi-convergence, pour tout y, il existe une suite $\{y_k\}$ convergeant vers y, telle que

$$\limsup_{k\in K} f_k^* + \epsilon \leq \limsup_{k\in K} f_k(y_k) + \epsilon \leq f(y) + \epsilon.$$

Il suit donc, en prenant l'infimum sur les y, que

$$f(x) \leq f^* + \epsilon$$

c'est-à-dire que $x \in \epsilon - A$. □

Théorème 2.3

Si f^* *est finie et si* $\{f_k\} \rightarrow_{epi} f$,
alors, on a

$$\{f_k^*\} \rightarrow f^* \Leftrightarrow A = \bigcap_{\epsilon>0} \text{Liminf}(\epsilon - A_k) \tag{2.2}$$

Preuve

($\Rightarrow$) Montrons d'abord que $A \supseteq \bigcap_{\epsilon>0} \text{Liminf}(\epsilon - A_k)$.
Par (2.1), on a Liminf $(\epsilon - A_k) \subset \epsilon - A$. En prenant l'intersection sur les $\epsilon > 0$, on a

$\bigcap_{\epsilon>0} \mathrm{Liminf}(\epsilon - A_k) \subseteq \bigcap_{\epsilon>0}(\epsilon - A)$. Comme $A = \bigcap_{\epsilon>0} \epsilon - A$, l'inclusion est évidente.
Montrons ensuite que $A \subseteq \bigcap_{\epsilon>0} \mathrm{Liminf}(\epsilon - A_k)$.
Supposons que A est non vide et considérons un point $x \in A$. On a donc $f(x) = f^*$.
Comme $\{f_k\} \rightarrow_{epi} f$, il existe une suite $\{(x_k, a_k)\}$ convergeant vers (x, f^*) et telle que, pour tout $k \in \mathrm{N}$, on a $(x_k, a_k) \in epi f_k$.
La thèse revient alors à montrer que, pour tout $\epsilon > 0$ et pour tout k, $x_k \in \epsilon - A_k$, c'est-à-dire que $f_k(x_k) \leq f_k^* + \epsilon$.
Par l'absurde, supposons qu'il existe un ensemble d'indices $K_\epsilon \subset \mathrm{N}$ tel que pour tout $k \in K_\epsilon$, on a $f_k^* + \epsilon < f_k(x_k) \leq a_k$.
On a alors

$$\lim_k f_k^* + \epsilon = f^* + \epsilon \leq f^* = \lim_k a_k,$$

ce qui est impossible car $\epsilon > 0$.

($\Leftarrow$) Par hypothèse $A = \bigcap_{\epsilon>0} \mathrm{Liminf}(\epsilon - A_k)$. Utilisant la définition de Liminf, on a alors que, pour tout $x \in A$ (i.e. $f(x) = f^*$) et pour tout $\epsilon > 0$, il existe une suite $\{x_k\}$ convergeant vers x et telle, pour tout $k \in \mathrm{N}$, on a $x_k \in \epsilon - A_k$, i.e.,

$$f_k(x_k) \leq \epsilon + f_k^* \text{ pour tout } k \in \mathrm{N}.$$

On a donc, par la défintion d'épi-convergence,

$$f^* = f(x) = (epi - liminf f)(x) \leq \liminf_{k \in N} f_k(x_k) \leq \epsilon + \liminf_{k \in N} f_k^*.$$

Comme par le théorème 1.2, on a toujours $f^* \geq \limsup f_k^*$, on en déduit que

$$f^* = \lim_k f_k^*.$$

$\square$

Chapitre 3

Application à des algorithmes d'optimisation

Le présent chapitre a pour but de donner deux exemples d'applications de l'épi-convergence en optimisation.

3.1 Méthodes barrières

On considère le problème non linéaire suivant:

$$(\mathrm{P})\begin{cases} \text{minimiser} & g_0(x) \\ \text{sous contraintes :} & g_i(x) \leq 0,\ i = 1, 2, ..., m, \end{cases}$$

où pour $i = 0, 1, ..., m$, les fonctions $g_i : \Re^n \to \Re$ sont continues.
Notons $S = \{x\ t.q.\ g_i(x) \leq 0,\ i = 1, 2, ...m\}$ l'ensemble admissible. On supposera que $cl\ int\ S = S$ c'est-à-dire que S est égal à la fermeture de son intérieur.
On définit les fonctions:

$$f(x) = \begin{cases} g_0(x) & \text{si } x \in S \\ +\infty & \text{sinon,} \end{cases}$$

et

$$f_k(x) = g_0(x) + q(\theta_k, x)$$

où les $\theta_k > 0$ convergent de façon strictement croissante vers $+\infty$, quand k tend vers l'infini,

$q :]0, \infty[\times \Re^n \longrightarrow]0, \infty]$ est continue,

$q(\theta, x)$ est finie si $x \in intS$ et vaut $+\infty$ sinon.

la fonction $\theta \longrightarrow q(\theta, x)$ est strictement décroissante vers 0 lorsque $x \in int\ S$.

La fonction q est appelée **fonction barrière**. Voici quelques fonctions barrières connues:

$$q(\theta, x) = -\theta^{-1} \sum_{i=1}^{i=m} [\min(0, g_i(x))]^{-1}$$

$$q(\theta, x) = \theta^{-2} \sum_{i=1}^{i=m} [\min(0, g_i(x))]^{-2}$$

$$q(\theta, x) = -\theta^{-1} \sum_{i=1}^{i=m} \ln[\min(0.5, -g_i(x))].$$

Remarquons que les fonctions f et f_k sont à valeurs dans $\Re \bigcup \{+\infty\}$.

Lemme 3.1

Pour tout $x \in S$ et pour tout $\epsilon > 0$, il existe une suite $\{x_k\}$ $\in S$ convergeant vers x et un nombre k_ϵ tels que

$$q(\theta_k, x_k) \leq \epsilon, \text{ pour tout } k \geq k_\epsilon.$$

Preuve

Soit $\epsilon > 0$. Définissons les ensembles $S_k \stackrel{def}{=} \{y \in S \ t.q. \ q(\theta_k, y) \leq \epsilon\}$.
Pour tout $k \in N$, on a $S_k \subset S_{k+1}$ puisque, par définition, les θ_k sont strictement croissants et, pour x fixé, les fonctions $q(\theta_k, x)$ sont strictement décroissantes.
De plus, on a $cl(\bigcup_{k \in N} S_k) = S$.
En effet,

$cl(\bigcup_{k \in N} S_k) \subseteq S$.
Comme S est fermé, il suffit de montrer que $\bigcup_{k \in N} S_k \subseteq S$. Soit $y \in \bigcup_{k \in N} S_k$. Alors il existe $k \in$ N tel que $y \in S_k$. D'où $q(\theta_k, y) \leq \epsilon$, et donc $y \in int\ S$ car $q(\theta_k, y)$ est finie si et seulement si $y \in int\ S$.

$cl(\bigcup_{k \in N} S_k) \supseteq S$.
Comme $S = cl\ int\ S$, il suffit de montrer que $int\ S \subseteq \bigcup_{k \in N} S_k$.
Soit $y \in int\ S$. Alors par définition de q, on a que $q(\theta_k, y) \to 0$ et donc la thèse.

Finalement la suite $\{S_k\}$ étant croissante, on a Lim $S_k = S$ et donc $x \in S$ est limite d'une suite $\{x_k\}$, $x_k \in S_k$ pour tout $k \in N$. D'où, la thèse. □

Théorème 3.1

La suite $\{f_k\}$ définie pour chaque k par $f_k(.) = g_0(.) + q(\theta_k, .)$ épi-converge vers f.

Preuve

La thèse revient à montrer que

$$\forall x \in \Re^n \quad (epi - liminf f_k)(x) = f(x) = (epi - limsup f_k)(x).$$

Or, on sait que l'on a toujours $(epi - liminf f_k)(x) \leq (epi - limsup f_k)(x)$. Il reste donc à montrer que $(epi - limsup f_k)(x) \leq f(x) \leq (epi - liminf f_k)(x)$.

- $(epi - limsup f_k)(x) \leq f(x)$.
  Si x n'est pas dans S, l'inégalité est évidente, puisque dans ce cas $f(x) = +\infty$.
  Si $x \in S$, par le lemme précédent, on sait que, pour $\epsilon > 0$ donné, on peut trouver une suite $\{x_k\}$ convergeant vers x et telle que, pour k suffisamment grand, on a

$$q(\theta_k, x_k) \leq \epsilon.$$

  De plus, comme g_0 est continue, on peut dire que, pour k suffisamment grand, on a

$$g_0(x_k) - g_0(x) \leq \epsilon.$$

  On a donc,

$$\begin{aligned}
(epi - limsup f_k)(x) &\leq \limsup_{k \in N} f_k(x_k) \\
&\leq \limsup_{k \in N} g_0(x_k) + \limsup_{k \in N} q(\theta_k, x_k) \\
&= \limsup_{k \in N} g_0(x_k) + \limsup_{k \in N} q(\theta_k, x_k) + g_0(x) - g_0(x) \\
&\leq 2\epsilon + f(x)
\end{aligned}$$

  Comme ceci est valable pour tous les $\epsilon > 0$, on a l'inégalité voulue.

- $f(x) \leq (epi - liminf f_k)(x)$.
  Si x n'est pas dans S, l'inégalité est évidente.
  Si $x \in S$, considérons une suite $\{x_k\}$ arbitraire convergeant vers x. Par continuité de g_0, pour tout $\epsilon > 0$ et pour k suffisamment grand, on a

$$g_0(x) - \epsilon \leq g_0(x_k).$$

  De plus, comme $q(\theta, x) \geq 0$, on a

$$\begin{aligned}
f(x) - \epsilon &= g_0(x) - \epsilon \\
&\leq g_0(x_k) + q(\theta_k, x_k) \\
&= f_k(x_k)
\end{aligned}$$

Donc,

$$f(x) - \epsilon \leq \liminf_k f_k(x_k).$$

Comme ceci est valable pour tout $\epsilon > 0$ et pour toute suite $\{x_k\}$ convergeant vers x, on que $f(x) \leq (epi - liminf f_k)(x)$. $\square$

La suite $\{f_k\}$ épi-convergeant vers f on a, par le théorème 1.3, l'affirmation suivante.

Théorème 3.2 *Convergence des méthodes barrières*

Si, pour tout k, x_k^* minimise f_k et si x^* est un point d'accumulation de la suite $\{x_k^*\}$, alors x^* minimise f et donc x^* est solution du problème (P). $\square$

3.2 Méthodes de pénalisations extérieures

On considère le problème d'optimisation:

$$(P) \begin{cases} \text{minimiser} & g_0(x) \\ \text{sous contraintes :} & g_i(x) \leq 0, i = 1, ..., m \\ & g_i(x) = 0, i = m+1, ..., l \end{cases}$$

où pour $i = 0, 1, ..., l$, les fonctions $g_i : \Re^n \rightarrow \Re$ sont continues.
Notons S l'ensemble des solutions admissibles.
On définit les fonctions

$$f(x) = \begin{cases} g_0(x) & \text{si } x \in S \\ +\infty & \text{sinon.} \end{cases}$$

et

$$f_k(x) = g_0(x) + p(\theta_k, x)$$

où les $\theta_k > 0$ convergent de façon strictement croissante vers $+\infty$, quand k tend vers l'infini,

$p :]0, \infty[\times \Re^n \longrightarrow [0, \infty[$ est continue, finie et non négative,

si $x \in S$, $p(\theta, x) = 0$,

la fonction $\theta \longrightarrow p(\theta, x)$ est croissante de façon uniforme vers $+\infty$ sur tout ensemble compact de $\Re^n \backslash S$.

La fonction p est appelée **fonction de pénalisation extérieure**. Voici un exemple de fonction de pénalisation extérieure:

$$p(\theta, x) = \theta \sum_{i=1}^{i=m} [\max(0, g_i(x))]^\alpha + \sum_{i=m+1}^{i=l} |g_i(x)|^\beta$$

avec $\alpha \geq 1$ et $\beta \geq 1$.

La convergence des méthodes de pénalisations extérieures est établie dans le théorème suivant.

Théorème 3.3

Si la suite $\{f_k\}$ est croissante et converge point par point vers f et si soit les fonctions f_k, $k = 1, 2, \ldots$, soit la fonction f sont semi-continues inférieurement, alors la suite $\{f_k\}$ épi-converge vers f.

Preuve

Étant donné la croissance des fonctions f_k, on a l'inclusion suivante:

$$epif_{k+1} \subset epif_k, \text{ pour tout } k = 1, 2, \ldots,$$

et donc $\mathrm{Lim}_k(epif_k) = cl(\bigcap_{k=1}^{\infty} epif_k)$.
D'autre part on peut facilement voir que

$$epif = \bigcap_{k=1}^{\infty} epif_k.$$

Par l'hypothèse de semi-continuité, soit $epif$ est fermé, soit chacun des $epif_k$, pour $k = 1, 2, \ldots$ est fermé. Dans tous les cas $\bigcap_{k=1}^{\infty} epif_k$ est fermé. Et donc $epif = \mathrm{Lim}_k(epif_k)$. Dès lors la suite $\{f_k\}$ épi-converge vers f et la conclusion est immédiate. $\square$

Dans notre cas, on a toujours que f_k est continue pour tout k puisque g_0 et p le sont. Dès lors les hypothèses du théorème 3.3 sont vérifiées et les méthodes de pénalisations extérieures sont bien convergentes, i.e., tout point d'accumulation de la suite $\{x_k^*\}$ est solution du problème (P).

Partie II

Emboîtement épigraphique

Chapitre 4

Emboîtement épigraphique

Dans la première partie, on a vu que si la suite $\{f_k\}$ est une suite de fonctions de $\Re^n$ dans $\Re\bigcup\{+\infty\}$ qui épi-converge vers une fonction f de $\Re^n$ dans $\Re\bigcup\{+\infty\}$, alors tout point d'accumulation de la suite des minima de $f_k, k = 1, 2, ...$, est un minimum de f. L'épi-convergence de la suite $\{f_k\}$ est une condition suffisante pour obtenir cette propriété. Cependant elle n'est pas une condition nécessaire comme l'indique l'exemple suivant:

Exemple:
Soit à résoudre le problème: $\min\{f(x) = |x| \ t.q.\ x \in \Re^n\}$, en considérant les fonctions approximantes $f_k(x) = \frac{|x|}{k} - \frac{1}{k}$. On a pour minima des fonctions f_k, $x_k = 0$ pour tout k et $\lim_{k\to\infty} x_k = 0$, qui est bien la solution du problème. De plus, remarquons que $\lim_{k\to\infty} f_k(x_k) = f(0)$ (Ce qui était une des conséquences de l'épi-convergence de la suite $\{f_k\}$ vers la fonction f). Cependant, la suite $\{f_k\}$ n'épi-converge pas vers f, quelque soit le voisinage de 0.

Le théorème suivant indique que l'épi-convergence de la suite $\{f_k\}$ vers f est trop stricte pour obtenir que tout point d'accumulation de la suite $\{x_k\}$ soit solution du problème. Avant de l'énoncer, introduisons un nouveau concept de convergence.

Défintion de l'emboîtement épigraphique

On dira que *l'épigraphe de f est emboîté dans la suite des épigraphes de f_k* si

$$epi f \subseteq \mathrm{Liminf}_{k\to\infty}(epi f_k).$$

Comme

$$\mathrm{Liminf}_{k\to\infty}(epi f_k) = \{(x, a)\ t.q.\ a \geq inf\{\limsup f_k(y_k);\ \{y_k\} \to x\}\},$$

cette propriété est équivalente à

$$f(x) \geq inf\{\limsup f_k(y_k);\ \{y_k\} \to x\} \quad \text{pour tout } x.$$

En quelque sorte, l'emboîtement épigraphique revient à imposer que les approximations successives f_k fournissent une borne inférieure sur la fonction f. De plus, il est immédiat que si $\{f_k\} \to_{epi} f$ alors l'emboîtement épigraphique a lieu.

Théorème 4.1

Soit $\{x_k\}$ *une suite définie par*

$$x_k \in \operatorname{argmin}\{f_k(x)\ t.q.\ x \in \Re^n\} \text{ pour tout } k.$$

Si

$$\lim_{k\in K} f_k(x_k) = f(\bar{x}) \text{ et } epi(f) \subseteq \operatorname{Liminf}_{k\in K} epi(f_k) \text{ quand } \lim_{k\in K} x_k = \bar{x},$$

alors tout point d'accumulation de la suite $\{x_k\}_{k=1}^{\infty}$ *est contenu dans l'ensemble*

$$\operatorname{argmin}\{f(x)\ t.q.\ x \in \Re^n\}.$$

Preuve

Soit x^* un point d'accumulation de la suite $\{x_k\}_{k=1}^{\infty}$ et soit K l'ensemble des indices tel que $\lim_{k\in K} x_k = x^*$. Montrons que $f(x^*) \leq f(x)$ pour tout $x \in \Re^n$.
Par hypothèse, on a

$$\lim_{k\in K} f_k(x_k) = f(x^*) \text{ et } epi(f) \subseteq \operatorname{Liminf}_{k\in K} epi(f_k).$$

Soit $x \in \Re^n$ et $\{y_k\}$ une suite convergeant vers x. Par définition de x_k, on a

$$f_k(x_k) \leq f_k(y_k) \text{ pour chaque } k \in K.$$

Dès lors

$$f(x^*) = \lim_{k\in K} f_k(x_k) \leq \limsup_{k\in K} f_k(y_k).$$

La suite $\{y_k\}$ étant arbitraire, on en déduit

$$f(x^*) \leq inf\{\limsup_{k\in K} f_k(y_k); \{y_k\} \to x\}.$$

Comme par hypothèse, $epi f$ est emboîté dans la suite des épigraphes $\{epi f_k\}_{k\in K}$, le membre de droite de l'inégalité précédente est inférieure à $f(x)$ et donc

$$f(x^*) \leq f(x).$$

Comme x a été choisi de façon arbitraire, on a la thèse. $\square$

Bien que l'emboîtement épigraphique soit moins strict que l'épi-convergence, il est à noter que cette hypothèse ne peut expliquer la convergence de nombreux algorithmes d'optimisation, notamment ceux qui recherchent des points stationnaires comme les méthodes de type Newton.

On va donc s'intéresser à une nouvelle forme de convergence qui sera applicable à une classe plus large d'algorithmes. On montrera que la convergence peut être assurée si les dérivées directionnelles de f et des fonctions approximantes f_k satisfont la propriété d'emboîtement épigraphique. Cette théorie englobe une large classe de problèmes en optimisation différentiable et non différentiable.

Commençons par quelques notations. Soit la fonction $g : \Re^n \times \Re^n \to \Re$ et considérons le problème suivant:
Déterminer $x \in S \subseteq \Re^n$, où S est défini comme suit:

$$S \stackrel{def}{=} \{x \in \Re^n \ t.q. \ g(x,d) \geq 0, \ \forall d \in \Re^n\}.$$

Dans une méthode d'approximations successives, on construit une suite d'approximations $g_k : \Re^n \times \Re^n \to \Re$ et une suite $\{x_k\}$ telle que pour chaque k on a,

$$x_k \in S_k \stackrel{def}{=} \{x \in \Re^n \ t.q. \ g_k(x,d) \geq 0, \ \forall d \in \Re^n\}. \tag{4.1}$$

Le résultat suivant fournit une condition pour que tout point d'accumulation de la suite $\{x_k\}$ soit contenu dans S. Notons qu'on supposera toujours S non vide.

Proposition 4.1

Soit $\{x_k\}$ une suite générée selon (4.1) et supposons que $\{x_k\}_{k\in K} \to \bar{x}$. Si pour tout $d \in \Re^n$, il existe une suite $\{d_k\}$ convergeant vers d telle que

$$\limsup_{k\in K} g_k(x_k, d_k) \leq g(\bar{x}, d),$$

alors $\bar{x} \in S$.

Preuve

Pour tout d_k, on a

$$g_k(x_k, d_k) \geq 0, \ \forall k \in K.$$

Donc $\limsup_{k\in K} g_k(x_k, d_k) \geq 0$ et $g(\bar{x}, d) \geq 0$. $\square$

Les hypothèses de ce lemme peuvent s'écrire de façon équivalente comme

$$epi(g(\bar{x},.)) \subseteq \text{Liminf}_{k\in K} epi(g_k(x_k,.)) \text{ quand } \{x_k\}_{k\in K} \to \bar{x}.$$

Donc les hypothèses de ce lemme ne sont ni plus ni moins qu'une forme d'emboîtement épigraphique des fonctions g et g_k.

On peut établir un résultat semblable quand des contraintes sont introduites. Dans ce cas, on note $X \subseteq \Re^n$ l'ensemble admissible, et le problème consiste à trouver un point x de l'ensemble

$$S \stackrel{def}{=} \{x \in X \ t.q. \ g(x,d) \geq 0, \ \forall d \in D(x;X)\},$$

où

$$D(x;X) \stackrel{def}{=} \{d \in \Re^n \ t.q. \ \exists\{x_k\} \subset X, \{x_k\} \to x \text{ et } d = \lim_{k\to\infty} \frac{(x_k - x)}{\|x_k - x\|}\}. \tag{4.2}$$

En plus des approximations g_k de g, on considèrera des approximations X_k de X. La suite $\{x_k\}$ est définie comme suit:

$$x_k \in S_k \stackrel{def}{=} \{x \in X_k \ t.q. \ g_k(x,d) \geq 0, \ \forall d \in D(x;X_k)\}. \tag{4.3}$$

On obtient un résultat analogue à la proposition 4.1

Proposition 4.2

Soit $\{x_k\}$ une suite générée selon (4.3), et supposons que $\{x_k\}_{k\in K} \to \bar{x}$. Si $\bar{x} \in X$ et si, pour tout $d \in D(\bar{x};X)$, il existe une suite $\{d_k\} \in D(x_k;X_k)$ convergeant vers d telle que

$$\limsup_{k\in K} g_k(x_k,d_k) \leq g(\bar{x},d),$$

alors $\bar{x} \in S$. □

La preuve est aussi évidente que la preuve précédente.

Dans les chapitres suivants on va montrer comment l'emboîtement épigraphique peut être utilisé pour prouver la convergence des méthodes de type Newton en optimisation sans contrainte (Ch5), des méthodes de programmation quadratique successive pour des problèmes avec contraintes (Ch6) et de la méthode du point proximal en optimisation convexe non différentiable(Ch7).

Chapitre 5

Approximations en optimisation différentiable sans contrainte

Dans ce chapitre, on considère le problème sans contrainte

$$\begin{cases} \text{minimiser} & f(x) \\ \text{sous contraintes} & x \in \Re^n \end{cases}$$

où f est une fonction différentiable de $\Re^n$ dans $\Re$.

On sait qu'une condition nécessaire d'optimalité d'un point $\bar{x} \in \Re^n$ est que le gradient de f en $\bar{x}$, noté $\nabla f(\bar{x})$, soit nul. Cette condition peut encore s'écrire

$$\nabla f(\bar{x})^T d \geq 0, \; \forall d \in \Re^n. \tag{5.1}$$

Sans hypothèse de convexité sur f, la plupart des méthodes numériques connues fournissent non pas un minimum mais un point $\bar{x}$ satisfaisant (5.1). Un tel point est appelé *point stationnaire de f*. Dès lors, si on pose $g(x,d) = \nabla f(x)^T d$, le problème à résoudre sera celui de la recherche d'un point x tel que

$$g(x,d) \geq 0, \; \forall d \in \Re^n.$$

D'autre part très souvent la suite $\{x_k\}$ des itérés est constuite de la façon suivante:

$$x_{k+1} = x_k + \alpha_k s_k, \tag{5.2}$$

où α_k est un scalaire positif ou nul et s_k une "direction de descente".

Le lemme suivant fournit des conditions suffisantes pour la convergence globale d'une méthode basée sur un tel schéma.

Lemme 5.1

Soit g une fonction de $\Re^n \times \Re^n$ dans $\Re$ et soit $\{g_k\}$ une suite de fonctions définies sur les mêmes espaces. Supposons qu'une suite de points $\{x_k\}$ est générée selon (5.2) et vérifie les propriétés suivantes:

(a) $s_k \in S_k = \{s \text{ t.q. } g_k(x_k + s, d) \geq 0, \ \forall d \in \Re^n\}$;

(b) pour toute direction $d \in \Re^n$, il existe une suite $\{d_k\}$ convergeant vers d telle que

$$\limsup_{k \in K} g_k(x_k, d_k) \leq g(\bar{x}, d), \text{ quand } \{x_k\}_{k \in K} \to \bar{x};$$

(c) $\lim_{k \to \infty} g_k(x_k + s_k, d_k) - g_k(x_k, d_k) = 0$ *pour toute suite $\{d_k\}$ convergeant vers d.*

Alors tout point d'accumulation de la suite $\{x_k\}$ est contenu dans

$$S = \{x \text{ t.q. } g(x, d) \geq 0, \ \forall d \in \Re^n\}.$$

Preuve

Soit $\bar{x}$ un point d'accumulation de la suite $\{x_k\}$, et soit K l'ensemble des indices tels que la suite $\{x_k\}_{k \in K}$ converge vers $\bar{x}$. L'hypothèse (a) assure que

$$g_k(x_k + s_k, d) \geq 0, \ \forall d \in \Re^n.$$

De plus, par (b) et (c), on a que pour tout $d \in \Re^n$, il existe une suite $\{d_k\}$ convergeant vers d telle que

$$0 \leq \limsup_{k \in K} g_k(x_k + s_k, d_k) = \limsup_{k \in K} g_k(x_k, d_k) \leq g(\bar{x}, d).$$

D'où la thèse. □

Remarquons que ce lemme ne contient aucune référence quant au choix de longueur de pas α_k. En fait, ils sont essentiellement choisis de façon à assurer l'existence d'un point d'accumulation de la suite $\{x_k\}$.

Examinons à présent quelques méthodes particulières qui cherchent un point stationnaire d'une fonction f différentiable.

Méthodes de type Newton

Dans ces méthodes, on considère en chaque point engendré x_k une approximation du second ordre f_k de la fonction f définie par

$$f_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + 1/2(x - x_k)^T Q_k(x - x_k). \tag{5.3}$$

Pour la méthode de Newton, Q_k représente la matrice hessienne de f au point x_k, mais de façon plus générale pour une méthode de type Newton, Q_k représente une matrice symétrique définie positive. On prend alors pour fonction g_k la dérivée directionnelle de f_k définie de la façon suivante:

$$g_k(x, d) = [\nabla f(x_k) + Q_k(x - x_k)]^T d. \tag{5.4}$$

Si on pose $s = x - x_k$, alors on a

$$f_k(x_k + s) = f(x_k) + \nabla f(x_k)^T s + 1/2 s^T Q_k s, \tag{5.5}$$

$$g_k(x_k + s, d) = [\nabla f(x_k) + Q_k s]^T d. \tag{5.6}$$

Dans la méthode de Newton, on calcule alors s_k un point stationnaire de (5.5) et on définit

$$x_{k+1} = x_k + s_k.$$

Dans une méthode de type Newton, on définit

$$x_{k+1} = x_k + \alpha_k s_k \text{ avec } \alpha_k > 0.$$

La convergence globale de la méthode est obtenue en s'assurant que la suite $\{f(x_k)\}$ est une suite décroissante et que l'ensemble $\{x \ t.q. \ f(x) \leq f(x_1)\}$ est un ensemble compact. En général, on obtient cette décroissance de la valeur de la fonction objectif via une recherche linéaire dans la direction s_k, avec le choix d'une longueur de pas α_k.

Le théorème suivant donne des conditions sous lesquelles la suite de points $\{x_k\}$ générée selon (5.2), avec les fonctions f_k et g_k définies comme en (5.5) et (5.6) respectivement, possède des points d'accumulation dans l'ensemble S. Avant cela, notons que, lorsque la matrice Q_k est définie positive, le minimum de (5.5) est donné par

$$s_k = -Q_k^{-1} \nabla f(x_k), \tag{5.7}$$

et que cette direction s_k est une direction de descente pour la fonction f en x_k. Si $\nabla f(x_k) \neq 0$, alors on a

$$f_k(x_k + s_k) - f_k(x_k) = -1/2 \nabla f(x_k)^T Q_k^{-1} \nabla f(x_k) < 0, \tag{5.8}$$

$$g_k(x_k + s_k, d) - g_k(x_k, d) = -\nabla f(x_k)^T d. \tag{5.9}$$

Théorème 5.1

Supposons que la fonction f soit continûment différentiable, bornée inférieurement et que l'ensemble $\{x \text{ t.q. } f(x) \leq f(x_1)\}$ soit compact. Considérons les fonctions f_k et g_k définies comme en (5.5) et (5.6), où les matrices Q_k, symétriques et définies positives, sont telles que, si λ_k est la plus petite valeur propre de Q_k, alors $\liminf_{k\to\infty} \lambda_k > 0$. Posons $\Delta_k = f_k(x_k + s_k) - f_k(x_k)$. Supposons de plus que

(a) $s_k \in S_k = \{s \text{ t.q. } g_k(x_k + s, d) \geq 0, \ \forall d \in \Re^n\}$;

(b) Les pas $\alpha_k \geq 0$ *sont choisis de telle sorte que* $f(x_k + \alpha_k s_k) \leq f(x_k)$ *pour tout* k;

(c) $\lim_{k\to\infty} \Delta_k = 0$.

Alors, tout point d'accumulation de la suite $\{x_k\}$ appartient à l'ensemble des points stationnaires

$$S = \{x \text{ t.q.} \nabla f(x)^T d \geq 0, \ \forall d \in \Re^n\}.$$

Preuve

La preuve se base sur le lemme 5.1. La conditon (a) est trivialement satisfaite, il reste donc à montrer que les conditons (b) et (c) du lemme 5.1 sont vérifiées.
Par hypothèse, la suite $\{f(x_k)\}$ étant décroissante, la suite $\{x_k\}$ est contenue dans un ensemble compact. Il existe donc un ensemble d'indices K tel que la suite $\{x_k\}_{k\in K}$ converge vers $\bar{x}$. De plus, la différentiabilité continue de la fonction f assure que la suite $\{\nabla f(x_k)\}_{k\in K}$ converge vers $\nabla f(\bar{x})$. La condition (b) du lemme 5.1 est donc satisfaite.
En vertu de (5.9), la condition (c) du lemme 5.1 sera satisfaite si la suite des gradients $\{\nabla f(x_k)\}$ converge vers 0. Cette propriété découle de (5.8) car $\lim_{k\to\infty} \Delta_k = 0$, et $\liminf_{k\to\infty} \lambda_k > 0$. □

L'hypothèse (b) de ce théorème ne spécifie pas de façon précise le choix de la longueur de pas. On va montrer que cette hypothèse est satisfaite pour des règles très connues telles que la règle d'Armijo, la règle de Goldstein et la règle de Wolfe. Rappelons d'abord comment α_k est choisi dans ces règles:

Règle d'Armijo

$$f(x_k + \alpha_k s_k) \leq f(x_k) + \epsilon[\nabla f(x_k)^T s_k]\alpha_k, \tag{5.10}$$

$$f(x_k + \eta\alpha_k s_k) > f(x_k) + \epsilon[\nabla f(x_k)^T s_k]\eta\alpha_k, \tag{5.11}$$

où $\epsilon \in (0,1)$ et $\eta > 1$ sont des paramètres.

Règle de Goldstein

$$f(x_k + \alpha_k s_k) \leq f(x_k) + \epsilon[\nabla f(x_k)^T s_k]\alpha_k, \tag{5.12}$$

$$f(x_k + \alpha_k s_k) \geq f(x_k) + (1-\epsilon)[\nabla f(x_k)^T s_k]\alpha_k, \tag{5.13}$$

où $\epsilon \in (0, 0.5)$.

Règle de Wolfe

$$f(x_k + \alpha_k s_k) \leq f(x_k) + \epsilon[\nabla f(x_k)^T s_k]\alpha_k, \tag{5.14}$$

$$\nabla f(x_k + \alpha_k s_k)^T s_k \geq \eta[\nabla f(x_k)^T s_k], \tag{5.15}$$

avec $0 < \epsilon < \eta < 1$.

Le théorème suivant montre que quand la suite $\{\alpha_k\}$ est choisie de façon à respecter ces règles, tout point d'accumulation de la suite $\{x_k\}$ est contenu dans S.

Théorème 5.2

Supposons que la fonction f soit continûment différentiable, bornée inférieurement et que l'ensemble $\{x \text{ t.q. } f(x) \leq f(x_1)\}$ soit compact. Considérons les fonctions f_k et g_k définies comme en (5.5) et (5.6), où les matrices Q_k, symétriques et définies positives, sont telles que, si λ_k est la plus petite valeur propre de Q_k, alors $\liminf_{k\to\infty} \lambda_k > 0$. Posons $\Delta_k = f_k(x_k + s_k) - f_k(x_k)$. Supposons de plus que

(a) $s_k \in S_k = \{s \text{ t.q. } g_k(x_k + s, d) \geq 0,\ \forall d \in \Re^n\}$;

(b) Les pas $\alpha_k \geq 0$ sont choisis selon une des trois règles citées précédemment.

Alors, tout point d'accumulation de la suite $\{x_k\}$ appartient à l'ensemble

$$S = \{x \text{ t.q.} \nabla f(x)^T d \geq 0,\ \forall d \in \Re^n\}.$$

Preuve

La preuve se base sur le théorème 5.1. La condition (a) du théorème 5.1 est trivialement satisfaite. La condition (b) est satisfaite puisque par (5.7)

$$\nabla f(x_k)^T s_k = -\nabla f(x_k) Q_k^{-1} \nabla f(x_k).$$

De plus, par la définie positivité des matrices Q_k, $\nabla f(x_k)^T s_k$ est strictement négatif. Par conséquent, le second terme de l'inégalité suivante est négatif

$$f(x_k + \alpha_k s_k) - f(x_k) \leq \epsilon[\nabla f(x_k)^T s_k]\alpha_k.$$

D'où la condition (b) est vérifiée.

Il reste ensuite à montrer que $\lim_{k\to\infty} \Delta_k = 0$ lorsque la suite $\{\alpha_k\}$ est déterminée selon une des trois règles précédentes. On peut supposer qu'aucun point x_k n'est stationnaire c'est-à-dire que $\nabla f(x_k)$ est non nul pour tout k. Comme la fonction f est bornée inférieurement, la suite décroissante $\{f(x_k)\}$ converge vers un nombre réel et on a

$$\lim_{k\to\infty} f(x_{k+1}) - f(x_k) = 0.$$

De plus, par la première condition des règles et par (5.7) et (5.8), on a les inégalités suivantes:

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\le \epsilon[\nabla f(x_k)^T s_k]\alpha_k, \\ &= \epsilon[-\nabla f(x_k) Q_k^{-1} \nabla f(x_k)]\alpha_k, \\ &= 2\epsilon[f_k(x_k + s_k) - f_k(x_k)]\alpha_k, \\ &= 2\epsilon\Delta_k\alpha_k, \\ &< 0. \end{aligned}$$

Donc

$$\lim_{k\to\infty} \alpha_k\Delta_k = 0.$$

Si $\liminf_{k\to\infty} \alpha_k > 0$, alors $\lim_{k\to\infty} \Delta_k = 0$. Supposons donc que $\liminf_{k\to\infty} \alpha_k = 0$. Étant donné que $\liminf_{k\to\infty} \lambda_k > 0$, la suite $\{Q_k^{-1}\}$ est bornée. Donc, comme la suite $\{x_k\}$ est bornée, elle possède une sous-suite convergente et il existe un ensemble d'indices K tel que

$$\{\alpha_k\}_{k\in K} \to 0.$$

et

$$\{x_k\}_{k\in K} \to \bar{x}.$$

De plus, comme la fonction f est continûment différentiable, on a que

$$\{\nabla f(x_k)\}_{k\in K} \to \nabla f(\bar{x}).$$

Comme la suite $\{Q_k^{-1}\}$ est bornée, elle a une sous-suite convergente, et donc, par la défintion de s_k en (5.7), on obtient

$$\{s_k\}_{k\in K} \to \bar{s}.$$

Utilisant ensuite la condition (5.11) de la règle d'Armijo, on a

$$\frac{f(x_k + \eta\alpha_k s_k) - f(x_k)}{\eta\alpha_k} > \epsilon[\nabla f(x_k)^T s_k] \quad \text{pour tout } k \in K.$$

Comme $\alpha_k \to 0$ et $s_k \to \bar{s}$, le membre de gauche de cette expression converge vers $\nabla f(\bar{x})^T\bar{s}$, et le membre de droite vers $\epsilon\nabla f(\bar{x})^T\bar{s}$. Mais comme $\epsilon \in (0,1)$ et $\nabla f(x_k)^T s_k \le 0$, on a

$\nabla f(\bar{x})^T \bar{s} = 0$.
Si on utilise la condition (5.13) de la règle de Goldstein, on a

$$\frac{f(x_k + \alpha_k s_k) - f(x_k)}{\alpha_k} \geq (1 - \epsilon)[\nabla f(x_k)^T s_k].$$

Et en passant à la limite, on a

$$\nabla f(\bar{x})^T \bar{s} \geq (1 - \epsilon) \nabla f(\bar{x})^T \bar{s}.$$

Comme précédemment, on en déduit que $\nabla f(\bar{x})^T \bar{s} = 0$.

Enfin, par passage à la limite dans la condition (5.15) de la règle de Wolfe, on a

$$\nabla f(\bar{x})^T \bar{s} \geq \eta \nabla f(\bar{x})^T \bar{s},$$

ce qui implique aussi que $\nabla f(\bar{x})^T \bar{s} = 0$.
Ainsi quelle que soit la règle pour les longueurs de pas, on a toujours

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

On a donc a $\limsup_{k\to\infty} \Delta_k = 0$, puisque, par (5.8), $\Delta_k = 1/2 \nabla f(x_k)^T s_k$. Il reste donc à vérifier que $\liminf_{k\to\infty} \Delta_k = 0$ pour obtenir la thèse.
Par l'absurde, supposons que $\liminf_{k\to\infty} \Delta_k < 0$. Alors il existe un ensemble d'indices K_1 tel que $\lim_{k\in K_1} \Delta_k < 0$.
Comme $\lim_{k\to\infty} \alpha_k \Delta_k = 0$, on a que $\lim_{k\in K_1} \alpha_k = 0$. En raisonnant comme précédemment il existe un ensemble d'indices $K_2 \subseteq K_1$ tel que $\{\Delta_k\}_{k\in K_2} \to 0$, ce qui contredit notre hypothèse. Donc $\liminf_{k\to\infty} \Delta_k = 0$. □

Remarquons enfin que le point de vue adopté ici est aussi applicable aux méthodes qui utilisent une minimisation inexacte à chaque étape. Par exemple, considérons les méthodes de Newton inexactes.

Méthodes de Newton inexactes

Pour un problème sans contraintes, les équations de Newton peuvent être résolues avec un certain degré d'incertitude, i.e.,

$$\nabla f(x_k) + Q_k(x - x_k) - \epsilon_k e = 0,$$

où e représente un vecteur unité et ϵ_k un scalaire. Pour que la suite $\{f(x_k)\}$ soit décroissante, i.e., vérifie la condition (b) du théorème 5.1, il faut qu'elle satisfasse $\nabla f(x_k)^T s_k < 0$, pour tout k. On trouve qu'il faut imposer que

$$\epsilon_k < \frac{\nabla f(x_k)^T Q_k^{-1} \nabla f(x_k)}{\nabla f(x_k)^T Q_k^{-1} e}.$$

Enfin, cherchons la condition à imposer pour vérifier la condition (c) du théorème 5.1. On a que

$$\begin{aligned}
\Delta_k &= f_k(x_k + s_k) - f_k(x_k), \\
&= \nabla f(x_k)^T s_k + 1/2 s_k^T Q_k s_k \text{ par (5.5)}, \\
&\leq \nabla f(x_k)^T s_k + s_k^T Q_k s_k, \\
&= \epsilon_k e s_k
\end{aligned}$$

Il faut donc imposer que la suite $\{\epsilon_k e d\}$ converge vers 0 pour toute direction $d \in \Re^n$ ou tout simplement que la suite $\{\epsilon_k\}$ converge vers 0. Dans ces conditions, le théorème 5.1 est encore valable pour la méthode de Newton inexacte.

Chapitre 6

Approximations en optimisation différentiable avec contraintes

Les résultats de convergence obtenus pour les problèmes sans contraintes peuvent se généraliser aux problèmes avec contraintes de la forme suivante:

$$\begin{cases} \text{minimiser} & f(x) \\ \text{sous contraintes} & x \in X. \end{cases}$$

où f est une fonction différentiable de $\Re^n$ dans $\Re$ et X une partie fermée non vide de $\Re^n$. On sait qu'une condition nécessaire d'optimalité d'un point $\bar{x} \in X$ est que

$$\nabla f(\bar{x})^T d \geq 0 \quad \forall d \in D(\bar{x}, X) \tag{6.1}$$

où $D(\bar{x}, X)$ désigne l'ensemble des directions admissibles en $\bar{x}$ pour X.
Sous hypothèse de convexité sur f et X, la plupart des méthodes numériques connues fournissent non pas un minimum mais un point $\bar{x}$ satisfaisant (6.1). Dès lors si on pose $g(x, d) = \nabla f(x)^T d$, le problème à résoudre sera celui de la recherche d'un point $\bar{x} \in X$ tel que

$$g(\bar{x}, d) \geq 0 \quad \forall d \in D(\bar{x}, X).$$

Comme dans le chapitre précédent la suite $\{x_k\}$ des itérés est construite suivant le schéma

$$x_{k+1} = x_k + \alpha_k s_k$$

où α_k est un scalaire positif et s_k une "direction de descente."

De façon équivalente au lemme 5.1, nous avons le lemme suivant:

Lemme 6.1

Soit g une fonction de $\Re^n \times \Re$ dans $\Re$ et soit $\{g_k\}$ une suite de fonctions définies sur les mêmes espaces, et soient une suite de points $\{x_k\}$ générée selon (5.2). Soit une suite $\{X_k\}$ d'approximations de l'ensemble X. Supposons qu'une sous-suite $\{x_k\}_{k\in K}$ converge vers $\bar{x}$ où $\bar{x} \in X$.
Supposons de plus que

(a) $s_k \in S_k = \{s \text{ t.q. } g_k(x_k + s, d) \geq 0, \ \forall d \in D(x_k + s; X_k) \text{ et } x_k + s \in X_k\}$;

(b) *pour toute direction $d \in D(\bar{x}; X)$, il existe $d_k \in D(x_k + s; X_k)$ telle que*

$$\{d_k\}_{k\in K} \to d \text{ et } \limsup_{k\in K} g_k(x_k, d_k) \leq g(\bar{x}, d);$$

(c) $\lim_{k\in K} g_k(x_k + s_k, d_k) - g_k(x_k, d_k) = 0$ *pour toute suite $\{d_k\}$ convergeant vers $d \in D(\bar{x}; X)$.*

Alors $\bar{x}$ est contenu dans l'ensemble

$$S = \{x \text{ t.q. } g(x, d) \geq 0, \ \forall d \in D(x; X)\}.$$

□

La preuve est tout à fait semblable à celle du lemme 5.1.

Afin d'illustrer le résultat précédent, considérons la méthode de programmation quadratique successive S.Q.P. pour résoudre le problème suivant:

$$\begin{cases} \text{minimiser} & f(x) \\ \text{sous contraintes} & c_i(x) \geq 0, \quad i \in I, \\ & c_i(x) = 0, \quad i \in E. \end{cases} \tag{6.2}$$

où toutes les fonctions f, c_i, où $i \in I \bigcup E$, sont continûment différentiables.
L'ensemble des solutions admissibles est noté X. Comme décrit précédemment, l'objectif est de trouver un point stationnaire, c'est-à-dire un point $\bar{x} \in X$ tel que

$$g(\bar{x}, d) = \nabla f(\bar{x})^T d \geq 0, \ \forall d \in D(\bar{x}; X), \tag{6.3}$$

où $D(\bar{x}; X)$ désigne l'ensemble des directions admissibles en $\bar{x}$ pour X.

En supposant qu'une condition de qualification de contraintes est satisfaite pour tout $\bar{x}$ satisfaisant (6.3), on recherche un point $\bar{x}$ satisfaisant les conditions nécessaires, c'est-à-dire, un point $\bar{x}$ pour lequel il existe y^i, $i \in I \bigcup E$ tel que

$$\begin{cases} \nabla f(\bar{x}) = \sum_{i \in I \bigcup E} y_i \nabla c_i(\bar{x}), & \\ c_i(\bar{x}) \geq 0, & \forall i \in I, \\ c_i(\bar{x}) = 0, & \forall i \in E, \\ y^i c_i(\bar{x}) = 0, & \forall i \in I, \\ y^i \geq 0, & \forall i \in I. \end{cases}$$

Dans le but de surmonter les difficultés due à la non linéarité de ce système, la méthode S.Q.P. utilise des approximations.
Ainsi, étant donné un point x_k, on définit

$$g_k(x_k + s, d) = [\nabla f(x_k) + Q_k s]^T d. \tag{6.4}$$

Comme dans le cas de l'optimisation sans contrainte, on suppose que les matrices Q_k sont symétriques définies positives. On cherche alors une direction s_k telle que

$$s_k \in S_k = \{s \ t.q. \ g_k(x_k + s; d) \geq 0, \ \forall d \in D(x_k + s; X_k)\}, \tag{6.5}$$

où

$$X_k = \left\{ x_k + s \ t.q. \ \begin{array}{l} c_i(x_k) + \nabla c_i(x_k)^T s \geq 0, \ i \in I, \\ c_i(x_k) + \nabla c_i(x_k)^T s = 0, \ i \in E \end{array} \right\}.$$

X_k est l'ensemble obtenu en linéarisant chaque contrainte $c_i(x)$, et l'ensemble $D(x_k + s; X_k)$ est défini comme en (4.2).
On remarque immédiatement que rechercher une direction satisfaisant à (6.5) revient à résoudre le problème quadratique suivant:

$$\begin{cases} \text{minimiser} & \nabla f(x_k)^T s + 1/2 s^T Q_k s \\ \text{sous contraintes} & c_i(x_k) + \nabla c_i(x_k)^T s \geq 0, \ \ \forall i \in I, \\ & c_i(x_k) + \nabla c_i(x_k)^T s = 0, \ \ \forall i \in E. \end{cases} \tag{6.6}$$

Les solutions primales et duales du problème (6.6) seront notées s_k et y_k respectivement. Une fois la direction s_k calculée, le nouvel itéré x_{k+1} sera obtenue en effectuant une recherche linéaire le long de s_k en utilisant comme fonction de mérite la fonction de pénalité exacte définie comme suit:

$$\begin{array}{rcl} f_k(x_k + s) & = & f(x_k) + \nabla f(x_k)^T s + 1/2 s^T Q_k s \\ & & + \sum_{i \in I} \sigma_k^i \max\{0, -c_i(x_k) - \nabla c_i(x_k)^T s\} \\ & & + \sum_{i \in E} \sigma_k^i |c_i(x_k) + \nabla c_i(x_k)^T s|, \end{array} \tag{6.7}$$

où

$$\sigma_i^k \geq y_k^i + \epsilon, \ i \in I \bigcup E, \ \epsilon > 0.$$

On obtient alors le théorème suivant.

Théorème 6.1

Supposons que les fonctions f et c_i, $i \in I \bigcup E$, sont continûment différentiables et que la fonction f est bornée inférieurement.
Soient les fonctions f_k et g_k définies comme en (6.4) et (6.7), où les matrices $\{Q_k\}$ sont symétriques et définies positives telles que, si λ_k est la plus petite valeur propre de Q_k, alors
$\liminf_{k\to\infty} \lambda_k > 0$.
Supposons que la suite $\{x_k\}$ est générée selon (5.2). Posons $\Delta_k = f_k(x_k + s_k) - f_k(x_k)$. Supposons de plus que

(a) $s_k \in S_k = \{s \text{ t.q. } g_k(x_k + s, d) \geq 0,\ \forall d \in D(x_k + s_k; X_k)\}$;

(b) La condition de qualification de contraintes d'Abadie, notée A.C.Q., est vérifiée pour tout point d'accumulation de la suite $\{x_k\}$, i.e., si

$$I(x) = \{i \in I \text{ t.q. } c_i(x) = 0,\ i \in I\}$$

et

$$D_L(x; X) = \{d \in \Re^n \text{ t.q. } \nabla c_i(x)^T d \geq 0,\ i \in I(x) \text{ et } \nabla c_i(x)^T d = 0,\ i \in E\},$$

alors $D_L(x; X) = D(x; X)$ pour tout point d'accumulation de la suite $\{x_k\}$;

(c) $\lim_{k\to\infty} \Delta_k = 0$.

Alors, tout point d'accumulation de la suite $\{x_k\}$ appartient à l'ensemble

$$S = \{x \in X \text{ t.q. } \nabla f(x)^T d \geq 0,\ \forall d \in D(x; X)\}.$$

<u>Preuve</u>

La preuve se base sur le lemme 6.1. Dans un premier temps, on montrera que, si $\bar{x}$ est un point d'accumulation de la suite $\{x_k\}$, alors $\bar{x} \in X$, i.e.,

$$\begin{aligned} c_i(\bar{x}) &\geq 0, \quad i \in I, \\ c_i(\bar{x}) &= 0, \quad i \in E. \end{aligned}$$

Ensuite, on vérifiera que les conditions (a), (b), (c) du lemme 6.1 sont satisfaites.
Notons que, pour la première partie, il suffit de montrer que

$$\liminf_{k\to\infty} c_i(x_k) \geq 0, \forall i \in I$$

et

$$\lim_{k\to\infty} c_i(x_k) = 0, \forall i \in E.$$

De (6.7), on a

$$f_k(x_k) = f(x_k) + \sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} + \sum_{i\in E} \sigma_k^i |c_i(x_k)|, \tag{6.8}$$

où la condition $\sigma_k^i > y_k^i$ assure que

$$\sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} + \sum_{i\in E} \sigma_k^i |c_i(x_k)| + \sum_{i\in I\bigcup E} y_k^i c_i(x_k) \geq 0. \tag{6.9}$$

D'autre part, comme s_k est solution de (6.6) avec y_k pour multiplicateur de Lagrange, on a l'inégalité suivante:

$$\begin{array}{rcl} -\sum_{i\in I\bigcup E} y_k^i c_i(x_k) & \geq & \nabla f(x_k)^T(s_k) + 1/2 s_k^T Q_k s_k \\ & & -\sum_{i\in I\bigcup E} y_k^i (c_i(x_k) + \nabla c_i(x_k)^T s_k), \end{array}$$

ou encore, par la condition de complémentarité du programme quadratique,

$$-\sum_{i\in I\bigcup E} y_k^i c_i(x_k) \geq \nabla f(x_k)^T(s_k) + 1/2 s_k^T Q_k s_k.$$

Enfin, par définition de la solution du problème (6.6), s_k est admissible et on a

$$\begin{array}{rcl} -\sum_{i\in I\bigcup E} y_k^i c_i(x_k) & \geq & \nabla f(x_k)^T(s_k) + 1/2 s_k^T Q_k s_k \\ & & +\sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k) - \nabla c_i(x_k)^T s_k\} \\ & & +\sum_{i\in E} \sigma_k^i |c_i(x_k) + \nabla c_i(x_k)^T s_k|. \end{array}$$

Par la définition de la fonction de pénalité exacte (6.7), le terme de droite de la dernière inégalité vaut $f_k(x_k + s_k) - f(x_k)$. On a donc

$$f_k(x_k + s_k) \leq f(x_k) - \sum_{i\in I\bigcup E} y_k^i c_i(x_k) \tag{6.10}$$

On obtient ainsi

$$\begin{array}{rcll} -\Delta_k & = & -f_k(x_k + s_k) + f_k(x_k) & \text{par hypothese,} \\ & = & -f_k(x_k + s_k) + f(x_k) + \sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} & \\ & & +\sum_{i\in E} \sigma_k^i |c_i(x_k)| & \text{par (6.8),} \\ & \geq & \sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} + \sum_{i\in E} \sigma_k^i |c_i(x_k)| + & \\ & & \sum_{i\in I\bigcup E} y_k^i c_i(x_k) & \text{par (6.10),} \\ & \geq & 0. & \text{par (6.9)} \end{array}$$

Comme $\lim_{k\to\infty} \Delta_k = 0$, on en déduit

$$\lim_{k\to\infty} \sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} + \sum_{i\in E} \sigma_k^i |c_i(x_k)| + \sum_{i\in I\bigcup E} y_k^i c_i(x_k) = 0. \tag{6.11}$$

Supposons, par l'absurde, qu'il existe $i \in I$ tel que $\liminf_{k\to\infty} c_i(x_k) < 0$. Alors il existe un ensemble d'indices K tel que

$$\lim_{k\in K} c_i(x_k) < 0.$$

Utilisant la condition $\sigma_k^i \geq y_k^i + \epsilon$, $\forall i, k$, on obtient

$$\lim_{k\in K} c_i(x_k)(y_k^i - \sigma_k^i) > 0.$$

Un raisonnement semblable peut être tenu pour $i \in E$. Si on considère ces conditions dans l'égalité (6.11), on obtient une somme de termes strictement positifs qui serait nulle, ce qui est impossible. D'où, on a

$$\liminf_{k\to\infty} c_i(x_k) \geq 0, \ \forall i \in I \text{ et } \lim_{k\to\infty} c_i(x_k) = 0 \ \forall i \in E.$$

Donc, tout point d'accumulation de la suite $\{x_k\}$ est contenu dans l'ensemble X.
Il reste ensuite à vérifier que les trois conditions du lemme 6.1 sont satisfaites. En ce qui concerne la condition (a), c'est évident. Notons que

$$g_k(x_k + s, d) = (\nabla f(x_k) + Q_k s)^T d.$$

Comme les fonctions f et c_i, $i \in I \bigcup E$ sont continûment différentiables, la condition (b) du lemme 6.1 découle immédiatement de l'hypothèse de qualification de contraintes A.C.Q.[3]
Comme s_k et y_k^i sont solutions du système (6.6), on a

$$\nabla f(x_k)^T(s_k) + s_k^T Q_k s_k = \sum_{i\in I\bigcup E} y_k^i \nabla c_i(x_k)^T s_k,$$

Et donc, en utilisant la définition de la fonction de pénalité exacte (6.7), on a

$$\begin{aligned}
\Delta_k &= f_k(x_k + s_k) - f_k(x_k) \\
&= \nabla f(x_k)^T s_k + 1/2 s_k^T Q_k s_k - \textstyle\sum_{i\in I\bigcup E} y_k^i [c_i(x_k) + \nabla c_i(x_k)^T s_k] \\
&\quad - \textstyle\sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} - \sum_{i\in E} \sigma_k^i |c_i(x_k)| \\
&= - \textstyle\sum_{i\in I\bigcup E} y_k^i c_i(x_k) - 1/2 s_k^T Q_k s_k - \sum_{i\in I} \sigma_k^i \max\{0, -c_i(x_k)\} \\
&\quad - \textstyle\sum_{i\in E} \sigma_i^k |c_i(x_k)|.
\end{aligned}$$

Ainsi de l'hypothèse (c) et du résultat (6.11), on en déduit que

$$\lim_{k\to\infty} \frac{1}{2} s_k^T Q_k s_k = 0.$$

Enfin, les matrices Q_k étant définies positives et $\liminf_{k\to\infty} \lambda_k > 0$, on obtient immédiatement que la suite $\{s_k\}$ converge vers 0.
Or,

$$g_k(x_k + s_k, d) - g_k(x_k, d) = (Q_k s_k)^T d.$$

D'où la condition (c) du lemme 6.1 est satisfaite et le théorème est démontré. $\square$

Chapitre 7

Approximations en optimisation convexe non différentiable

Dans ce chapitre, on s'intéresse à la résolution du problème d'optimisation non différentiable suivant:

$$(P)\begin{cases} \text{minimiser} & f(x) \\ \text{sous contraintes} & x \in X \end{cases}$$

où X est une partie convexe fermée non vide de $\Re^n$ et f est une fonction convexe non différentiable de $\Re^n$ dans $\Re$. Comme dans les chapitres précédents, on définit l'ensemble des solutions S comme suit:

$$S = \{x \in X \ t.q. \ g(x,d) \geq 0, \ \forall d \in D(x;X)\},$$

où $D(x;X) = \{d \ t.q. \ \exists \{x_k\} \subset X \ t.q. \ \{x_k\} \to x, \ et \ d = \lim_{k\to\infty} \frac{(x_k - x)}{\|x_k - x\|}\}$,

$g(x,d)$ représente la dérivée directionnelle de la fonction f au point x dans la direction d.

Rappelons à présent la définition et quelques propriétés du sous différentiel d'une fonction convexe f. Le sous différentiel de f en x est défini par

$$\partial f(x) \stackrel{def}{=} \{x^* \in \Re^n \ t.q. \ f(z) \geq f(x) + \langle x^*, z - x\rangle \ \forall z \in \Re^n\},$$

et le $\epsilon-$sous différentiel de f en x par

$$\partial_\epsilon f(x) \stackrel{def}{=} \{x^* \in \Re^n \ t.q. \ f(z) \geq f(x) + \langle x^*, z - x\rangle - \epsilon \ \forall z \in \Re^n\}.$$

Le sous différentiel $\partial f(x)$ satisfait les propriétés suivantes:

(1) $\partial f(x) = \{\xi \ t.q. \ \xi^T d \leq f'(x;d), \ \forall d \in D(x;X)\}$;
(2) $0 \in \partial f(x) \Leftrightarrow f'(x;d) \geq 0, \ \forall d \in D(x;X)$;
(3) $f'(x;d) = \max\{\xi^T d \ t.q. \ \xi \in \partial f(x)\}$.

Avant de considérer l'implication de l'emboîtement épigraphique pour des problèmes d'optimisation, on considère quelques résultats généraux concernant une méthode en optimisation non différentiable.

7.1 Résultats de base

Soit une fonction $f : \Re^n \to \Re \bigcup \{+\infty\}$ convexe, propre et semi-continue inférieurement. En vue de rechercher le minimum de f, on considère la suite $\{x_k\}$ définie de la façon suivante

$$x_{k+1} = x_k - \alpha_k \gamma_k, \ \ k \in N_0$$

où $\alpha_k > 0, \ \gamma_k \in \partial_{\epsilon_k} f(x_k), \ \epsilon_k \geq 0$.
Indiquons quelques résultats théoriques concernant cette suite.

Lemme 7.1

$\forall k, \ \forall y \in X$, on a

$$||x_{k+1} - y||^2 \leq ||x_k - y||^2 + \alpha_k^2 ||\gamma_k||^2 + 2\alpha_k [f(y) - f(x_k) + \epsilon_k].$$

Preuve

En utilisant la définition de la suite $\{x_k\}$ et le fait que $\gamma_k \in \partial_{\epsilon_k} f(x_k)$, on a

$$\begin{aligned} ||x_{k+1} - y||^2 &= ||x_{k+1} - x_k + x_k - y||^2 \\ &= ||x_{k+1} - x_k||^2 + 2\langle x_{k+1} - x_k, x_k - y\rangle + ||x_k - y||^2 \\ &= \alpha_k^2 ||\gamma_k||^2 + 2\langle \alpha_k \gamma_k, -x_k + y\rangle + ||x_k - y||^2 \\ &\leq \alpha_k^2 ||\gamma_k||^2 + 2\alpha_k [f(y) - f(x_k) + \epsilon_k] + ||x_k - y||^2 \end{aligned}$$

□

Proposition 7.1

Si $\sum_{k=1}^{\infty} \alpha_k = +\infty, \ \{\epsilon_k\} \to 0, \ \{\alpha_k ||\gamma_k||^2\} \to 0$,
alors $\liminf_{k\to\infty} f(x_k) = inf_{\{x \in \Re^n\}} f(x) \stackrel{not}{=} f^*$.

Preuve

Par l'absurde, supposons que $\liminf_{k\to\infty} f(x_k) > f^*$. Alors,

$$\exists\delta > 0,\ \exists y \in \Re^n,\ \lim_{k\to\infty} inf_{\{q\geq k\}} f(x_q) > f(y) + \delta,$$

ce qui peut aussi s'écrire

$$\exists\delta > 0,\ \exists y \in \Re^n,\ \exists k_0\ \forall k \geq k_0\ inf_{\{q\geq k\}} f(x_q) > f(y) + \delta.$$

D'où,

$$\exists\delta > 0,\ \exists y \in \Re^n,\ \exists k_0\ t.q.\ \forall k \geq k_0\quad f(x_k) > f(y) + \delta.$$

Or $\{\alpha_k\|\gamma_k\|^2 + \epsilon_k\} \to 0$. Donc,

$$\alpha_k\|\gamma_k\|^2 + \epsilon_k < \frac{\delta}{2} \text{ pour k assez grand.}$$

Sans perte de généralité, on peut supposer que k_0 est suffisamment grand. Dès lors, par le lemme 7.1,

$$\begin{aligned} \|x_{k+1} - y\|^2 &\leq \|x_k - y\|^2 + \alpha_k[2\alpha_k\|\gamma_k\|^2 + 2\{f(y) - f(x_k) + \epsilon_k\}] \\ &\leq \|x_k - y\|^2 + \alpha_k[\delta - 2\delta] \\ &= \|x_k - y\|^2 - \delta\alpha_k. \end{aligned}$$

Donc, on a pour tout $k > k_0$,

$$0 \leq \|x_k - y\|^2 \leq \|x_{k_0} - y\|^2 - \delta\sum_{i=k_0}^{k-1}\alpha_i.$$

Quand on fait tendre k vers l'infini, on obtient que la série $\sum_{i=k_0}^{+\infty}\alpha_i$ est convergente, ce qui contredit l'hypothèse. D'où la thèse. □

Proposition 7.2

Soit $\bar{x}$ une valeur d'adhérence de la suite $\{x_k\}$ vérifiant

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 + \delta_k$$

où $\{\delta_k\}$ est une suite positive ou nulle telle que $\sum_{k=1}^{\infty}\delta_k < +\infty$,
alors toute la suite $\{x_k\}$ converge vers $\bar{x}$.

Preuve

Soit $\delta > 0$. $\bar{x}$ étant une valeur d'adhérence de la suite $\{x_k\}$, il existe un ensemble d'indices K tel que $\{x_k\}_K \to \bar{x}$. Il existe donc un indice $k_1 \in K$ tel que

$$\|x_{k_1} - \bar{x}\| \leq \frac{\delta}{2} \text{ et } \sum_{k=k_1}^{\infty} \delta_i \leq \frac{\delta}{2}.$$

Par hypothèse, pour tout k_2 plus grand que k_1, on a

$$\|x_{k_2+1} - \bar{x}\|^2 \leq \|x_{k_1} - \bar{x}\|^2 + \sum_{k=k_1}^{k=k_2} \delta_k \leq \delta,$$

d'où la thèse. □

7.2 Itération proximale

Soit $\alpha_k > 0$ et $x_k \in \Re^n$. On perturbe la fonction f de la façon suivante:

$$\bar{f}(y) = f(y) + \frac{1}{2\alpha_k}\|y - x_k\|^2$$

On a

$$\partial \bar{f}(y) = \partial f(y) + \frac{y - x_k}{\alpha_k}.$$

On suppose la suite $\{\alpha_k\}$ fixée à priori. *L'itération proximale* consiste à définir la suitę $\{x_k\}$ de la façon suivante

$$x_{k+1} = p_f(x_k),$$

où

$$p_f(x_k) = \text{argmin}\{\bar{f}(y) \ t.q.\ y \in \Re^n\}.$$

On a

$$\begin{array}{lcl} x_{k+1} = p_f(x_k) & \Leftrightarrow & 0 \in \partial \bar{f}(x_{k+1}) \\ x_{k+1} = p_f(x_k) & \Leftrightarrow & \frac{x_k - x_{k+1}}{\alpha_k} \in \partial f(x_{k+1}). \end{array}$$

L'itération proximale peut donc se mettre sous la forme

$$\begin{cases} x_{k+1} = x_k - \alpha_k \gamma_k \\ \text{avec } \gamma_k \in \partial f(x_{k+1}). \end{cases}$$

La proposition suivante montre que l'itération proximale rentre dans le shéma général.

Proposition 7.3

$$x_{k+1} = p_f(x_k) \Leftrightarrow \begin{cases} x_{k+1} = x_k - \alpha_k \gamma_k \\ \text{avec } \gamma_k \in \partial_{\epsilon_k} f(x_k) \\ \text{et } \epsilon_k = f(x_k) - f(x_{k+1}) - \alpha_k \|\gamma_k\|^2 \geq 0 \end{cases}$$

Preuve

Supposons d'abord que $x_{k+1} = p_k(x_k)$. Alors $x_{k+1} = x_k - \alpha_k \gamma_k$ avec $\gamma_k \in \partial f(x_{k+1})$.
Par définition du sous différentiel, on a $f(x_k) \geq f(x_{k+1}) + \langle \gamma_k, x_k - x_{k+1} \rangle$, qui peut encore s'écrire $f(x_k) - f(x_{k+1}) - \alpha_k \|\gamma_k\|^2 \geq 0$.
Il reste à voir que $\gamma_k \in \partial_{\epsilon_k} f(x_k)$ avec $\epsilon_k = f(x_k) - f(x_{k+1}) - \alpha_k \|\gamma_k\|^2$.
Pour tout y, on a

$$f(y) \geq f(x_{k+1}) + \langle \gamma_k, y - x_{k+1} \rangle.$$

Dès lors

$$f(y) \geq f(x_{k+1}) + f(x_k) - f(x_k) + \langle \gamma_k, y - x_k \rangle + \langle \gamma_k, x_k - x_{k+1} \rangle$$

et donc

$$f(y) \geq f(x_k) - \epsilon_k + \langle \gamma_k, y - x_k \rangle.$$

D'où la thèse.

Pour la réciproque, il suffit de montrer que $\gamma_k \in \partial f(x_{k+1})$.
Comme $\gamma_k \in \partial_{\epsilon_k} f(x_k)$ et $\epsilon_k = f(x_k) - f(x_{k+1}) - \alpha_k \|\gamma_k\|^2$, on a pour tout y, l'inégalité suivante

$$f(y) \geq f(x_k) - [f(x_k) - f(x_{k+1}) - \alpha_k \|\gamma_k\|^2] + \langle \gamma_k, y - x_k \rangle,$$

qui peut encore s'écrire

$$f(y) \geq f(x_{k+1}) + \alpha_k \|\gamma_k\|^2 + \langle \gamma_k, y - x_k \rangle,$$

$$f(y) \geq f(x_{k+1}) + \langle \gamma_k, x_k - x_{k+1} \rangle + \langle \gamma_k, y - x_k \rangle,$$

$$f(y) \geq f(x_{k+1}) + \langle \gamma_k, y - x_{k+1} \rangle.$$

D'où la thèse. □

Proposition 7.4 Convergence de l'algorithme proximal

Soit $\{x_k\}$ la suite engendrée par l'algorithme proximal.
Si $\sum_{k=1}^{+\infty} \alpha_k = +\infty$, alors la suite $\{f(x_k)\}$ est décroissante et converge vers $inf_{\{x \in \Re^n\}} f(x)$.

Preuve

Par la proposition précédente, on a

$$f(x_k) - f(x_{k+1}) = \epsilon_k + \alpha_k \|\gamma_k\|^2 \geq 0.$$

Par conséquent, la suite $\{f(x_k)\}$ est décroissante. Si $\{f(x_k)\}$ converge vers $-\infty$, la preuve est terminée. Si $\{f(x_k)\}$ converge f^*, alors

$$\epsilon_k + \alpha_k \|\gamma_k\|^2 = f(x_k) - f(x_{k+1}) \to f^* - f^* = 0.$$

D'où $\epsilon_k \to 0$ et $\alpha_k \|\gamma_k\|^2 \to 0$. Il suffit alors d'appliquer la proposition 7.1. □

Considérons ensuite une situation un peu plus générale.

Proposition 7.5

Soit la suite $\{x_k\}$ définie par

$$\begin{cases} x_{k+1} = x_k - \alpha_k \gamma_k \\ \text{avec } \alpha_k > 0, \gamma_k \in \partial_{\epsilon_k} f(x_k), \ \epsilon_k \geq 0. \end{cases}$$

et soit $m > 0$.

(i) Si $\sum_{k=1}^{+\infty} \alpha_k = +\infty$, $\{\epsilon_k\} \to 0$, et $f(x_{k+1}) \leq f(x_k) - m\alpha_k \|\gamma_k\|^2$, alors $\{f(x_k)\} \searrow inf_{\{x \in \Re^n\}} f(x)$.

(ii) Si en plus, la suite $\{\alpha_k\}$ est bornée et si $\sum_{k=1}^{+\infty} \epsilon_k < +\infty$, alors, s'il existe un minimum, $\{x_k\} \to x^*$ où x^* est un minimum de la fonction f.

Remarquons que le cas où m vaut 1 revient à la méthode proximale.

<u>**Preuve**</u>

(i) D'abord la suite $\{f(x_k)\}$ est une suite décroissante. Si elle converge vers $-\infty$, la preuve est achevée. Si elle converge de façon décroissante vers f^*, alors $\{\alpha_k \|\gamma_k\|^2\} \to 0$ et il suffit d'appliquer la proposition 7.1.

(ii) (a) la suite $\{x_k\}$ est bornée.

Supposons qu'il existe un minimum x^*. Par le lemme 7.1, on a

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &\leq \|x_k - x^*\|^2 + \alpha_k^2 \|\gamma_k\|^2 + 2\alpha_k [f(x_{k+1}) - f(x_k) + \epsilon_k] \\ &\leq \|x_k - x^*\|^2 + \alpha_k^2 \|\gamma_k\|^2 + 2\alpha_k [m\alpha_k \|\gamma_k\|^2 + \epsilon_k] \\ &\leq \|x_k - x^*\|^2 + (1+2m)\alpha_k^2 \|\gamma_k\|^2 + 2\alpha_k \epsilon_k. \end{aligned}$$

Donc, on a

$$\|x_k - x^*\|^2 \leq \|x_1 - x^*\|^2 + (1+2m) \sum_{i=1}^{+\infty} \alpha_i^2 \|\gamma_i\|^2 + 2 \sum_{i=1}^{+\infty} \alpha_i \epsilon_i.$$

La série $\sum_{i=1}^{+\infty} \alpha_i^2 \|\gamma_i\|^2$ est convergente car la suite $\{\alpha_i\}$ est bornée et la série $\sum_{i=1}^{+\infty} \alpha_i \|\gamma_i\|^2$ est convergente. En effet,

$$\begin{aligned}
\sum_{i=1}^{+\infty} \alpha_i \|\gamma_i\|^2 &\leq 1/m \sum_{i=1}^{+\infty} f(x_i) - f(x_{i+1}) \\
&= 1/m \lim_{k\to\infty} [f(x_1) - f(x_k)] \\
&= \frac{f(x_1) - f^*}{m} \in \Re.
\end{aligned}$$

La série $\sum_{i=1}^{+\infty} \alpha_i \epsilon_i$ est convergente, car la suite $\{\alpha_i\}$ est bornée et la série $\sum_{i=1}^{+\infty} \epsilon_i$ est convergente. Par conséquent, la suite $\{x_k\}$ est bornée.

(b) Soit $\bar{x}$ une valeur d'adhérence de la suite $\{x_k\}$.
Alors $\{x_{q_k}\} \to \bar{x}$, donc $\{f(x_{q_k})\} \searrow f(\bar{x})$ et

$$f(\bar{x}) \leq \liminf_{k\to\infty} f(x_{q_k}) = \lim_{k\to\infty} f(x_{q_k}) = \lim_{k\to\infty} f(x_k) = inf_{\{x\in\Re^n\}} f(x).$$

D'où $\bar{x}$ est un minimum. De plus, par le lemme 7.1, on a

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 + (1+2m)\alpha_k^2 \|\gamma_k\|^2 + 2\alpha_k \epsilon_k.$$

Comme la série $\sum (1+2m)\alpha_k^2 \|\gamma_k\|^2 + 2\alpha_k \epsilon_k$ est convergente, on a par la proposition 7.2 que toute la suite $\{x_k\}$ converge vers $\bar{x}$. □

Pour l'algorithme proximal, on peut améliorer la partie (ii) de cette proposition.

Proposition 7.6

Soit $\{x_k\}$ la suite engendrée par l'algorithme proximal.
Si la série $\sum_{k=1}^{\infty} \alpha_k$ est divergente, alors $\{f(x_k)\} \searrow inf_{\{x\in\Re^n\}} f(x)$ et la suite $\{x_k\}$ converge vers un minimum x^*, si il existe.

Preuve

Il faut montrer seulement la deuxième conclusion, puisque la première a déjà été vue. Dans la preuve précédente, on remplace ϵ_k par sa valeur et on obtient

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2.$$

D'où, la suite $\{x_k\}$ est bornée. On applique la proposition 7.2 avec $\delta_k = 0$ et on a que toute la suite $\{x_k\}$ converge vers x^*. □

7.3 Application de l'emboîtement épigraphique à la méthode proximale

En vue de résoudre le problème (P) décrit en début de chapitre, considérons une généralisation de la méthode proximale sans recherche linéaire

$$x_{k+1} = x_k + s_k$$

où $s_k \in \operatorname{argmin}\{f_k(x_k + s) + \rho(s) \ t.q. \ x_k + s \in X\}$, f_k est une approximation convexe de f et $\rho(s) = \frac{1}{2\sigma}\|s\|^2$ avec $\sigma > 0$. Comme précédemment, on notera $g_k(x_k + s, d)$ la dérivée directionnelle de la fonction $f_k(x_k + s)$ au point $x_k + s$ dans la direction d.

Dans le théorème suivant, on donne des conditions sur la suite $\{f_k\}$ pour que tout point d'accumulation de la suite $\{x_k\}$ soit optimal.

Théorème 7.1

Considérons une suite $\{x_k\}$ telle que $x_{k+1} = x_k + s_k$ et supposons qu'elle admette une sous-suite $\{x_k\}_{k\in K} \to \bar{x}$.
Si les conditions suivantes sont satisfaites

(a) $s_k \in S_k = \{s \ t.q. \ g_k(x_k + s, d) + \rho'(s; d) \geq 0, \ \forall d \in D(x_k + s; X), \ et \ x_k + s \in X\}$;

(b) pour tout $d \in D(\bar{x}, X)$, *il existe une suite* $\{d_k\}_K$ *telle que* $d_k \in D(x_k + s_k; X)$,

$$\{d_k\}_K \to d \text{ et } \limsup_{k\in K} g_k(x_k + s_k, d_k) \leq g(\bar{x}; d);$$

(c) si $\Delta_k = f_k(x_{k+1}) - f_k(x_k)$, *alors la suite* $\{\Delta_k\}_K$ *converge vers 0;*

(d) il existe un scalaire $\epsilon > 0$ *tel que* $\Delta_k \leq -\epsilon\rho(s_k)$.

Alors $\bar{x}$ appartient à l'ensemble S.

Preuve

Comme X est un ensemble fermé et comme la suite $\{x_k\}$ est incluse dans l'ensemble X, on a que $\bar{x} \in X$. Soit $d \in D(\bar{x}; X)$ et considérons la suite $\{d_k\}$ donnée par l'hypothèse (b). On a alors

$$0 \leq g_k(x_k + s_k, d_k) + \rho'(s_k; d_k). \tag{7.1}$$

Par les conditions (c) et (d), la suite $\{\rho(s_k)\}_K$ converge vers 0 et donc la suite $\{s_k\}_K$ converge vers 0. Or

$$\rho'(s_k; d_k) = \frac{(s_k)^T d_k}{\sigma}.$$

D'où, on a

$$\{\rho'(s_k; d_k)\}_K \to 0.$$

D'autre part la condition (b) combinée avec (7.1) implique que

$$0 \leq \limsup_{k\to\infty} g_k(x_k + \alpha_k s_k, d_k) \leq g(\bar{x}, d).$$

Donc, $\bar{x} \in S$. □

Remarquons qu'on peut également vérifier ce théorème quand une recherche linéaire est effectuée pour un problème (P) sans containte.

Vérifions à présent, que les conditions du théorème 7.1 sont satisfaites quand $f_k = f$ pour tout k c'est-à-dire lorsque la fonction objectif f n'est pas remplacée par une approximation f_k. Dans ce cas on retrouve *la méthode originale du point proximal*. Notons qu'on se restreint au cas sans contrainte.

Théorème 7.2

Considérons la suite $\{x_k\}$ engendrée par la méthode du point proximal sans recherche linéaire. Supposons que la suite $\{x_k\}$ est bornée et notons K l'ensemble des indices tel que $\{x_k\}_K \to \bar{x}$. Alors

$$\bar{x} \in S \stackrel{def}{=} \{x \in \Re^n \ t.q. \ f'(x; d) \geq 0, \ \forall d \in \Re^n\}.$$

Preuve

Il suffit de vérifier les quatre hypothèses du théorème 7.1.
La condition (a) s'écrit

$$s_k \in S_k = \{s \ t.q. \ f'(x_k + s; d) + \rho'(s; d) \geq 0, \ \forall d \in \Re^n\}.$$

Elle est trivialement satisfaite car, par définition de la direction s_k, on a

$$s_k \in \text{argmin}\{f(x_k + s) + \rho(s)\},$$

Pour satisfaire la condition (b), il faut vérifier que

$$\forall d \in \Re^n, \ \exists \{d_k\}_K \subseteq \Re^n \ t.q. \ \{d_k\}_K \to d, \ et \ \limsup_{k\in K} f'(x_k + s_k; d_k) \leq f'(\bar{x}; d).$$

Soit $d \in \Re^n$ et soit la suite $\{d_k\}$ où $d_k = d$ pour tout k, qui converge bien vers d. On sait que la fonction f est continue. Par conséquent la fonction f_t définie pour tout $t > 0$, par

$$f_t(x) = \frac{f(x + td) - f(x)}{t}$$

est aussi continue. Comme $f'(.;d) \stackrel{def}{=} inf_{t>0} f_t(.)$, la fonction $f'(.;d)$ est semi-continue supérieurement, et donc $\limsup_{k \in K} f'(x_k; d_k) \leq f'(\bar{x}; d)$.

Ensuite par la proposition 7.4, on a que la suite $\{f(x_k)\}$ converge vers f^* puisque ici $\alpha_k = 1$ pour tout k. Par conséquent, on a

$$\Delta_k = f(x_{k+1}) - f(x_k) \to f^* - f^* = 0.$$

La condition (c) du théorème 7.1 est donc bien satisfaite.

Enfin, par la proposition 7.3, on a que

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\leq -\|\gamma_k\|^2 \\ &\leq -\|x_k - x_{k+1}\|^2. \end{aligned}$$

Par conséquent, il suffit de prendre pour ϵ la valeur $2\sigma > 0$ pour obtenir la condition (d)

$$\Delta_k \leq -\epsilon\rho(s_k).$$

La thèse découle alors du théorème 7.1. □

On peut également montrer que toute la suite $\{x_k\}$ engendrée par la méthode originale du point proximal converge vers $\bar{x}$.

Théorème 7.3

Soit la suite $\{x_k\}$ engendrée par l'algorithme du point proximal.
Alors toute la suite $\{x_k\}$ converge vers $\bar{x}$ où $\bar{x} \in S$.

Preuve

Par la proposition 7.2, la thèse sera obtenue si on montre que

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2.$$

En effet, il suffit alors de prendre pour suite $\{\delta_k\}$, la suite nulle qui engendre bien une série convergente. Il reste donc à montrer l'inégalité précédente.
Par définition de x_{k+1}, on a successivement

$$x_{k+1} = \operatorname{argmin}\{f(x) + \frac{1}{2\sigma}\|x - x_k\|^2\},$$

$$i.e., \ 0 \in \partial f(x_{k+1}) + 1/\sigma(x_{k+1} - x_k),$$

$$\frac{x_k - x_{k+1}}{\sigma} \in \partial f(x_{k+1}).$$

De plus par le théorème 7.2, on sait que $\bar{x} \in S$, i.e.,

$$0 \in \partial f(\bar{x}).$$

Comme ∂f est monotone, on a

$$\langle \frac{x_k - x_{k+1}}{\sigma}, x_{k+1} - \bar{x} \rangle \geq 0,$$

et donc

$$\langle x_k - x_{k+1}, x_{k+1} - \bar{x} \rangle \geq 0.$$

D'autre part comme la propriété suivante

$$\langle u, v \rangle = \frac{1}{2}[\|u+v\|^2 - \|u\|^2 - \|v\|^2]$$

est vraie pour tout $u, v \in \Re^n$, on a

$$\frac{1}{2}[\|x_k - \bar{x}\|^2 - \|x_k - x_{k+1}\|^2 - \|x_{k+1} - \bar{x}\|^2] \geq 0,$$

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \|x_k - x_{k+1}\|^2.$$

On en déduit

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2$$

et donc la thèse. □

Le théorème suivant exprime des conditions plus faibles sous lesquelles la suite $\{x_k\}$ converge vers $\bar{x} \in S$.

Théorème 7.4

Soit la suite $\{x_k\}$ générée par l'algorithme proximal avec recherche linéaire

$$x_{k+1} = x_k + \alpha_k s_k,$$

où $\{\alpha_k\}$ est une suite de scalaires positifs. Si la série $\sum_{k=1}^{\infty} \alpha_k = +\infty$ et si
$f(x_{k+1}) \leq f(x_k) - \alpha_k \|\frac{x_k - x_{k+1}}{\alpha_k}\|^2$,
alors toute la suite $\{x_k\}$ converge vers $\bar{x} \in S$.

Preuve

Montrons que les conditions du théorème 7.1 sont vérifiées. Les conditions (a) et (b) du théorème 7.1 sont satisfaites (cfr la preuve du théorème 7.2).

En ce qui concerne la condition (c), elle est satisfaite puisque par la proposition 7.4, la suite $\{f(x_k)\}$ converge vers f^*. Donc

$$\{\Delta_k\} \to f^* - f^* = 0.$$

Quant à la condition (d), elle vérifiée en vertu de la propriété

$$f(x_{k+1}) \leq f(x_k) - \alpha_k \| \frac{x_k - x_{k+1}}{\alpha_k} \|^2.$$

Il suffit de prendre pour ϵ la valeur $\frac{1}{2\alpha_k \sigma}$ quel que soit k. D'où la thèse. □

7.4 Discussion

L'emboîtement épigraphique permet de prouver le convergence de nombreux algorithmes aussi bien en optimisation différentiable que non différentiable. La condition imposée sur les données du problème se trouve résumée dans la proposition 4.1 et peut s'exprimer

$$inf_{\{\{d_k\} \to d\}} \limsup_{k \in K} g_k(x_k, d_k) \leq g(\bar{x}, d) \text{ quand} \{x_k\}_K \to \bar{x}.$$

Une question naturelle est alors de savoir quand cette condition est satisfaite. Pour y répondre, introduisons la notion de $\partial-$compatibilité sous-différentielle.

Définition de $\partial-$compatibilité

On dira qu'*une suite de fonctions $\{f_k\}$ et une fonction f sont $\partial-$compatibles, en respectant une suite de points $\{x_k\}$ si*

$$\text{Limsup}_{k \in K} \partial f_k(x_k) \subseteq \partial f(\bar{x}) \text{ quand } \{x_k\}_K \to \bar{x}.$$

Cette propriété sera notée $\{f_k\} \xrightarrow{\partial} f \text{ (wrt} \{x_k\})$.
De plus, si pour tout $x \in X \subseteq \Re^n$ et pour toute suite $\{x_k\} \subset X$ telle que $\lim_{k \to \infty} x_k = x$, on a $\{f_k\} \xrightarrow{\partial} f \text{ (wrt } \{x_k\})$, on dira alors que $\{f_k\} \xrightarrow{\partial} f(\text{wrt } X)$.

Comme $f'(x; d) = \max\{\xi^T d \ t.q. \ \xi \in \partial f(x)\}$, on a la proposition suivante.

Proposition 7.7

Si $f_k \xrightarrow{\partial} f(wrt\{x_k\})$, alors

$$inf_{\{\{d_k\} \to d\}} \limsup_{k \in K} g_k(x_k, d_k) \leq g(\bar{x}, d) \text{ quand} \{x_k\}_K \to \bar{x},$$

où g_k et g représentent les dérivées directionnelles de f_k et de f respectivement. □

Dans le contexte de la convergence d'algorithmes, la $\partial-$compatiblité, une propriété relativement facile à vérifier, offre une information intéressante pour la construction des fonctions approximantes.

Dans le théorème suivant, on établit le lien entre la notion de $\partial-$compatibilité et celle d'épi-convergence.

Théorème 7.5

Soit $\{f_k\}$ une suite de fonctions convexes telle que $\{f_k\} \to_{epi} f$ sur X où f est une fonction convexe et X une partie de $\Re^n$.
Alors $\{f_k\} \xrightarrow{\partial} f$ (wrt X).

Preuve

Soit $x \in X$ et soit $\{x_k\} \subseteq X$ une suite vérifiant $\{x_k\} \to x$, montrons que

$$\mathrm{Limsup}_{k\to\infty} \partial f_k(x_k) \subseteq \partial f(x).$$

Si $\mathrm{Limsup}_{k\to\infty} \partial f_k(x_k) = \emptyset$, le résultat est trivial.
Supposons donc qu'il existe $\mu \in \mathrm{Limsup}_{k\to\infty} \partial f_k(x_k)$ et montrons que $\mu \in \partial f(x)$ i.e. que

$$\forall y \in X,\ \ f(y) \geq f(x) + \langle \mu, (y - x) \rangle.$$

Prenons $y \in X$ arbitraire.
Par la définition de μ, il existe une suite $\{\mu_k\}_{k\in K} \to \mu$ telle que $\mu_k \in \partial f_k(x_k)$ pour tout $k \in K$.
Comme $\{f_k\} \to_{epi} f$ sur X, il existe une suite $\{y_k\} \subseteq X$ convergeant vers y et telle que

$$\lim_{k\to\infty} f_k(y_k) = f(y) \text{ et } \liminf_{k\to\infty} f_k(x_k) \geq f(x).$$

Dès lors, comme $\mu_k \in \partial f_k(x_k)$, on a, pour tout k,

$$f_k(y_k) \geq f_k(x_k) + \langle \mu_k, y_k - x_k \rangle$$

et donc en passant à la limite inférieure

$$\liminf_{k\in K} f_k(y_k) \geq \liminf_{k\in K} [f_k(x_k) + \langle \mu_k (y_k - x_k) \rangle].$$

On en déduit alors que

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

et donc la thèse. □

Ainsi la notion de $\partial-$compatibilité est une condition suffisante pour l'emboîtement épigraphique et nécessaire à l'épi-convergence.

Conclusion

Dans la première partie, on a vu que la convergence des méthodes barrières et de pénalisations extérieures pouvait être établie grâce à la notion d'épi-convergence. Si la limite des épigraphes des fonctions approximantes f_k est l'épigraphe de la fonction objectif f, alors la suite $\{x_k\}$ des minima des $\{f_k\}$ converge vers la solution du problème.

Dans la seconde partie, on a montré que l'emboîtement épigraphique des dérivées directionnelles des fonctions approximantes f_k permettait de prouver la convergence des méthodes de type Newton, des méthodes de programmation quadratique successive ainsi que de la méthode du point proximal en optimisation convexe non différentiable.

Finalement, on a introduit le concept de $\partial-$compatibilité, qui est une propriété facilement vérifiable et qui implique l'emboîtement épigraphique. De plus, l'épi-convergence entraîne la $\partial-$ compatibilité.

Bibliographie

[1] Hedy Attouch et Michel Théra, "Convergences en analyse multivoque et variationnelle", pp. 23-40

[2] Hedy Attouch et Roger J.-B. Wets, "Approximation and convergence in nonlinear optimization", Nonlinear programming, 1981, pp. 367-394

[3] Bazaraa, M. S., et Shetty, C. M.,"Nonlinear Programming: Theory and Algorithms", John Wiley et Sons, New York, 1976

[4] Rafael Correa et Claude Lemaréchal, "Convergence of some algorithms for convex minimisation", Mathematical Programming 62, 1993, pp. 261-275

[5] Julia L.Higle et Suvrajet Sen, " On the convergence of algorithms with implication for stochastic and nondifferentiable optimization", mathematics of operation research, Vol 17, n 1, févirer 1992, pp.112-131

[6] J.L. Higle et S. Sen, "Epigraphical Nesting: a unifying Theory for the Convergence of Algorithms", journal of optimization theory and application; vol 84, n 2, fevrier 1995, pp. 339-360

[7] Peter Kall, "Approximation to optimization problems: an elementary revieuw", mathematics of operations research; vol 11,n 1; février 1986, pp. 9-18