University of Namur

MASTER EN SCIENCES MATHÉMATIQUES

Résolution d'inéquations variationnelles par des méthodes barrières

Verstraeten, Virginie

Award date:
1999

FACULTES UNIVERSITAIRES NOTRE DAME DE LA PAIX
NAMUR
FACULTES DES SCIENCES

Résolution de problèmes d'inéquations variationnelles par des méthodes barrières

Mémoire présenté pour l'obtention du grade
de Licencié en Sciences
Mathématiques
par

Promoteur : Jean-Jacques Strodiot

Verstraeten Virginie

Année académique 1998-1999

Remerciements

Je remercie spécialement Monsieur le Professeur Jean-Jacques Strodiot pour l'aide précieuse qu'il m'a apportée tout au long de l'élaboration de ce mémoire.

Je tiens également à remercier mes parents, Frédéric et mes amis qui m'ont soutenue durant mes études, sans oublier Jean-Yves qui m'a souvent aidé à résoudre les problèmes rencontrés lors de la rédaction de ce mémoire.

Résumé

Dans ce mémoire, nous définissons d'abord de façon générale le problème d'inéquations variationnelles et nous démontrons sous quelles hypothèses il existe une ou plusieurs solutions.

Ensuite, nous précisons notre cadre de travail et nous considérons un premier problème d'inéquations variationnelles, noté $\mathrm{VI}(F, \Omega)$. Afin de résoudre ce dernier, nous définissons une famille de sous-problèmes barrières et nous présentons un algorithme convergeant vers l'unique solution de chacun de ces sous-problèmes. En ajoutant une hypothèse supplémentaire, nous remarquonsque l'on peut modifier cet algorithme pour en obtenir un second dont la convergence est quadratique. Après cela, nous élaborons un algorithme convergeant vers une solution du problème $\mathrm{VI}(F, \Omega)$.

Dans une seconde partie, nous étudions le problème d'inéquations variationnelles précédent dans un cadre plus général en introduisant la notion de self-concordance. De manière analogue à ce qui a été fait dans la première partie, nous généralisons les deux algorithmes convergeant vers l'unique solution de chaque sous-problème barrière, ainsi que l'algorithme convergeant vers une solution de l'inéquation variationnelle étudiée.

Abstract

First, we define the variational inequality problem we consider in this work and we determine which assumptions are required to prove the existence and the uniqueness of solutions for this problem.

Then, we specify our framework and we study a first variational inequality, denoted by $\mathrm{VI}(F, \Omega)$. To solve this problem, we define a family of barrier subproblems and we present an algorithm converging to the unique solution of each subproblem. By adding an additional assumption, we remark that we can establish an improved algorithm whose convergence is quadratic. After that, we present an algorithm converging to a solution of the problem $\mathrm{VI}(F, \Omega)$.

In a second part, we study the previous variational inequality problem in a more general framework by introducing the concept of self-concordance. Similarly to what was done in the first part, we generalize the two algorithms converging to the unique solution of each barrier subproblem and we also generalize the algorithm converging to a solution of the studied variational inequality problem.

Table des matières

Introduction

Dans ce mémoire, nous considérons les problèmes d'inéquations variationnelles VI(F, Ω) du type

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \forall x \in \Omega \quad F(x^*)^t(x - x^*) \geq 0,$$

où $\Omega \subseteq I\!R^n$ est un ensemble fermé, convexe, non vide et $F\colon \Omega \to I\!R^n$ est une fonction monotone.

Dans un premier temps, nous établirons quelques résultats d'existence et d'unicité. Nous déterminerons ainsi sous quelles hypothèses le problème VI(F, Ω) admet au moins une solution.

Ensuite, à l'instar de ce qui a été fait dans l'article de J.H. Wu [4], nous considérerons des ensembles Ω du type

$$\Omega = \{x | Ax = b, \ x \geq 0\} \subseteq I\!R^n$$

où A est une matrice de $I\!R^{m \times n}$ $(m \leq n)$ et b est un vecteur de $I\!R^m$. La résolution de ce problème d'inéquations variationnelles sera le principal objet du deuxième chapitre.
Pour résoudre ce problème, nous emploierons la méthode barrière et nous définirons sur l'intérieur de l'ensemble Ω une famille de fonctions $\{\bar{F}(x) = F(x) - X^{-1}w\}_{w>0}$. Nous présenterons une fonction de mérite pour chaque sous-problème barrière VI$(\bar{F}, \Omega_{int})$

$$\text{Trouver } x(w) \in \Omega_{int} \quad \text{tel que} \quad \forall x \in \Omega_{int} \quad \bar{F}(x(w))^t(x - x(w)) \geq 0,$$

ainsi qu'une fonction ϕ qui nous permettra de vérifier si un vecteur $x \in I\!R^n$ est proche de la solution du problème VI$(\bar{F}, \Omega_{int})$. Nous allons élaborer un premier algorithme (l'algorithme 2.1) pour résoudre le problème VI$(\bar{F}, \Omega_{int})$. Sous des hypothèses appropriées, nous montrerons que cet algorithme est convergent. Ensuite, nous ajouterons une hypothèse supplémentaire et grâce à celle-ci, nous serons en mesure d'élaborer une nouvelle version de l'algorithme 2.1. L'avantage de ce nouvel algorithme par rapport au précédent est qu'il converge quadratiquement vers la solution de VI($\bar{F}, \Omega_{int}$). Finalement, nous élaborerons un algorithme pour résoudre le problème VI(F, Ω) en se basant sur l'idée suivante : à l'itération k, si l'itéré x^k obtenu par la méthode de Newton est suffisamment proche de la solution de VI($\bar{F}, \Omega_{int}$) avec $w > 0$, alors nous diminuons la valeur du paramètre w et nous passons à l'itération suivante. Sinon, nous appliquons l'algorithme 2.1 pour "centrer" l'itéré. Pour clôturer ce chapitre, nous démontrerons la convergence de cet algorithme.

Dans le Chapitre 3, nous introduirons la notion de self-concordance. La plupart des résultats qui y seront présentés sont extraits du livre de Nesterov Y. et Nemirovskii A. [2]. Seuls les résultats les plus utiles dans ce mémoire y seront démontrés.

Dans le Chapitre 4, nous allons étendre l'ensemble Ω du Chapitre 2 de manière analogue à ce qui a été fait dans l'article de F.Sharifi-Mokhatarian et J.L. Goffin [1] : nous remplacerons les contraintes de non-négativité par des contraintes convexes. L'ensemble Ω sera donc de la forme

$$\Omega = \{x \in I\!R^n |\ Ax = b,\ f_i(x) \leq 0\ i = 1, ..., m\},$$

où $f_i : I\!R^n \to I\!R$ sont des fonctions convexes de classe C^3 pour tout i appartenant à l'ensemble $\{1, ..., m\}$, A est une matrice de $I\!R^{p\times n}$ $(p < n)$ et b est un vecteur de $I\!R^p$.
La démarche suivie lors de l'élaboration de l'algorithme pour résoudre ce problème d'inéquations variationnelles est semblable à celle présentée au Chapitre 2. En effet, nous définirons d'abord une famille de fonctions en ajoutant à la fonction objectif la dérivée d'une fonction barrière self-concordante multipliée par un facteur μ. Ensuite, nous élaborerons deux algorithmes pour résoudre chaque problème d'inéquations variationnelles "barrière". En suivant la même idée que précédemment, nous présenterons un algorithme pour résoudre le problème VI(S, Ω) où S est un opérateur monotone. Finalement, nous montrerons que cet algorithme est convergent.

Chapitre 1

Notions préliminaires

1.1 Définition du problème d'inéquations variationelles

Soient $\Omega \subseteq I\!R^n$ un ensemble fermé, convexe, non vide et $F : \Omega \to I\!R^n$.

Nous considérons dans ce mémoire le problème d'inéquations variationnelles VI(F, Ω) suivant :

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \forall x \in \Omega \quad F(x^*)^t(x - x^*) \geq 0.$$

Ce problème est une généralisation du problème

$$(P) \equiv \begin{cases} \min f(x) \\ s.c.\ x \in \Omega \end{cases}$$

où $f : I\!R^n \to I\!R$.

En effet, si on prend $F(x) = \nabla f(x)$,

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \nabla f(x^*)^t(x - x^*) \geq 0 \quad \forall x \in \Omega$$

est équivalent à

$$\text{Trouver } x^* \text{ solution de} \quad \begin{cases} \min f(x) \\ s.c.\ x \in \Omega. \end{cases}$$

Des exemples typiques d'ensemble Ω sont :

$\Omega = \{x |\ Ax = b,\ x \geq 0\}$ où A est une matrice $m \times n$ et $b \in I\!R^m$.
$\Omega = \{x |\ g_i(x) \leq 0\ i = 1, ..., m\}$ où g_i est une fonction convexe de $I\!R^n$ dans $I\!R$, pour tout i appartenant à l'ensemble $\{1, ..., m\}$.

1.2 Définitions et notions préliminaires

Avant d'aborder l'étude du problème VI(F, Ω), nous avons besoin de rappeler quelques définitions et notions de base.

1.2.1 Définitions d'un opérateur monotone et fortement monotone

On dit que l'opérateur $F : \Omega \rightarrow I\!R^n$ est **monotone** sur Ω si

$$[F(x) - F(y)]^t(x - y) \geq 0 \quad \forall x, y \in \Omega.$$

Lorsque F est différentiable, cette propriété est encore équivalente à

$$z^t F'(x) z \geq 0 \quad \forall x \in \Omega,\ \forall z \in I\!R^n.$$

où $F'(x)$ désigne la matrice jacobienne de F en x.

On dit que l'opérateur $F : \Omega \rightarrow I\!R^n$ est **fortement monotone** sur Ω s'il existe un scalaire $\lambda > 0$ tel que

$$[F(x) - F(y)]^t(x - y) \geq \lambda \parallel x - y \parallel^2 \quad \forall x, y \in \Omega.$$

Lorsque F est différentiable, cette propriété est encore équivalente à la propriété suivante :

$$\textit{Il existe } \lambda > 0 \ \textit{ tel que } \quad z^t F'(x) z \geq \lambda \parallel z \parallel^2 \quad \forall\, x \in \Omega,\ z \in I\!R^n.$$

1.2.2 Normes matricielles

Soit B une matrice symétrique et définie positive. A cette matrice, nous pouvons associer la norme $\| \, . \, \|_B$ définie par

$$\| x \|_B^2 = x^t B x \quad \forall \, x \in I\!R^n.$$

Lorsque B est la matrice identité I, la norme $\| \, . \, \|_B$ est la norme euclidienne, notée $\| \, . \, \|$.

Rappelons également la définition de la norme euclidienne d'une matrice carrée C :

$$\| C \| = \max\{\|Cx\| \mid \|x\| = 1\}.$$

1.2.3 Les fonctions barrières

Une méthode classique de résolution d'un problème d'optimisation avec contraintes consiste à remplacer ce problème par un problème sans contrainte en ajoutant une fonction barrière à la fonction objectif. Comme nous utiliserons cette méthode pour résoudre le problème VI(F, Ω), rappelons d'abord le principe de celle-ci dans le cas particulier de l'optimisation.

Considérons un problème de minimisation avec contrainte du type

$$(P) \equiv \left\{ \begin{array}{l} \min f(x) \\ s.c. \; x \in \Omega \end{array} \right.$$

où $f : I\!R^n \to I\!R$ est convexe et $\Omega \subseteq I\!R^n$ est un ensemble convexe, fermé et non vide.

Si l'on note $I(x)$ la fonction indicatrice de l'ensemble Ω (c'est-à-dire $I(x) = 0$ si $x \in \Omega$ et $I(x) = +\infty$ si $x \notin \Omega$) alors, il est facile de voir que ce problème peut se mettre sous la forme du problème de minimisation sans contrainte suivant :

$$\left\{ \begin{array}{l} \min f(x) + I(x) \\ s.c. \; x \in I\!R^n \end{array} \right.$$

La méthode consiste à approximer la fonction $I(x)$ par une fonction $\frac{1}{w}\hat{I}(x)$ vérifiant les propriétés suivantes :

1. $\hat{I}(x)$ est convexe et de classe C^1 sur l'intérieur de Ω;
2. Pour toute suite $\{x^k\}$ dans l'intérieur de Ω convergeant vers un point de la frontière de Ω, $\hat{I}(x^k)$ converge vers $+\infty$;
3. $w > 0$.

Pour certaines valeurs de w, on va résoudre le problème de minimisation sans contrainte suivant :

$$\begin{cases} \min wf(x) + \hat{I}(x) \\ s.c. \quad x \in I\!R^n \end{cases}$$

Notons $x(w)$ la solution obtenue. Sous des hypothèses appropriées, on peut montrer que, si $\{w^k\}$ est une suite tendant vers $+\infty$, alors la suite $\{x(w^k)\}$ converge vers la solution du problème (P).

La fonction $\hat{I}(x)$ est la **fonction barrière** associée au problème (P).

Remarquons que par la deuxième propriété de la fonction $\hat{I}$, on a que, si l'on applique une méthode de descente en partant d'un point $x^0 \in \Omega_{int}$, tous les itérés x^k vont rester strictement admissibles.

Exemple : La fonction barrière logarithmique.

Soit $\Omega = \{x | \; a_i^t x - b_i \geq 0 \;\; \forall i = 1, ..., m\}$ où $a_i \in I\!R^n$ et $b_i \in I\!R$, $i = 1, ..., m$.

La fonction barrière logarithmique associée à Ω est définie par

$$\hat{I}(x) = -\sum_{i=1}^{m} \log(a_i^t x - b_i).$$

Cette fonction vérifie bien les propriétés énoncées ci-dessus.
En effet, $\hat{I}$ est convexe et de classe C^1 sur l'ensemble Ω_{int}, et pour toute suite $\{x^k\}$ convergeant vers un point de la frontière de Ω, $-\sum_{i=1}^{m} \log(a_i^t x^k - b_i)$ converge vers $+\infty$.

1.2.4 Les fonctions de mérite

Lors de la résolution de VI(F, Ω), on va transformer le problème

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \forall x \in \Omega \quad F(x^*)^t(x - x^*) \geq 0$$

en un problème équivalent du type

$$\begin{cases} \min f(x) \\ s.c. \ x \in \Omega \end{cases}$$

La fonction $f(x)$ sera appelée la **fonction de mérite ou fonction gap** associée au problème VI(F, Ω).
C'est un procédé qui s'emploie régulièrement lors de la résolution d'inéquations variationnelles. L'avantage est que ce dernier problème peut être résolu grâce à des méthodes de descente dont la convergence est globale.

Exemples:

1. On a déjà vu au paragraphe 1.1 que lorsque $F(x) = \nabla f(x)$, le problème VI(F, Ω) est équivalent à
$$\begin{cases} \min f(x) \\ s.c. \ x \in \Omega \end{cases}$$
$f(x)$ est donc une fonction de mérite associée à VI$(\nabla f, \Omega)$.

2. Une fonction de mérite associée au problème VI(F, Ω) est donnée par
$$g(x) = \max\{ \ \langle F(x), x - y\rangle \mid y \in \Omega\}$$
Cette fonction vérifie les propriétés suivantes :
(a) $g(x) \geq 0 \quad \forall x \in \Omega$.
En effet, $\forall x \in \Omega \quad g(x) \geq \langle F(x), x - x\rangle = 0$.
(b) $g(x^*) = 0$ lorsque x^* satisfait $\langle F(x^*), x - x^*\rangle \geq 0 \quad \forall x \in \Omega$.
En effet, comme $\langle F(x^*), x^* - x\rangle \leq 0$ pour tout x de l'ensemble Ω, on a que $g(x^*) \leq 0$. Et donc par le point (a), on a que $g(x^*) = 0$.

1.2.5 Notes sur la différentiation

Afin de pouvoir définir certaines notions et notamment la propriété de self-concordance, nous devons au préalable définir les opérateurs de différentiation. Ce sera l'objet de ce paragraphe.

Considérons E un espace vectoriel réel de dimension finie ainsi qu'une fonction $F: E \to I\!R$ et un élément x de E tel que $F(x)$ est fini.

On dit que F **est différentiable** en x s'il existe une transformation linéaire $DF(x)$ sur E telle que

$$F(x+h) = F(x) + DF(x)[h] + o(\|h\|)$$

où $o(\|h\|)$ est une fonction en h telle que $o(\|h\|)$ converge uniformément vers 0 quand h converge vers 0. On appelle $DF(x)$ la **transformation dérivée de F** en x.

On peut ainsi définir la dérivée de F en x, $F'(x) \in E^*$ par la relation suivante :

$$DF(x)[h] = \langle F'(x), h \rangle = \frac{\partial}{\partial t_1} F(x + t_1 h) \mid_{t_1=0} .$$

Lorsqu'elle existe, cette dérivée est unique.

Si la dérivée de F existe pour tout x et est continue, alors on dit que F est continûment différentiable.

Supposons maintenant que F est différentiable en x et qu'il existe une transformation symétrique et bilinéaire $D^2F(x)$ sur $E \times E$ telle que

$$F(x+h) = F(x) + \langle F'(x), h \rangle + \frac{1}{2} D^2F(x)[h,h] + o(\|h\|^2).$$

On dit alors que F est **deux fois différentiable** en x. On appelle $D^2F(x)$ la **transformation dérivée seconde.**

On peut définir également $F''(x)$ par la relation suivante :

$$D^2F(x)[h_1, h_2] = \langle F''(x)h_1, h_2 \rangle = \frac{\partial^2}{\partial t_1 \partial t_2} F(x + t_1 h_1 + t_2 h_2) \mid_{t_1, t_2 = 0} .$$

On peut recommencer un raisonnement analogue pour la dérivée troisième. On a alors

$$D^3F(x)[h_1, h_2, h_3] = \frac{\partial^3}{\partial t_1 \partial t_2 \partial t_3} F(x + t_1 h_1 + t_2 h_2 + t_3 h_3) \mid_{t_1, t_2, t_3 = 0} .$$

Le raisonnement précédent peut se généraliser pour la dérivée quatrième, cinquième... Cependant, nous aurons besoin uniquement des dérivées d'ordre inférieur ou égal à 3.

Chapitre 2

Elaboration d'un algorithme pour résoudre le problème VI(F, Ω)

2.1 Cadre de travail

2.1.1 Définition du problème d'inéquations variationnelles VI(F, Ω)

Soient $\Omega = \{x | Ax = b,\ x \geq 0\} \subseteq \mathbb{R}^n$ où A est une matrice de $\mathbb{R}^{m\times n}$ $(m \leq n)$, b est un vecteur de $\mathbb{R}^m$ et $F : \Omega \to \mathbb{R}^n$ une fonction monotone. Le problème VI(F, Ω) est le suivant :

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \forall x \in \Omega \quad F(x^*)^t(x - x^*) \geq 0. \tag{2.1}$$

Dans ce chapitre, nous établirons un algorithme pour résoudre le problème VI(F, Ω). Pour cela, nous définirons par la méthode barrière une famille de fonctions et nous résoudrons le problème d'inéquations variationnelles associé à ces fonctions. Mais d'abord, il faudra poser certaines hypothèses sur F ainsi que sur l'ensemble Ω pour pouvoir démontrer la convergence des algorithmes étudiés.

2.1.2 Notations

Dans la suite, on notera

$\Omega_{int} = \{x|\ Ax = b,\ x > 0\}$, $\qquad \min(w) = \min_i\{w_i\}$,
$\Omega' = \{x|\ Ax = b\}$, $\qquad \max(w) = \max_i\{w_i\}$,
$X = \mathrm{diag}(x_1, ..., x_n)$, $\qquad e = (1, \ldots, 1)^t$,
$W = \mathrm{diag}(w_1, ..., w_n)$.

2.1.3 Les hypothèses sur A, Ω et F

Comme il a été annoncé précédemment, nous allons poser certaines hypothèses qui seront valables durant tout ce chapitre.

(A1) La matrice A est de rang plein;

(A2) $\Omega_{int} \neq \emptyset$;

(A3) L'ensemble Ω est borné,
(on notera $\gamma = \max\{x_i | \ x \in \Omega\}$ et $\Gamma = \min\{x_i | \ x \in \Omega\}$);

(A4) F est monotone;

(A5) F est continûment différentiable.

2.2 Existence de solutions

2.2.1 Les conditions de Karush-Kuhn-Tucker (KKT)

Remarquons d'abord que les hypothèses (A1)-(A2) sont équivalentes à l'hypothèse de qualification de contraintes de Slater. Dès lors, grâce aux conditions de KKT associées au problème VI(F, Ω), on sait que, si x^* est une solution de VI(F, Ω) alors il existe deux vecteurs $y^* \in I\!\!R^m$ et $s^* \in I\!\!R^n$ tels que

$$
\begin{aligned}
&F(x^*) - A^t y^* - s^* = 0 \\
&Ax^* = b \\
&X^* s^* = 0 \\
&x^* \geq 0 \\
&s^* \geq 0.
\end{aligned}
$$

De plus, comme F est monotone, la réciproque est également vraie.

Dans la suite de ce paragraphe, nous allons définir ce qu'est une solution faible. Cette notion nous sera très utile pour conclure sur l'existence et l'unicité des différents problèmes d'inéquations variationnelles que nous rencontrerons.

2.2.2 Les solutions faibles

Définition

$x \in \Omega$ est une solution faible de VI(F, Ω) si

$$\langle F(y), y - x \rangle \geq 0 \quad \forall y \in \Omega.$$

Proposition 2.2.1

Si Ω est un ensemble fermé convexe à intérieur non vide et si F est monotone, on a les deux résultats suivants:

1. *Toute solution de VI(F,Ω) est une solution faible de VI(F,Ω).*
2. *Si F est continue, alors toute solution faible de VI(F,Ω) est une solution de VI(F,Ω).*

preuve

1. Démontrons d'abord que toute solution de VI(F,Ω) est une solution faible de VI(F,Ω).

   Soit x^* une solution de VI(F,Ω).
   Par définition, x^* vérifie la relation $\langle F(x^*), x - x^* \rangle \geq 0 \;\; \forall x \in \Omega$.

   Comme F est monotone, on a également

   $$\langle F(x^*) - F(x), x^* - x \rangle \geq 0 \quad \forall x \in \Omega$$

   et donc

   $$0 \leq \langle F(x^*), x - x^* \rangle \leq \langle F(x), x - x^* \rangle \quad \forall x \in \Omega,$$

   c'est-à-dire, x^* est une solution faible de VI(F,Ω).

2. Démontrons maintenant que si F est continue, toute solution faible de VI(F,Ω) est une solution de VI(F,Ω).
   Si x^* est une solution faible de VI(F,Ω), alors, pour tout $y \in \Omega$, on a

   $$\langle F(x^* + t(y - x^*)), y - x^* \rangle \geq 0 \quad \text{et} \quad 0 \leq t \leq 1.$$

   La fonction F étant continue, on en déduit en passant à la limite lorsque t tend vers 0, que

   $$\langle F(x^*), y - x^* \rangle \geq 0 \quad \forall y \in \Omega,$$

   c'est-à-dire que x^* est une solution de VI(F,Ω).

□

Lemme 2.2.1

Soient $G(F)=\{ (x,y) \mid y = F(x),\ x \in \Omega\}$, *le graphe de la fonction* F;
$S=\{ (x_i, F(x_i)),\ i = 1,...,m\}$ *un sous-ensemble fini du graphe de* F;
$X(S)=\{x \in \Omega \mid \langle F(x_i), x_i - x\rangle \geq 0 \quad i = 1,...,m\}$.

Si $X(S)$ *est vide et si* Ω *est un ensemble convexe fermé et borné, alors il existe des scalaires* $\lambda_i \geq 0$, $i = 1,...,m$ *et* $\delta > 0$ *tels que* $\sum_{i=1}^{m} \lambda_i = 1$ *et*

$$\sum_{i=1}^{m} \lambda_i \langle F(x_i), x_i - x\rangle \leq -\delta < 0 \quad \forall x \in \Omega.$$

preuve

Considérons le problème

$$(P) \equiv \begin{cases} \max\ t \\ s.c.\ t \leq \langle F(x_i), x_i - x\rangle \quad i = 1,...,m \\ \quad x \in \Omega. \end{cases}$$

Notons t^* la valeur optimale de la fonction objectif de (P). Vu que nous avons supposé que $X(S)$ est vide, t^* doit être strictement négatif.

Pour démontrer la thèse, nous devons construire le problème dual de (P). Par définition, la fonction duale est de la forme suivante :

$$d(\lambda) = \sup_{x\in\Omega,\ t\in I\!R} L(x,t,\lambda) = \sup_{x\in\Omega,\ t\in I\!R} \left\{ t + \sum_{i=1}^{m} \lambda_i[\langle F(x_i), x_i - x\rangle - t] \right\}$$

où les composantes du multiplicateur de Lagrange λ sont positives.
Développons l'expression de $d(\lambda)$:

$$\begin{aligned} d(\lambda) &= \sup_{t\in I\!R} (1 - \sum_{i=1}^{m} \lambda_i)t + \sup_{x\in\Omega} \sum_{i=1}^{m} \lambda_i \langle F(x_i), x_i - x\rangle \\ &= \begin{cases} \sup_{x\in\Omega} \sum_{i=1}^{m} \lambda_i \langle F(x_i), x_i - x\rangle & \text{si } \sum_{i=1}^{m} \lambda_i = 1 \\ +\infty & \text{sinon.} \end{cases} \end{aligned}$$

Vu que Ω est un ensemble fermé et borné, le supremum est atteint en un point de Ω. Le problème dual peut donc s'écrire comme suit

$$(D) \equiv \begin{cases} \min \; \{\sup_{x\in\Omega} \sum_{i=1}^{m} \lambda_i \langle F(x_i), x_i - x\rangle\} \\ \text{s.c.} \;\; \lambda \geq 0, \\ \qquad \sum_{i=1}^{m} \lambda_i = 1. \end{cases}$$

Or, on a par la dualité forte que $d^* = t^* < 0$ où d^* représente la valeur optimale de la fonction objectif du problème dual.
Dès lors, il existe $\lambda^* \geq 0$ tel que $\sum_{i=1}^{m} \lambda_i^* = 1$ et

$$\sup_{x\in\Omega} \sum_{i=1}^{m} \lambda_i^* \langle F(x_i), x_i - x\rangle < 0,$$

ce qui peut encore s'énoncer de la façon suivante :

il existe $\lambda^* \geq 0$ et $\delta > 0$ tels que $\sum_{i=1}^{m} \lambda_i^* = 1$ et

$$\sum_{i=1}^{m} \lambda_i^* \langle F(x_i), x_i - x\rangle \leq -\delta < 0 \quad \forall x \in \Omega.$$

□

Proposition 2.2.2

Si en plus des hypothèses de la Proposition 2.2.1, on ajoute l'hypothèse (A3), alors le problème VI(F, Ω) admet des solutions faibles.

preuve

Soient $G(F)$, S et $X(S)$, les ensembles définis dans l'énoncé du Lemme 2.2.1.

1. Montrons d'abord que l'ensemble $X(S)$ est non vide.
   Supposons par l'absurde que X(S) est vide.
   Grâce au Lemme 2.2.1, nous savons qu'il existe des scalaires $\lambda_i \geq 0$, $\sum_{i=1}^{m} \lambda_i = 1$, tels que la fonction linéaire $g(x) = \sum_{i=1}^{m} \lambda_i \langle F(x_i), x_i - x\rangle$ est négative sur Ω, c'est-à-dire

   $$g(x) \leq -\delta < 0 \quad \forall x \in \Omega, \text{ où } \delta > 0.$$

   Posons $x^* = \sum_{i=1}^{m} \lambda_i x_i$. Comme Ω est convexe, $x^* \in \Omega$.

Soient $z \in \Omega_{int}$ et $x(t) = x^* + t(z - x^*), \quad 0 < t \leq 1$.
Dès lors, $x(t) \in \Omega_{int} \subseteq \mathrm{Dom}(F)$, $t \in]0,1]$ et on a pour tout $0 < t \leq 1$:

$$\begin{aligned}
-\delta \geq g(x(t)) &= \textstyle\sum_{i=1}^m \lambda_i \langle F(x_i), x_i - x(t)\rangle \\
&\geq \textstyle\sum_{i=1}^m \lambda_i \langle F(x(t)), x_i - x(t)\rangle \\
&= \langle F(x(t)), x^* - x(t)\rangle = -t\langle F(x(t)), z - x^*\rangle
\end{aligned}$$

où l'inégalité a été obtenue par monotonie de F.

Si l'on pose $w(t) = \langle F(x(t)), z - x^*\rangle$, l'inégalité précédente peut être réécrite de la façon suivante :

$$tw(t) \geq \delta > 0 \quad \forall t \in (0,1].$$

Montrons que la fonction w est croissante sur $(0,1]$.

En effet, prenons $0 < t_1 < t_2 \leq 1$ et montrons que $w(t_1) \leq w(t_2)$.
Par définition de $w(t_1)$ et de $w(t_2)$, on a

$$w(t_1) = \langle F(x(t_1)), z - x^*\rangle \quad \text{et} \quad w(t_2) = \langle F(x(t_2)), z - x^*\rangle.$$

Comme F est monotone, on a la relation suivante :

$$\begin{aligned}
\langle F(x(t_2)) - F(x(t_1)), x(t_2) - x(t_1)\rangle &= \langle F(x(t_2)) - F(x(t_1)), (t_2 - t_1)(z - x^*)\rangle \\
&= (t_2 - t_1)(w(t_2) - w(t_1)) \geq 0
\end{aligned}$$

et donc $w(t)$ est croissant sur $(0,1]$.

On obtient alors :

$$\lim_{\substack{t \to 0 \\ >}} tw(t) \leq \lim_{\substack{t \to 0 \\ >}} t\, w(1) = 0.$$

Mais ceci n'est pas possible car $tw(t) \geq \delta > 0, \ 0 < t \leq 1$.
On vient donc de démontrer que $X(S)$ est non vide pour tout S.

2. Montrons ensuite que $\mathrm{VI}(F, \Omega)$ admet des solutions faibles.
   Considérons S_i, $i \in \{1, ..., p\}$, p sous-ensembles finis de $G(F)$.
   Ces ensembles sont fermés et vérifient la propriété suivante :

$$\bigcap_{i=1}^{p} X(S_i) \neq \emptyset.$$

$G(F)$ étant compact, on a alors

$$\bigcap_{S \subseteq G(F)} X(S) \neq \emptyset.$$

Or, cette dernière intersection est l'ensemble de toutes les solutions faibles de VI(F, Ω).

$\square$

Grâce aux propositions précédentes, on peut facilement démontrer que le problème VI(F, Ω) admet au moins une solution.

2.3 Les sous-problèmes barrières VI$(\bar{F}, \Omega_{int})$

2.3.1 Définition

Soit w, un vecteur dont les composantes sont positives. Définissons la famille suivante de fonctions :

$$\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{X}^{-1}\mathbf{w}.$$

Le vecteur w est le paramètre de pénalité associé à la fonction F.

Pour w positif quelconque, considérons le sous-problème barrière VI$(\bar{F}, \Omega_{int})$ défini comme suit :

$$\text{Trouver } x(w) \in \Omega_{int} \quad \text{tel que} \quad \bar{F}(x(w))^t[x - x(w)] \geq 0 \quad \forall x \in \Omega_{int}. \tag{2.2}$$

L'ensemble $\{x(w),\ w > 0\}$, s'il existe, s'appelle trajectoire centrale. La méthode proposée pour résoudre le problème VI(F, Ω) va consister à suivre cette trajectoire centrale.

Le lemme suivant nous indique que, sous les hypothèses posées au début de ce chapitre, $\bar{F}$ est fortement monotone.

Lemme 2.3.1

Sous les hypothèses (A3)-(A5), $\bar{F}(x)$ est fortement monotone sur Ω de constante $\lambda = min(w)/\gamma^2$.

preuve

Il faut montrer que $\langle \bar{F}(x) - \bar{F}(y), x-y\rangle \geq \lambda \|x-y\|^2 \quad \forall x, y \in \Omega$.

$$\begin{aligned}\langle \bar{F}(x) - \bar{F}(y), x-y\rangle &= \langle F(x) - F(y), x-y\rangle - \langle X^{-1}w - Y^{-1}w, x-y\rangle \\ &\geq \langle (Y^{-1} - X^{-1})w, x-y\rangle. \end{aligned} \tag{2.3}$$

$$\text{Comme } (Y^{-1} - X^{-1})w = \begin{pmatrix} \frac{w_1(x_1-y_1)}{x_1y_1} \\ \vdots \\ \frac{w_n(x_n-y_n)}{x_ny_n} \end{pmatrix}$$

on obtient $\langle (Y^{-1} - X^{-1})w, x-y\rangle = \frac{w_1}{x_1y_1}(x_1-y_1)^2 + \ldots + \frac{w_n}{x_ny_n}(x_n-y_n)^2$,

et constatant que $\frac{w_i}{x_iy_i} \geq \frac{\min(w)}{\gamma^2} \quad \forall i \in \{1, \ldots, n\}$,

$$\begin{aligned}\langle (Y^{-1} - X^{-1})w, x-y\rangle &\geq \frac{\min(w)}{\gamma^2}\{(x_1-y_1)^2 + \ldots + (x_n-y_n)^2\} \\ &= \frac{\min(w)}{\gamma^2}\|x-y\|^2. \end{aligned} \tag{2.4}$$

La thèse se déduit alors des relations (2.3) et (2.4).

□

Théorème 2.1

Il existe une et une seule solution au problème (2.2).

preuve

1. L'existence d'une solution se démontre facilement en considérant les hypothèses (A1)-(A5) ainsi que les Propositions 2.2.1 et 2.2.2.

2. Démontrons l'unicité de la solution.
   Procédons par l'absurde. Supposons qu'il existe deux solutions au problème (2.2) et notons-les $x(w)$ et $y(w)$.
   Par le Lemme 2.3.1, on a

$$\lambda\|x(w) - y(w)\|^2 \leq [\bar{F}(x(w)) - \bar{F}(y(w))]^t(x(w) - y(w)). \tag{2.5}$$

Or, $x(w)$ et $y(w)$ vérifient respectivement

$$\bar{F}(x(w))(x - x(w)) \geq 0 \quad \forall x \in \Omega_{int} \tag{2.6}$$

et

$$\bar{F}(y(w))(x - y(w)) \geq 0 \quad \forall x \in \Omega_{int}. \tag{2.7}$$

En prenant $x = y(w)$ dans (2.6), on a

$$\bar{F}(x(w))(x(w) - y(w)) \leq 0.$$

En prenant $x = x(w)$ dans (2.7) on a

$$\bar{F}(y(w))(x(w) - y(w)) \geq 0.$$

Donc, la relation (2.5) devient

$$\lambda \|x(w) - y(w)\|^2 \leq [\bar{F}(x(w)) - \bar{F}(y(w))]^t(x(w) - y(w)) \leq 0,$$

ce qui implique que $x(w) = y(w)$.

□

2.3.2 Reformulation de VI($\bar{F}, \Omega_{int}$)

Dans ce paragraphe, nous reformulons le problème VI($\bar{F}, \Omega_{int}$) grâce à une fonction de mérite. Pour définir cette fonction de mérite, il va falloir définir une fonction intermédiaire $P(x)$.

La fonction de mérite associée à VI($\bar{F}, \Omega_{int}$)

Pour chaque $x \in \Omega_{int}$, considérons $\mathbf{P(x)} \in \Omega$ la solution du problème d'inéquations variationnelles suivant :

$$[\bar{F}(x) + B(P(x) - x)]^t[y - P(x)] \geq 0, \quad \forall y \in \Omega, \tag{2.8}$$

où B est une matrice symétrique et définie positive. Ce problème est équivalent au problème suivant :

$$\min_{y \in \Omega} \|y - (x - B^{-1}\bar{F}(x))\|_B^2,$$

dont la solution optimale est

$$P(x) = \text{Proj}_{\Omega,B}(x - B^{-1}\bar{F}(x)).$$

Comme Ω est un ensemble fermé, borné et convexe et B est une matrice symétrique et définie positive, on peut facilement démontrer que ce problème admet une solution unique.

Nous pouvons alors définir la fonction de mérite $f_w : \Omega_{int} \to I\!R$ pour chaque $x \in \Omega_{int}$:

$$f_w(x) = -\bar{F}(x)^t(P(x) - x) - \frac{1}{2}[P(x) - x]^t B[P(x) - x], \tag{2.9}$$

et considérer le problème d'Optimisation suivant :

$$\min_{x\in\Omega_{int}} f_w(x) = \min_{x\in\Omega_{int}} \left[-\bar{F}(x)^t(P(x) - x) - \frac{1}{2}[P(x) - x]^t B[P(x) - x]\right]. \tag{2.10}$$

Avant d'étudier les propriétés de f_w, définissons, pour chaque $x \in \Omega_{int}$, le vecteur $\mathbf{Z(x)} \in \Omega'$ comme étant la solution du problème suivant :

$$[\bar{F}(x) + \bar{F}'(x)^t(Z(x) - x)]^t[y - Z(x)] \geq 0 \quad \forall y \in \Omega'. \tag{2.11}$$

On peut facilement démontrer que ce problème admet une solution vu que $\bar{F}'(x)$ est une matrice définie positive et A est une matrice de rang plein [7].

Propriétés de la fonction de mérite f_w

Proposition 2.3.1

Soient la fonction f_w définie ci-dessus et $x(w)$ la solution du problème $VI(\bar{F}, \Omega_{int})$. La fonction f_w vérifie les propriétés suivantes:

1. $f_w(x) \geq 0$, $\forall x \in \Omega_{int}$;
2. $f_w(x) = 0$ *si et seulement si x est une solution du problème $VI(\bar{F}, \Omega_{int})$ et du problème (2.10);*

3. $f_w(x) \geq [\frac{min(w)}{\gamma^2} - \frac{1}{2}\|B\|\ ]\ \|x(w) - x\|^2 \quad \forall x \in \Omega_{int}$;

4. *$f_w(x)$ est continûment différentiable et son gradient est donné par la formule suivante:*

$$f'_w(x) = \bar{F}(x) - [\bar{F}'(x) - B][P(x) - x] \quad \forall x \in \Omega_{int};$$

5. *Une direction de descente pour $f_w(x)$ en x est donnée par $d = Z(x) - x$ lorsque $Z(x) \neq x$. Plus précisément,*

$$f'_w(x)^t[Z(x) - x] \leq -f_w(x), \quad \forall x \in \Omega_{int}, \quad \text{si } \|B\| \leq 2\frac{min(w)}{\gamma^2}.$$

preuve

1. Selon la relation (2.8) où l'on remplace y par x, on a

$$\bar{F}(x)^t[P(x) - x] \leq -\|P(x) - x\|_B^2. \tag{2.12}$$

Il suit alors de la définition de f_w que

$$f_w(x) = -\bar{F}(x)^t(P(x) - x) - \tfrac{1}{2}[P(x) - x]^t B[P(x) - x]$$
$$\geq \tfrac{1}{2}\|P(x) - x\|_B^2 \geq 0.$$

2. (a) Démontrons d'abord que si $f_w(x) = 0$ où $x \in \Omega$ alors x résout le problème VI($\bar{F}, \Omega_{int}$) ainsi que le problème (2.10).
Il suit de la définition de f_w que

$$\begin{aligned}\bar{F}(x)^t(P(x) - x) &= -\frac{1}{2}[P(x) - x]^t B[P(x) - x] \\ &= -\frac{1}{2}\|P(x) - x\|_B^2.\end{aligned} \tag{2.13}$$

En combinant (2.12) et (2.13), on a que

$$-\frac{1}{2}\|P(x) - x\|_B^2 \leq -\|P(x) - x\|_B^2$$

et donc

$$\frac{1}{2}\|P(x) - x\|_B^2 \leq 0$$

ce qui entraîne que $P(x) = x$ vu que B est une matrice définie positive. Dès lors, par la relation (2.8), on a que x est la solution de VI($\bar{F}, \Omega_{int}$).

(b) Montrons maintenant que si $x \in \Omega$ est la solution de VI($\bar{F}, \Omega_{int}$) et du problème (2.10) alors $f_w(x) = 0$.
Prenons $x \in \Omega_{int}$, la solution de VI($\bar{F}, \Omega_{int}$). Par passage à la limite dans l'inégalité (2.2), on a que

$$\bar{F}(x)^t(y - x) \geq 0 \quad \forall y \in \Omega.$$

Prenons $y = P(x) \in \Omega$. On obtient alors

$$\bar{F}(x)^t[P(x) - x] \geq 0.$$

Par la relation (2.8) où l'on remplace y par $x \in \Omega$, on a que

$$[\bar{F}(x) + B(P(x) - x)]^t[x - P(x)] \geq 0$$

ce qui implique que

$$-[P(x) - x]^t B[P(x) - x] \geq \bar{F}(x)^t[P(x) - x] \geq 0.$$

Donc, $P(x) = x$ car la matrice B est symétrique et définie positive. On a donc que $f_w(x) = 0$.

3. Prenons $x(w)$ la solution du problème VI($\bar{F}, \Omega_{int}$). Vu que $\bar{F}(x)$ est fortement monotone, on a :

$$[\bar{F}(x) - \bar{F}(x(w))]^t[x - x(w)] \geq \frac{\min(w)}{\gamma^2}\|x(w) - x\|^2 \quad \forall x \in \Omega_{int},$$

et par définition de $x(w)$,

$$\bar{F}(x(w))[x - x(w)] \geq 0 \quad \forall x \in \Omega_{int}.$$

Les deux inégalités précédentes entraînent alors que

$$\bar{F}(x)[x - x(w)] \geq \frac{\min(w)}{\gamma^2}\|x(w) - x\|^2 \quad \forall x \in \Omega_{int}.$$

On a alors successivement les relations suivantes :

$$\begin{aligned} f_w(x) &= -\bar{F}(x)^t[P(x) - x] - \frac{1}{2}[P(x) - x]^t B[P(x) - x] \\ &= -\min_{y \in \Omega}\left[\bar{F}(x)^t(y - x) + \frac{1}{2}(y - x)^t B(y - x)\right] \end{aligned}$$

$$
\begin{aligned}
f_w(x) &\geq -\bar{F}(x)^t[x(w)-x] - \frac{1}{2}[x(w)-x]^t B[x(w)-x] \qquad (2.14)\\
&\geq \frac{\min(w)}{\gamma^2}\|x(w)-x\|^2 - \frac{1}{2}\|B\|\;\|x(w)-x\|^2\\
&= \left[\frac{\min(w)}{\gamma^2} - \frac{1}{2}\|B\|\right]\;\|x(w)-x\|^2,
\end{aligned}
$$

où la première inégalité vient du fait que $x(w) \in \Omega_{int}$. On a ainsi démontré la troisième propriété de la proposition.

4. Tout d'abord, nous savons que $\bar{F}$ est continûment différentiable sur l'ensemble Ω_{int}. De plus, pour tout $x \in \Omega_{int}$, $P(x) = \mathrm{Proj}_{\Omega,B}(x - B^{-1}\bar{F}(x))$. On a donc immédiatement que $P(x)$ est continue.

   Définissons ensuite une fonction intermédiaire $h : \Omega_{int} \times I\!R^n \to I\!R$ par la relation suivante :

$$
h(x,y) = \langle \bar{F}(x), y-x\rangle + \frac{1}{2}\langle y-x, B(y-x)\rangle.
$$

   Vu que $\bar{F}$ est continûment différentiable, h l'est également.
   Par définition,

$$
\begin{aligned}
f_w(x) &= -\min_{y\in\Omega}\left(\bar{F}(x)^t(y-x) + \frac{1}{2}(y-x)^t B(y-x)\right)\\
&= -\min\{h(x,y)|\; y \in \Omega\}.
\end{aligned}
$$

   Comme ce minimum est atteint en $y = P(x)$, f_w est différentiable sur l'ensemble Ω_{int}.

   Finalement, pour trouver l'expression du gradient de f_w, il suffit de considérer les égalités suivantes qui sont valables pour tout x appartenant à l'ensemble Ω_{int} :

$$
\begin{aligned}
\nabla f_w(x) &= -\nabla_x h(x, P(x))\\
&= \bar{F}(x) - \bar{F}'(x)^t(P(x)-x) + B(P(x)-x)\\
&= \bar{F}(x) - [\bar{F}'(x) - B][P(x)-x].
\end{aligned}
$$

5. Pour démontrer le dernier point, remarquons d'abord que

$$\begin{aligned} f'_w(x)^t[Z(x)-x] &= \bar{F}(x)^t[Z(x)-x] - [P(x)-x]^t[\bar{F}'(x)^t - B][Z(x)-x] \\ &= T_1 + T_2 \end{aligned}$$

où $T_1 = \bar{F}(x)^t[Z(x)-x]$ et $T_2 = -[P(x)-x]^t[\bar{F}'(x)^t - B][Z(x)-x]$.

Développons d'abord T_2 :

$$\begin{aligned} T_2 &= [\bar{F}(x) + \bar{F}'(x)^t(Z(x)-x)]^t[x-P(x)] - \bar{F}(x)^t[x-P(x)] \\ &\quad +[P(x)-x]^tB[Z(x)-x] \qquad (2.15) \\ &= -[\bar{F}(x) + \bar{F}'(x)^t(Z(x)-x)]^t[P(x)-Z(x)] \\ &\quad +[\bar{F}(x) + \bar{F}'(x)^t(Z(x)-x)]^t[x-Z(x)] \\ &\quad -\bar{F}(x)^t[x-P(x)] + [P(x)-x]^tB[Z(x)-x] \qquad (2.16) \\ &\leq [\bar{F}(x) + \bar{F}'(x)^t(Z(x)-x)]^t[x-Z(x)] - \bar{F}(x)^t[x-P(x)] \\ &\quad +[P(x)-x]^tB[Z(x)-x] \qquad (2.17) \\ &\leq -\bar{F}(x)^t[Z(x)-x] - [Z(x)-x]^t\bar{F}'(x)[Z(x)-x] - \bar{F}(x)^t[x-P(x)] \\ &\quad -\frac{1}{2}\|Z(x)-P(x)\|_B^2 + \frac{1}{2}\|Z(x)-x\|_B^2 + \frac{1}{2}\|P(x)-x\|_B^2 \qquad (2.18) \\ &= -\bar{F}(x)^t[Z(x)-x] - [Z(x)-x]^t\bar{F}'(x)[Z(x)-x] - f_w(x) \\ &\quad -\frac{1}{2}\|Z(x)-P(x)\|_B^2 + \frac{1}{2}\|Z(x)-x\|_B^2 \qquad (2.19) \end{aligned}$$

où la première inégalité vient de la relation (2.11) et du fait que $P(x) \in \Omega \subseteq \Omega'$. Pour obtenir la deuxième inégalité, il suffit de suivre le raisonnement suivant :

Vu que $(Z(x)-P(x)) = (Z(x)-x) - (P(x)-x)$, on a que

$\|Z(x)-P(x)\|_B^2 = \|Z(x)-x\|_B^2 + \|P(x)-x\|_B^2 - 2(P(x)-x)^tB(Z(x)-x)$.

Prenons ensuite $d = Z(x) - x$. Finalement, on obtient :

$$\begin{aligned}
T_1 + T_2 &\leq -d^t \bar{F}'(x)^t d - f_w(x) - \frac{1}{2}\|Z(x) - P(x)\|_B^2 + \frac{1}{2} d^t B d \\
&\leq -\left(\frac{\min(w)}{\gamma^2} - \frac{1}{2}\|B\|\right) \|Z(x) - x\|^2 - f_w(x) - \frac{1}{2}\|Z(x) - P(x)\|_B^2 \\
&\leq -f_w(x),
\end{aligned}$$

où la dernière inégalité est due à l'hypothèse $\frac{\min(w)}{\gamma^2} \geq \frac{1}{2}\|B\|$.

$\square$

Commentaire

Lorsque le vecteur w change, la condition sur la matrice B change aussi. Donc, pour tout nouveau vecteur w, il se peut qu'on doive changer la matrice B afin de satisfaire l'hypothèse $\frac{\min(w)}{\gamma^2} \geq \frac{1}{2}\|B\|$. On peut choisir $B = 2\frac{\min(w)}{\gamma^2} I$ où I est la matrice identité $n \times n$.

2.3.3 La méthode de Newton

Pour élaborer un algorithme, il faut connaître la direction dans laquelle nous nous dirigerons pour définir chaque nouvel itéré. La direction que nous emploierons est celle obtenue par la méthode de Newton.
L'objet de ce paragraphe sera donc de développer la méthode de Newton afin de trouver l'expression de la direction. Pour cela, nous avons besoin des conditions de KKT.

Les conditions de KKT associées au problème VI$(\bar{F}, \Omega_{int})$ sont

$$\begin{aligned}
&F(x) - X^{-1}w - A^t y = 0, \\
&Ax = b, \quad x > 0 \\
&x \in I\!R^n, \quad y \in I\!R^m.
\end{aligned}$$

Posons

$$H(x,y) = \begin{bmatrix} F(x) - X^{-1}w - A^t y \\ Ax - b \end{bmatrix} \qquad H'(x,y) = \begin{bmatrix} F'(x)^t + X^{-2}W & -A^t \\ A & 0 \end{bmatrix}$$

Etant donné un couple (x, y), où $x \in \Omega_{int}$ et $y \in I\!R^m$, pour trouver la direction de Newton il faut résoudre le système suivant :

$$H(x,y) + H'(x,y) \begin{bmatrix} \hat{x} - x \\ \hat{y} - y \end{bmatrix} = 0.$$

Définissons $\Delta x = \hat{x} - x$, $\Delta y = \hat{y} - y$.
Le sytème ci-dessus peut-être écrit sous la forme :

$$\begin{cases} F(x) - X^{-1}w - A^t y + [F'(x)^t + X^{-2}W]\Delta x - A^t \Delta y = 0 \\ Ax - b + A\Delta x = 0, \end{cases}$$

ce qui revient à :

$$\begin{cases} F(x) - X^{-1}w + [F'(x)^t + X^{-2}W]\Delta x - A^t \hat{y} = 0 \\ A\hat{x} = b. \end{cases} \tag{2.20}$$

Vu que la matrice A est de rang plein et que $F'(x)$ est une matrice définie positive, ce sytème est bien défini.

Δx est la direction de Newton que nous cherchons. En fait, ce sytème représente les conditions de KKT du problème (11). On a donc que $\hat{x} = Z(x)$ et donc, la direction de Newton Δx est également définie par

$$\Delta x = Z(x) - x.$$

2.3.4 Premier algorithme pour résoudre les sous-problèmes VI$(\bar{F}, \Omega_{int})$: l'algorithme 2.1

Dans ce paragraphe, un premier algorithme pour résoudre VI$(\bar{F}, \Omega_{int})$ sera présenté. Grâce aux hypothèses posées au début de ce chapitre, nous pourrons montrer que cet algorithme est convergent.

Initialisation : Choisir $x^0 \in \Omega_{int}$, $0 < \rho, \sigma < 1$, et poser $k = 0$.

Itération k : Effectuer les pas 1 à 4 ci-dessous.

Pas 1: Critère d'arrêt

Si un critère de convergence est satisfait, alors on s'arrête.
Ce critère de convergence peut être

$$f_w(x^k) \leq \varepsilon \quad \text{ou} \quad \|Z(x^k) - x^k\| \leq \varepsilon \ \text{ avec } \varepsilon > 0$$

Pas 2: **La direction de Newton**

Trouver $\hat{x} = Z(x^k)$ la solution du système suivant :

$$\begin{cases} F(x^k) - X_k^{-1}w + [F'(x^k)^t + X_k^{-2}W](\hat{x} - x^k) - A^t\hat{y} = 0 \\ A\hat{x} = b. \end{cases}$$

où $X_k \equiv \text{diag}(x_1^k, \ldots, x_n^k)$. On obtient ainsi $d^k = \hat{x} - x^k$.

Pas 3: **Recherche linéaire inexacte**

Choisir t^k, le plus grand nombre de l'ensemble $\{1, \rho, \rho^2, ...\}$ tel que :

$$\begin{cases} f_w(x^k + t^k d^k) \leq f_w(x^k) + \sigma t^k f_w'(x^k)^t d^k, \text{ [règle d'Armijo]} \\ x^k + t^k d^k \in \Omega_{int} \end{cases}$$

Pas 4: **Mise à jour**

Poser $x^{k+1} = x^k + t^k d^k, \ y^{k+1} = \hat{y}, \ k = k + 1.$

Théorème 2.2 Convergence de l'algorithme.

Soient $\|B\| \leq 2\frac{min(w)}{\gamma^2}$ *et* $x^0 \in \Omega_{int}$.
L'algorithme 2.1 engendre une suite d'itérés $\{x^k\}$ *convergeant vers l'unique solution* $x(w)$ *du problème* $VI(\bar{F}, \Omega_{int})$.

preuve

On sait déjà que $\bar{F}(x)$ est continûment différentiable et fortement monotone sur Ω_{int} et que la suite des itérés $\{x^k\}$ reste dans l'ensemble Ω_{int} (car l'algorithme 2.1 l'impose).

1. Montrons d'abord que la suite $\{f_w(x^k)\}$ est décroissante et que la suite $\{x_k\}$ est bornée.
   Par le cinquième résultat de la Proposition 2.3.1 et la définition de la règle d'Armijo, la suite $\{f_w(x^k)\}$ est décroissante. En effet,

$$\begin{aligned} f_w(x^k + t^k d^k) &\leq f_w(x^k) + \sigma t^k f_w'(x^k)^t d^k, \\ &\leq f_w(x^k) + \sigma t^k(-1) f_w(x^k), \end{aligned}$$

par conséquent,

$$\begin{aligned} f_w(x^k + t^k d^k) &\leq (1 - \sigma t^k) f_w(x^k) \\ &\leq f_w(x^k). \end{aligned}$$

De plus, si on considère la troisième propriété de la Proposition 2.3.1, on peut montrer que la suite $\{x^k\}$ est bornée.
En effet,

$$\begin{aligned} \|x(w) - x^k\|^2 &\leq \frac{1}{\frac{\min(w)}{\gamma^2} - \frac{1}{2}\|B\|} f_w(x^k) \\ &\leq \frac{1}{\frac{\min(w)}{\gamma^2} - \frac{1}{2}\|B\|} f_w(x^0). \end{aligned}$$

2. Montrons ensuite que la fonction $f_w(x)$ converge vers $+\infty$ lorsque x converge vers un point de la frontière de Ω_{int}.

   Soient $\bar{\Omega}_{int}$ la fermeture de Ω_{int} ,
   $M = \max_{x,y \in \Omega} \|x - y\|$,
   $\bar{L} = \max_{x \in \Omega} \|F(x)\|$,
   $\lambda_{\max}(B)$, la plus grande valeur propre de la matrice B .

   Par la relation (2.14) de la démonstration de la troisième propriété de la Proposition 2.3.1, on a que

$$f_w(x) \geq -\bar{F}(x)^t[x(w) - x] - \frac{1}{2}[x(w) - x]^t B[x(w) - x].$$

D'où,

$$\begin{aligned} f_w(x) &\geq -[F(x) - X^{-1}w]^t[x(w) - x] - \frac{1}{2}\|x(w) - x\|_B^2 \\ &\geq -F(x)^t[x(w) - x] + (X^{-1}w)^t[x(w) - x] - \frac{1}{2}\|M\|^2 \lambda_{\max}(B) \\ &\geq -\bar{L}M + w^t[X^{-1}x(w)] - w^t e - \frac{1}{2}\|M\|^2 \lambda_{\max}(B) \end{aligned}$$

où la deuxième inégalité vient d'une propriété des quotients de Rayleigh :

$$\frac{(x(w) - x)^t B(x(w) - x)}{\|x(w) - x\|^2} \leq \lambda_{\max}(B)$$

Vu que $w, x(w) > 0$, on a donc que $\lim_{x \to \bar{\Omega}_{int}} f_w(x) = +\infty$.

Or, on a montré que la fonction $f_w(x)$ est bornée sur l'ensemble Ω_{int}. Par conséquent, la suite des itérés va toujours rester hors de la frontière de l'ensemble Ω_{int}. La suite possède donc au moins un point d'accumulation dans un sous-ensemble compact de Ω_{int}.

Prenons $\{x^k\}_{k \in \bar{K}}$ une sous-suite convergente de $\{x^k\}$ et notons sa limite $\bar{x} \in \Omega_{int}$. $Z(x)$ étant une projection, elle est continue sur Ω_{int}. Par conséquent, la suite $\{d^k\}_{k \in \bar{K}}$ converge vers le vecteur $\bar{d} = Z(\bar{x}) - \bar{x}$.

3. Pour démontrer la thèse, nous devons finalement prouver que $f_w(\bar{x}) = 0$.

   En effet, si $f_w(\bar{x}) = 0$, par la deuxième propriété de la Proposition 2.3.1, on sait que $\bar{x}$ est une solution du problème (2.10).
   Par l'absurde, supposons que $f_w(\bar{x})>0$. Il existe donc $\varepsilon>0$ et un indice $\bar{k}$ tels que
   $$f_w(x^k) > 0, \quad k \geq \bar{k}.$$
   On obtient donc pour $k \geq \bar{k}$:
   $$f'_w(x^k)^t d^k < -f_w(x^k) < 0, \tag{2.21}$$
   $$f_w(x^{k+1}) - f_w(x^k) \leq \sigma t^k f'_w(x^k)^t d^k \tag{2.22}$$
   $$f_w(x^k + (\tfrac{t^k}{\rho} d^k)) - f_w(x^k) > \sigma \tfrac{t^k}{\rho} f'_w(x^k)^t d^k, \tag{2.23}$$
   où la relation (2.23) est due à la règle d'Armijo et à l'inégalité $t^k/\rho > t^k$.

   Vu que la suite $\{f_w(x^k)\}$ est non négative et décroissante, on a
   $$t^k f'_w(x^k)^t d^k \to 0.$$
   Grâce à cette dernière relation et à l'inégalité (2.21), on sait que t^k converge vers 0 lorsque k converge vers $+\infty$.
   En divisant les deux membres de (2.23) par $\frac{t^k}{\rho}$ et en passant à la limite, on obtient
   $$f'_w(\bar{x})^t \bar{d} > \sigma f'_w(\bar{x})^t \bar{d}.$$
   Vu que $0 < \sigma < 1$, on a
   $$f'_w(\bar{x})^t \bar{d} > 0,$$

ce qui contredit l'inégalité (2.21). Il suit que tout point d'accumulation de $\{x^k\}$ est une solution de (2.10). De plus, vu que $\bar{F}(x)$ est fortement monotone sur l'ensemble Ω_{int}, cette solution est unique. On en conclut que la suite entière converge vers l'unique solution du problème (2.10).

$\square$

Nous venons de démontrer que l'algorithme 2.1. est convergent. Cependant, nous ne pouvons rien conclure sur sa vitesse de convergence. Nous allons définir dans les paragraphes suivants un nouvel algorithme dont la convergence est quadratique. Pour cela, il suffira d'ajouter une hypothèse supplémentaire. Mais avant d'entrer dans les détails, nous devons définir la condition de centrage.

2.3.5 La condition de centrage

Comme on vient de le mentionner, cette condition nous servira pour définir le nouvel algorithme. Elle nous servira également lors de l'élaboration de l'algorithme pour résoudre le problème $\mathrm{VI}(F, \Omega)$.

Soit la fonction

$$\phi_w(x, y) = w - X[F(x) - A^t y].$$

On peut facilement montrer que si le problème $\mathrm{VI}(\bar{F}, \Omega_{int})$ est résolu exactement, alors $\|\phi_w(x(w), y)\| = 0$, où y est le vecteur dual obtenu en résolvant le système de (KKT). C'est pourquoi, on peut employer $\|\phi_w(x, y)\|$ pour vérifier si un vecteur x est proche de la solution $x(w)$.

Définition

Soient $0 < \beta < 1$ et $w > 0$.
On dit que le couple (x, y) vérifie la condition de centrage si

$$\frac{\|\phi_w(x, y)\|}{\min(w)} \leq \beta.$$

Dans ce cas, on dit également que (x, y) se trouve sur la trajectoire-β.

Lors de l'élaboration de l'algorithme pour résoudre $\mathrm{VI}(F, \Omega)$, il faudra vérfier si l'itéré courant vérifie la condition de centrage. Voici la définition qui sera employée :

On dira que le couple (x^k, y^k) vérifie la condition de centrage si

$$\frac{\|\phi_{w^k}(x^k, y^k)\|}{\min(w^k)} \leq \beta^k, \quad \beta^k = \min\{\beta, \min(w^k)\}, \ 0 < \beta < 1.$$

On imposera alors que

$$\lim_{k\to\infty} w^k = 0.$$

Dans ce cas on dit aussi que (x^k, y^k) est sur la trajectoire-β^k.

2.3.6 Lemmes préliminaires

Dans ce paragraphe, nous allons établir trois résultats qui vont permettre de démontrer la convergence quadratique de ce nouvel algorithme. Pour cela, il va falloir ajouter une nouvelle hypothèse aux cinq hypothèses de départ.

Définition: F est "Scaled Lipschitz Continuous" (SLC).

On dit que la fonction F vérifie la propriété (SLC) sur Ω s'il existe une constante positive $\tau > 0$ telle que

$$\|X(F(\hat{x}) - F(x) - F'(x)(\hat{x} - x))\| \leq \tau(\hat{x} - x)^t F'(x)(\hat{x} - x), \quad \forall \hat{x}, x \in \Omega,$$

où $\|X^{-1}(\hat{x} - x)\| < 1$.

Dans ce paragraphe, nous supposerons que F vérifie cette propriété et nous allons l'ajouter aux cinq hypothèses de départ :

(A6) F vérifie la propriété (SLC).

Démontrons maintenant les trois propriétés qui nous serviront pour définir le nouvel algorithme.

Lemme 2.3.2

Soient $x \in \Omega_{int}$ et $y \in I\!R^m$. On a

$$\|X^{-1}\Delta x\| \leq \frac{\|\phi_w(x, y)\|}{min(w)}.$$

preuve

Considérons les égalités suivantes :

$$\begin{aligned} \phi_w(x,y) &= w - X[F(x) - A^t y] \\ &= X[F'(x)^t + X^{-2}W]\Delta x - XA^t\Delta y. \end{aligned}$$

où la première égalité est due à la définition de ϕ_w et la deuxième est obtenue en considérant la première relation du système (2.20).
En multipliant la dernière relation par $(X^{-1}\Delta x)^t$, on obtient :

$$(X^{-1}\Delta x)^t \phi_w(x,y) = \Delta x^t [F'(x)^t + X^{-2}W]\Delta x, \tag{2.24}$$

où l'on a employé l'égalité $A\Delta x = 0$. Dès lors,

$$\begin{aligned} \|X^{-1}\Delta x\|^2 &= \Delta x^t X^{-2} \Delta x \\ &\leq \frac{1}{\min(w)} \Delta x^t X^{-2} W \Delta x \\ &\leq \frac{1}{\min(w)} \Delta x^t [F'(x)^t + X^{-2}W]\Delta x \qquad (2.25) \\ &= \frac{1}{\min(w)} (X^{-1}\Delta x^t)^t \phi_w(x,y) \\ &\leq \frac{\|X^{-1}\Delta x\| \ \|\phi_w(x,y)\|}{\min(w)} \qquad (2.26) \end{aligned}$$

où la deuxième inégalité vient du fait que F est fortement monotone et la deuxième égalité suit de la relation (2.24).

En divisant par $\|X^{-1}\Delta x\|$ les deux membres de l'inégalité, on obtient :

$$\|X^{-1}\Delta x\| \leq \frac{\|\phi_w(x,y)\|}{\min(w)}.$$

□

Conséquence

Grâce à ce résultat, on sait que, si la condition de centrage est vérifiée alors $\hat{x}$ appartient à l'ensemble Ω_{int}.

En effet, si la condition de centrage est vérifiée, on a

$$\|X^{-1}\Delta x\| \leq \frac{\|\phi_w(x,y)\|}{\min(w)} \leq \beta < 1.$$

Ceci implique que

$$-1 < \frac{(\Delta x)_i}{x_i} < 1 \quad \forall i = 1...n,$$

d'où

$$x_i + \Delta x_i = \hat{x}_i > 0.$$

Par la deuxième relation du système (2.20), on a finalement que $\hat{x} \in \Omega_{int}$.

Lemme 2.3.3

Soient $x \in \Omega_{int}$, $y \in I\!R^m$, $0 < \beta < 1$ *et* $\tau > 0$.

Si $\frac{\|\phi_w(x,y)\|}{\min(w)} \leq \beta$, *alors on a*

$$\|\phi_w(\hat{x},\hat{y})\| \leq max\{(1+\beta)\tau, 1\}\Delta x^t(F'(x) + WX^{-2})\Delta x.$$

preuve

Il suit de la définition de ϕ_w que

$$\|\phi_w(\hat{x},\hat{y})\| = \|\hat{X}(F(\hat{x}) - A^t\hat{y}) - w\|. \tag{2.27}$$

La preuve sera divisée en deux parties. Nous allons d'abord démontrer quelques résultats assez techniques qui nous serviront à développer, dans la deuxième partie de la preuve, le membre de droite de l'équation (2.27).

1. Commençons donc par démontrer les quelques résultats suivants.

   (a) Par la première égalité du système (2.20), on a que

$$F(\hat{x}) - A^t\hat{y} = F(\hat{x}) - F(x) + X^{-1}w - [F'(x)^t + X^{-2}W]\Delta x.$$

(b) Soit D, la matrice diagonale dont les élèments sont égaux à $(X^{-1}\Delta x)_i$.
On va montrer que $De = X^{-1}\Delta x$ et $\hat{X} = X(I + D)$.

En effet, par définition de D, on a que

$$De = \begin{pmatrix} \frac{\hat{x}_1 - x_1}{x_1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & \frac{\hat{x}_n - x_n}{x_n} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\hat{x}_1 - x_1}{x_1} \\ \vdots \\ \vdots \\ \frac{\hat{x}_n - x_n}{x_n} \end{pmatrix} = X^{-1}\Delta x$$

La deuxième égalité se montre facilement en constatant que

$$X(I+D) = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & \dots & x_n \end{pmatrix} \begin{pmatrix} \frac{\hat{x}_1}{x_1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \frac{\hat{x}_n}{x_n} \end{pmatrix} = \mathrm{diag}(\hat{x}_1, ..., \hat{x}_n)$$

(c) $\|D^2 e\| \leq \Delta x^t X^{-2} \Delta x$.

En effet, $\|D^2 e\|^2 = (\frac{\Delta x_1}{x_1})^4 + ... + (\frac{\Delta x_n}{x_n})^4$,

et $(\Delta x^t X^{-2} \Delta x)^2 = ((\frac{\Delta x_1}{x_1})^2 + ... + (\frac{\Delta x_n}{x_n})^2)^2$.

2. Développons maintenant le membre de droite de la relation (2.27)

$$\begin{aligned} \|\phi_w(\hat{x}, \hat{y})\| &= \|X(I+D)(F(\hat{x}) - F(x) - F'(x)\Delta x) + (I+D)W(e - X^{-1}\Delta x) - w\| \\ &\leq \|X(I+D)(F(\hat{x}) - F(x) - F'(x)\Delta x)\| \\ &\quad + \|W[(I+D)(e - X^{-1}\Delta x) - e]\| \end{aligned}$$

Or, par l'hypothèse (A6), on sait que

$$\|X(F(\hat{x}) - F(x) - F'(x)\Delta x)\| \leq \tau(\Delta x^t F'(x) \Delta x).$$

De plus, vu que $De = X^{-1}\Delta x$, on a que

$$(I + D)(e - X^{-1}\Delta x) = e - D^2 e.$$

Finalement, on obtient

$$
\begin{aligned}
\|\phi_w(\hat{x},\hat{y})\| &\leq \tau\|(I+D)\|\Delta x^t F'(x)\Delta x + \|WD^2e\| \\
&\leq (1+\beta)\tau\Delta x^t(F'(x)+X^{-2}W)\Delta x - (1+\beta)\tau\Delta x^t X^{-2}W\Delta x + \|WD^2e\| \\
&\leq (1+\beta)\tau\Delta x^t(F'(x)+X^{-2}W)\Delta x \\
&\quad + \max\{0, -(1+\beta)\tau+1\}\Delta x^t X^{-2}W\Delta x \\
&\leq ((1+\beta)\tau + \max\{0, -(1+\beta)\tau+1\})\Delta x^t(F'(x)+WX^{-2})\Delta x \\
&= \max\{(1+\beta)\tau, 1\}\Delta x^t(F'(x)+WX^{-2})\Delta x.
\end{aligned}
$$

□

Lemme 2.3.4

1. *Soient* $w > 0$, $x \in \Omega_{int}$, $y \in I\!R^m$ *et* $0 < \beta < 1$ *tels que*

$$\|\phi_w(x,y)\|/min(w) \leq \beta.$$

Le problème (2.2) peut être résolu avec un taux de convergence quadratique, c'est-à-dire

$$\|\phi_w(\hat{x},\hat{y})\| \leq \gamma\|\phi_w(x,y)\|^2,$$

si la condition suivante est satisfaite :

$$\|\phi_w(x,y)\| \leq \frac{1}{\gamma}, \tag{2.28}$$

où $\gamma = \frac{\max\{(1+\beta)\tau,1\}}{\min(w)}$ $\tau > 0, w > 0,$

2. *Si* (x,y) *vérifie la relation (2.28) alors*
   (a) $(\hat{x},\hat{y})$ *vérifie également cette relation ;*
   (b) $\hat{x} \in \Omega_{int}$.

preuve

1. Pour démontrer le premier résultat, il suffit de considérer les inégalités suivantes :

$$\begin{aligned} \|\phi_w(\hat{x},\hat{y})\| &\leq \gamma \ \min(w)\Delta x^t(F'(x) + WX^{-2})\Delta x \\ &\leq \gamma \ \min(w)\|X^{-1}\Delta x\|\|\phi_w(x,y)\| \\ &\leq \gamma\|\phi_w(x,y)\|^2, \end{aligned}$$

   où la première inégalité vient du Lemme 2.3.3, la deuxième de la relation entre (2.25) et (2.26) et la dernière résulte du Lemme 2.3.2.

2. (a) Montrons d'abord que $(\hat{x},\hat{y})$ vérifie la relation (2.28).
   Soit (x,y) vérifiant l'inégalité (2.28). On a par le point 1

$$\frac{\|\phi_w(\hat{x},\hat{y})\|}{\|\phi_w(x,y)\|} \leq \gamma\|\phi_w(x,y)\| < 1.$$

   On a alors le résultat voulu, c'est-à-dire

$$\|\phi_w(\hat{x},\hat{y})\| \leq \|\phi_w(x,y)\| \leq \frac{1}{\gamma}.$$

   (b) Ce dernier résultat est une conséquence du Lemme 2.3.2.

□

2.3.7 Deuxième algorithme pour résoudre VI($\bar{F},\Omega_{int}$) : l'algorithme 2.2

Etablissons le nouvel algorithme. La différence entre cet algorithme et le précédent réside au pas 3. En effet, après un certain nombre d'itérations, on va poser $t^k = 1$. Ceci nous garantira le convergence quadratique de l'algorithme.

Initialisation :

Choisir $\quad x^0 \in \Omega_{int}, \quad 0 < \rho, \beta, \sigma < 1,$

et poser $\quad k = 0, \quad \gamma = \frac{\max\{(1+\beta)\tau,1\}}{\min(w)}.$

<u>Itération k</u> : Effectuer les pas 1 à 4 ci-dessous.

<u>Pas 1:</u> Critère d'arrêt

Si un critère de convergence est satisfait, alors arrêter.

<u>Pas 2</u>: La direction de Newton

Trouver $(\hat{x}, \hat{y})$ en résolvant le système suivant :

$$\begin{cases} F(x^k) - X_k^{-1}w + [F'(x^k)^t + X_k^{-2}W](\hat{x} - x^k) - A^t\hat{y} = 0 \\ A\hat{x} = b. \end{cases}$$

où $X_k \equiv \text{diag}(x_1^k, \dots, x_n^k)$. On obtient alors $d^k = \hat{x} - x^k$.

<u>Pas 3</u>: Vérification de la condition (2.28)

1. Si $\frac{\|\phi_w(\hat{x},\hat{y})\|}{\min(w)} < \beta$ et $\|\phi_w(\hat{x},\hat{y})\| < \frac{1}{\gamma}$ alors poser $t^k = 1$;
2. Sinon, prendre t^k, le plus grand nombre dans $\{1, \rho, \rho^2, ...\}$ tel que :

$$f_w(x^k + t^k d^k) \leq f_w(x^k) + \sigma t^k f_w'(x^k)^t d^k, \qquad \text{[règle d'Armijo]}$$
$$x^k + t^k d^k \in \Omega_{int}.$$

<u>Pas 4</u>: Mise à jour

Poser $x^{k+1} = x^k + t^k d^k$, $y^{k+1} = \hat{y}$, $k = k + 1$.

Théorème 2.3

Il existe un entier $\bar{k}$ dépendant de w et de n tel que pour tout $k \geq \bar{k}$, la suite $\{x^k\}$ engendrée par l'algorithme 2.2 converge quadratiquement vers la solution de $VI(\bar{F}, \Omega_{int})$ avec $t^k = 1$.

preuve

Supposons par l'absurde qu'il n'existe pas de $\bar{k}$ tel que la condition (2.28) soit satisfaite. Dans ce cas, l'algorithme 2.2 se réduit à l'algorithme 2.1. Par le théorème de convergence de l'algorithme 2.1, l'algorithme 2.2 engendre une suite qui converge vers la solution $x(w)$ tel que $\|\phi_w(x^k, y^k)\|$ converge vers 0 lorsque k converge vers $+\infty$. On obtient ainsi la contradiction voulue. Dès lors, $\bar{k}$ existe.

Selon le Lemme 2.3.4, pour tout $k \geq \bar{k}$, c'est-à-dire après que la condition (2.28) soit satisfaite, $\|\phi_w(x^k, y^k)\|$ converge quadratiquement vers 0 avec $x^k \in \Omega_{int}$ pour tout $k \geq \bar{k}$, ce qui implique que la suite $\{x^k\}$ converge vers la solution $x(w)$.

□

2.4 Algorithme pour résoudre VI($F\Omega$)

L'idée que nous allons suivre pour établir cet algorithme est la suivante : à chaque itération, nous vérifierons si l'itéré courant satisfait la condition de centrage. Si c'est le cas, on passe à l'itération suivante après avoir multiplié w par un facteur θw, $0 < \theta < 1$. Sinon, nous appliquons l'algorithme 2.1 pour "centrer" l'itéré. Si l'hypotothèse (A6) est vérifiée, on peut évidemment remplacer l'algorithme 2.1 par l'algorithme 2.2.

Avant d'établir l'algorithme, certaines considérations sont nécéssaires afin d'en faciliter la compréhension.

Soient K l'ensemble des indices de tous les itérés engendrés par l'algorithme 2.4, $x^0 \in \Omega_{int}$, $y^0 \in I\!R^m$ et $S_K = \{(x^k, y^k)\}_{k \in K}$. A chaque itération k, nous allons résoudre le système suivant pour trouver le nouvel itéré $\hat{x}$:

$$\begin{cases} F(x^k) - X_k^{-1} w^k + [F'(x^k)^t + X_k^{-2} W^k](\hat{x} - x^k) - A^t \hat{y} = 0 \\ A\hat{x} = b. \end{cases}$$

où $X_k \equiv \mathrm{diag}(x_1^k, \dots, x_n^k)$ et $W^k \equiv \mathrm{diag}(w_1^k, \dots, w_n^k)$.

Définissons deux ensembles S_{K^1} et S_{K^2} comme suit :

$$S_K = S_{K^1} \cup S_{K^2} \quad \text{et} \quad S_{K^1} \cap S_{K^2} = \emptyset.$$

où l'ensemble S_{K^1} est composé des indices des itérés qui vérifient la condition de centrage et qui sont strictement admissibles.

Donc, $\forall k \in S_{K^1}$, $0 < \beta^k < 1$, $w^k > 0$, on a

1. $\|\phi_{w^k}(x^k, y^k)\| / \min(w^k) \leq \beta_k$;
2. $x^k \in \Omega_{int}$.

A chaque indice k^1 dans S_{K^1}, nous ferons correspondre un sous-ensemble de S_{K^2} que l'on notera $S_{K^2_{k^1}}$. Cet ensemble peut être vide.

La fonction g dont il est question lors du critère d'arrêt est celle définie au premier chapitre par la relation suivante:

$$g(x) = \max\{\langle F(x), x - y\rangle \mid y \in \Omega\}.$$

C'est une fonction de mérite pour le problème VI(F, Ω).

2.4.1 L'algorithme 2.4

<u>**Initialisation**</u>

Choisir $x^0 \in \Omega_{int}$, $0 < \theta, \beta, \sigma, \rho < 1$, $\varepsilon > 0$,
w^0 tel que $\min(w^0) > 0$, B^0 tel que $\|B^0\| \leq 2\frac{\min(w^0)}{\gamma^2}$;
et poser $k = 0$, $k^1 = 0$, $K = \emptyset$, $K^1 = \emptyset$, $K^2_{k^1} = \emptyset$.

<u>**Itération k**</u> Effectuer les pas 1 à 4 ci-dessous.

<u>**Pas 1**</u>**: Critère d'arrêt**

Si un critère de convergence est satisfait, alors arrêter.
Ce critère d'arrêt peut être

$$g(x^k) \leq \varepsilon.$$

<u>**Pas 2**</u>**: La direction de Newton**

Trouver $(\hat{x},\hat{y})$ en résolvant le système suivant:

$$\begin{cases} F(x^k) - X_k^{-1}w^k + [F'(x^k)^t + X_k^{-2}W^k](\hat{x} - x^k) - A^t\hat{y} = 0 \\ A\hat{x} = b. \end{cases}$$

<u>**Pas 3**</u>**: Vérification de la condition de centrage**

Si $\frac{\|\phi_{\hat{w}}(\hat{x},\hat{y})\|}{\min(\hat{w})} \leq \hat{\beta}$ où $\hat{w} = w^k$, $\hat{\beta} = \min\{\beta, \min(\hat{w})\}$, $\hat{x} \in \Omega_{int}$,

alors aller au pas 3a,
sinon, aller au pas 3b.

<u>**Pas 3a**</u>**: Mise à jour**

Poser $x^{k+1} = \hat{x}$, $w^{k+1} = \theta\hat{w}^2$
Choisir B^{k+1} tel que $\|B^{k+1}\| \leq 2\frac{\min(w^{k+1})}{\gamma^2}$.
Poser $K^1 = K^1 \cup \{k+1\}$, $k^1 = k^1 + 1$.
Aller au pas 4.

Pas 3b: Recherche linéaire

Poser $d^k = \hat{x} - x^k$ et prendre t^k le plus grand nombre de l'ensemble $\{1, \rho, \rho^2, ...\}$ tel que

$$\begin{cases} f_w(x^k + t^k d^k) \leq f_w(x^k) + \sigma t^k f'_w(x^k)^t d^k, \\ x^k + t^k d^k \in \Omega_{int}. \end{cases}$$

Poser $x^{k+1} = x^k + t^k d^k$, $w^{k+1} = w^k$,
$K^2_{k^1} = K^2_{k^1} \cup \{k+1\}, K^2 = K^2 \cup \{k+1\}$.
Aller au pas 4.

Pas 4: Mise à jour

Poser $y^{k+1} = \hat{y}$, $K = K \cup \{k+1\}$, $k = k+1$.

L'ensemble S_K est engendré lors des pas 1 à 4. L'ensemble S_{K^1} est engendré lors des pas 1, 2, 3a et 4. L'ensemble S_{K^2} est engendré lors des pas 1, 2, 3b et 4. Pour chaque valeur de k^1, la valeur de w reste la même et le sous-ensemble $S_{K^2_{k^1}}$ est exactement engendré par l'algorithme 2.1 avec comme critère de convergence celui indiqué au pas 3. On va donc s'arrêter lorsque la condition de centrage sera satisfaite et dans ce cas, on change la valeur de w. On ne résout donc pas le problème (2.2) tant que la condition de centrage est satisfaite. Si elle n'est pas satisfaite, on va appliquer l'algorithme 2.1. pour "centrer" l'itéré courant.

2.4.2 La convergence de l'algorithme 2.4

Dans ce paragraphe, nous établissons le théorème de convergence de l'algorithme 2.4. Auparavant, nous démontrerons un lemme technique qui nous servira dans la démonstration du théorème.

Lemme 2.4.1

Si $x(w)$ est l'unique solution du problème $VI(\bar{F}, \Omega_{int})$, alors la relation suivante est vérifiée:

$$\langle F(x(w)), x - x(w) \rangle \geq n \frac{\min(w)(\Gamma - \gamma)}{\gamma}, \quad \forall x \in \Omega_{int} \tag{2.29}$$

preuve

Vu que $x(w)$ est la solution de VI($\bar{F}, \Omega_{int}$) et par définition de $\bar{F}$, on a pour tout $x \in \Omega_{int}$:

$$\langle F(x(w)) - X_w^{-1}w, x - x(w)\rangle \geq 0,$$

et par conséquent

$$\langle F(x(w)), x - x(w)\rangle \geq \langle X_w^{-1}w, x - x(w)\rangle, \tag{2.30}$$

où $X_w^{-1} = \text{diag}\{\frac{1}{x_1(w)}, ..., \frac{1}{x_n(w)}\}$.

Développons le membre de droite de l'équation (2.30).

$$\begin{aligned}
\langle X_w^{-1}w, x - x(w)\rangle &= \begin{bmatrix} \frac{w_1}{x_1(w)} & \cdots & \frac{w_n}{x_n(w)} \end{bmatrix} \begin{bmatrix} x_1 - x_1(w) \\ \vdots \\ x_n - x_n(w) \end{bmatrix} \\
&= \frac{w_1(x_1 - x_1(w))}{x_1(w)} + \ldots + \frac{w_n(x_n - x_n(w))}{x_n(w)} \\
&\geq \min(w)\left(\frac{(x_1 - x_1(w))}{x_1(w)} + \ldots + \frac{(x_n - x_n(w))}{x_n(w)}\right) \\
&\geq n\frac{\min(w)(\Gamma - \gamma)}{\gamma},
\end{aligned}$$

d'où la thèse.

□

Conséquence
Nous savons déjà que, si le vecteur w converge vers 0 alors $\bar{F}$ converge vers F. Vu que Ω est un ensemble compact, il existe une sous-suite de $\{x(w)\}_{w>0}$ qui converge vers $x(0)$ lorsque w converge vers 0. Par la relation (2.29), cette limite est une solution du problème VI(F, Ω).

Théorème 2.4 Convergence de l'algorithme 2.4.

Supposons que les hypothèses (A1)-(A5) sont vérifiées et que pour tout k, la matrice B^k est telle que $\|B^k\| \leq 2\frac{min(w^k)}{\gamma^2}$.
Dès lors, la sous-suite $\{x^k\}_{k\in K^1}$ engendrée par l'algorithme 2.4 appartient à l'ensemble Ω_{int} et converge vers une solution de VI(F, Ω) pour n'importe quel itéré initial $x^0 \in \Omega_{int}$.

preuve

Cette preuve se divise en deux parties : tout d'abord, nous montrons que l'ensemble $K^2_{k^1}$ est fini et ensuite, nous montrerons que $x^k \in \Omega_{int}$ pour tout $k \in K^1$ et que la sous-suite $\{x(w^k)\}_{k\in K^1}$ converge vers une solution de VI(F, Ω).

1. Démontrons par l'absurde que $K^2_{k^1}$ est un ensemble fini.
   Comme on l'a dit précédemment, w^k reste identique pour un k^1 fixé. En effet, si la condition de centrage n'est pas satisfaite on ne change ni k^1, ni la valeur de w^k .
   Donc , la sous-suite correspondant à k^1 est exactement celle engendrée par l'agorithme 2.1 avec w^k et le critère d'arrêt est la condition de centrage. Par le Théorème 2.2, si $K^2_{k^1}$ est infini, alors la sous-suite $\{x^k\}_{k\in K^2_{k^1}}$ converge vers l'unique solution du problème (2.2), ce qui implique que
   $$\lim_{k\in K^2_{k^1}\to\infty} \|\phi_{w^k}(x^k, y^k)\| = 0. \tag{2.31}$$
   Prenons $\varepsilon_k = \min(w^k)\beta_k$. Par l'égalité précédente, on a qu'il existe $\bar{k}$ tel que
   $$\|\phi_{w^k}(x^k, y^k)\| \leq \varepsilon_k \quad \forall k \geq \bar{k}, \quad k \in K^2_{k^1}.$$
   On a donc que $\forall k \geq \bar{k}$, (x^k, y^k) vérifie la condition de centrage. Or ceci est absurde car, par définition, $K^2_{k^1}$ ne contient pas les indices des itérés qui vérifient la condition de centrage. Dès lors, $K^2_{k^1}$ est fini.

2. Grâce au Lemme 2.3.2 et à la définition de l'algorithme 4.1, on sait que la suite des itérés $\{x^k\}_{k\in K^1}$ appartient à l'ensemble Ω_{int}. Montrons que la sous-suite $\{x^k\}_{k\in K^1}$ converge vers une solution de VI(F, Ω).
   Par le pas 3 de l'algorithme 2.4, si x^k satisfait la condition de centrage, alors l'indice k est ajouté à K^1; sinon, on obtient une sous-suite engendrée par l'algorithme 2.1 avec comme point initial x^k. Comme il a été démontré au point 1, la suite va atteindre un point qui va satisfaire la condition de centrage après un nombre fini d'itérations.
   Vu que Ω est compact, par le Lemme 2.4.1, on sait qu'il existe une sous-suite de $\{x(w)\}_{w>0}$ convergeant vers une solution x^* de VI(F, Ω) lorsque w tend vers 0. De plus, nous avons que
   $$\lim_{k\in K^1\to\infty} \beta_k = 0,$$
   et par conséquent,
   $$\lim_{k\in K^1\to\infty} \|\phi_{w^k}(x^k, y^k)\| = \|\phi_{w^\infty}(x^\infty, y^\infty)\| = 0.$$

Vu que la sous-suite $\{x^k\}_{k \in K^1}$ appartient à l'ensemble compact Ω, il existe une sous-suite de $\{x^k\}_{k \in K^1}$ convergeant vers x^∞ solution du sous-problème $\mathrm{VI}(\bar{F}, \Omega_{int})$ où w^∞ vaut 0. Il suit que $x(w^\infty) = x(0) = x^\infty$ est une solution de $\mathrm{VI}(F, \Omega)$.

$\square$

Chapitre 3

Les fonctions et les opérateurs self-concordants

Dans le Chapitre 2, nous avons élaboré un algorithme convergent pour le problème VI(F,Ω) où $\Omega = \{x| \ Ax = b, \ x \geq 0\}$, A étant une matrice (m,n) et b un vecteur de $I\!R^m$. Dans le Chapitre 4, nous allons étendre cet ensemble en remplaçant les contraintes de non négativité par des contraintes convexes. L'ensemble Ω sera donc de la forme

$$\Omega = \{x|Ax = b, \ f_i(x) \leq 0, \ i = 1, ..., m\}$$

où $f_i : I\!R^n \to I\!R$ sont des fonctions convexes, A est une matrice (p,n) et b est un vecteur de $I\!R^p$. Evidemment, des hypothèses supplémentaires vont être nécessaires pour prouver la convergence de l'algorithme. Dans ce chapitre, nous allons motiver l'introduction des nouvelles propriétés et en particulier, nous allons définir la notion de self-concordance.

3.1 Motivation

Dans le Chapitre 4, nous utiliserons la méthode barrière pour résoudre l'inéquation variationnelle et la fonction barrière sera self-concordante. Dans de ce paragraphe, nous expliquons l'origine de la définition de self-concordance.

Lorsque l'on veut résoudre un problème de programmation convexe par une méthode de points intérieurs, deux résultats principaux sont requis pour que l'algorithme soit polynomial.

Tout d'abord, si l'on applique la méthode de Newton à un sous-problème barrière et si le premier itéré est proche de la solution, il faut que le nombre

d'itérations nécessaires pour trouver approximativement cette solution est borné. Ensuite, si le paramètre de pénalité de la fonction barrière ne varie pas trop rapidement, il faut que la solution approximative d'un sous-problème ne soit pas trop "éloignée" de la solution du sous-problème suivant.
Ces deux résultats peuvent sembler évidents et pourtant, ils ne le sont pas.

En effet, en général en programmation convexe, la solution se trouve sur la frontière du domaine admissible. Or, la fonction barrière est singulière sur la frontière du domaine de sorte que la matrice hessienne est de plus en plus malconditionnée [1] au fur et à mesure que l'on s'approche de la solution.
De plus, en examinant le résultat suivant, on s'aperçoit que la constante de convergence c dépend du conditionnement de la matrice hessienne.

Considérons un sous-problème barrière du type $min\{tc^t x + F(x) | \ x \in I\!R^n\}$ *dont la solution est* x^* *et supposons que*

1. F *est une fonction convexe de classe* C^2 *sur une boule euclidienne* V *centrée en* x^* *et de rayon* r,
2. $|F''(x) - F''(x^*)| \leq L|x - x^*|, \quad x \in V$,
3. $F''(x^*)$*est non dégénéré et* $0 < \lambda_{min} < \lambda_{max}$.

On peut alors montrer les résultats suivants:

1. $|x^+ - x^*| \leq c|x - x^*|$ *où* x^+ *est l'itéré obtenu par la méthode de Newton;*
2. c *dépend du conditionnement de la matrice hessienne* $F''(x^*)$, *de* L *et de* r.

Pour analyser le comportement de la méthode de Newton dans ce cas, il faut en quelque sorte prendre en compte ces singularités. Pour cela, il serait utile de définir une norme $\| \ . \ \|_{x,F}$ en terme de la matrice hessienne de la fonction barrière F évaluée en un point x.
Vu que cette norme dépend de la matrice hessienne de la fonction barrière, elle varie lorsque les variables changent. Si la matrice hessienne varie rapidement, il se peut qu'il soit impossible d'employer les valeurs de $\|x - x^*\|_{x,F}$ pour tirer des conclusions sur la convergence de la méthode. Il faut donc que la matrice hessienne ne varie pas de trop. Cela nous amène à imposer une borne à la troisième dérivée de F en terme de la matrice hessienne. L'existence de cette borne permet de prouver le premier résultat.

1. Le conditionnement d'une matrice est le rapport en valeur absolue entre sa plus grande et sa plus petite valeur propre.

Pour prouver le deuxième résultat, c'est-à-dire que la solution approximative d'un sous-problème ne sera pas trop éloignée de la solution du sous-problème suivant, il est nécessaire que la valeur de la fonction barrière F ne varie pas trop rapidement lorsque le paramètre de pénalité varie. Pour cela, nous imposerons une borne sur la dérivée première de F en fonction de la matrice hessienne.
Si la fonction barrière possède de telles propriétés, alors il existe une méthode de points intérieurs de complexité polynomiale qui résout le problème initial.

3.2 Définitions et exemples

Lors de la résolution des sous-problèmes VI$(\bar{F}, \Omega)$, nous avons utilisé la méthode de Newton. Si l'on applique cette méthode à une fonction quadratique, alors la méthode converge en une itération. Si l'on étend ce résultat, on peut dire que, si la matrice hessienne ne varie pas trop rapidement, alors la méthode de Newton converge rapidement. On peut donc dire que le rayon de convergence de la méthode de Newton pour minimiser une fonction F est inversément proportionnel à la "non-linéarité" de F. La méthode de Newton donne de bons résultats si une petite variation en x entraîne une petite variation de la dérivée seconde de F. Or on peut mesurer la variation de la dérivée seconde grâce à la dérivée troisième. Intuitivement, la dérivée troisième devrait être petite par rapport à la dérivée seconde. La self-concordance reflète cette propriété.

3.2.1 Les fonctions self-concordantes

Dans ce chapitre, nous supposons que E est un espace vectoriel réel de dimension finie, Q un sous-ensemble ouvert, convexe, non vide de E, et $G \subseteq E$ un ensemble convexe, fermé, à intérieur non vide (l'intérieur de G sera noté G_{int}). Les définitions ci-dessous sont données en toute généralité. Cependant, dans le chapitre suivant, nous considérerons uniquement le cas où $E = \mathbb{R}^n$.

Le cas unidimensionnel

A une dimension, on dit qu'une fonction convexe f : Domf $\rightarrow \mathbb{R}$ est **self-concordante** si elle vérifie l'inégalité suivante :

$$|f'''(x)| \leq 2f''(x)^{3/2} \quad \forall x \in \text{int}\{\text{domf}\}.$$

Un exemple trivial est la fonction barrière logarithmique $f(x) = -\log(x)$ définie pour tout $x > 0$. En effet, $f''(x) = \frac{1}{x^2}$ et $f'''(x) = -\frac{2}{x^3}$. L'inégalité ci-dessus est dès lors immédiatement satisfaite.

Dans la définition de la self-concordance, la constante 2 du second membre n'a pas beaucoup d'importance. On peut la remplacer par n'importe quelle constante C. La définition peut encore s'écrire :

$$|f'''(x)| \leq Cf''(x)^{3/2} \quad \forall x \in \text{int}\{\text{domf}\}.$$

Le cas multi-dimensionnel

Dans le cas multi-dimensionnel, une fonction est self-concordante si sa restriction à chaque droite intersectant son domaine est self-concordante. Donnons une définition plus précise.

Considérons $F : Q \to I\!R$ une fonction convexe, de classe C^3.

On dit que **F est (a-)self-concordante sur Q** si la propriété suivante est vérifiée :

$$|F'''(x)[h,h,h]| \leq 2a^{-1/2}\langle F''(x)h, h\rangle^{3/2} \quad \forall x \in Q \quad \text{et} \quad \forall h \in E. \tag{3.1}$$

On dit que **F est fortement (a-)self-concordante sur Q** si pour toute suite (x_k) convergeant vers la frontière de Q, $F(x_k)$ converge vers $+\infty$.

Exemple

Reprenons comme exemple la fonction barrière logarithmique. Elle est définie sur l'ensemble $\Omega = \{x \mid a_i^t x - b_i \geq 0,\ i = 1, ..., m\}$ par la relation suivante :

$$F(x) = -\sum_{i=1}^{m} \log(a_i^t x - b_i),$$

où a_i est un vecteur de $I\!R^n$ et $b_i \in I\!R$, pour tout $i \in \{1, ..., m\}$.
On peut montrer que cette fonction est self-concordante sur l'ensemble Ω.

La semi-norme induite par le hessien

Comme nous l'avons déjà mentionné dans le premier paragraphe, nous allons définir une norme en fonction de la matrice hessienne de F en un point x.

Soient $x \in Q$ et $F''(x)$ la matrice hessienne de F en un point x. La semi-norme induite par le hessien est définie par la relation suivante:

$$\|h\|_{x,F}^2 = h^t F''(x)h. \tag{3.2}$$

Si la matrice $F''(x)$ est régulière pour tout x dans l'ensemble Q alors $\| \ . \ \|_{x,F}$ est une norme.

3.2.2 Les fonctions barrières ν-self-concordantes

Prenons $\nu \geq 0$. On dit qu'une fonction $F : G_{int} \to I\!R$ est **une fonction barrière ν-self-concordante** sur G si F est 1-fortement self-concordante sur G_{int} et si pour tout $h \in E$ et pour $x \in G_{int}$, la relation suivante est vérifiée:

$$|\langle F'(x), h\rangle| \leq \nu^{1/2}\|h\|_{x,F}. \tag{3.3}$$

ν est le paramètre de la fonction barrière F.

Propriétés

Les résultats cités dans cette section sont tirés du deuxième chapitre du livre de Nesterov et Nemirovskii [2].

Soient F une fonction barrière ν-self-concordante sur l'ensemble G, $x \in G_{int}$ et $E_F = \{h \in E | \ \|h\|_{x,F} = 0\}$.

Sous ces conditions, on peut démontrer que

a) *E_F ne dépend pas de x;*

b) *si G est borné, alors F est non dégénérée, c'est-à-dire $E_F \neq \{0\}$, et la semi-norme $\| \ . \ \|_{x,F}$ est une norme.*

Dès lors, si l'ensemble G est borné, la semi-norme définie par la realtion (3.2) est une norme.

On peut également démontrer le résultat suivant

Si F est une fonction barrière ν-self-concordante et $x, y \in G_{int}$ alors la relation suivante est vérifiée:

$$\langle F'(x), y - x \rangle \leq \nu. \tag{3.4}$$

Cette propriété nous sera très utile lors de la démonstration de la convergence de l'algorithme pour résoudre le problème d'inéquations variationnelles défini au chapitre suivant.

3.2.3 Les opérateurs self-concordants

Soient $T : Q \to E^*$ un opérateur monotone de classe C^2 et $x \in Q$. La dérivée première $T'(x)$ définie une forme bilinéaire $\langle T'(x)h, e\rangle \ : \ E \times E \to I\!R$. Comme T est un opérateur monotone, on a que

$$\langle T'(x)h, h\rangle \geq 0 \quad \forall h \in E.$$

Définissons $\hat{T}$, un opérateur symétrique et semi-défini positif par la relation suivante:

$$\hat{T}(x) = [T'(x) + (T')^*(x)]/2,$$

ainsi que le produit scalaire sur E

$$(h, e)_{T,x} = \langle \hat{T}(x)h, e\rangle,$$

et la semi-norme euclidienne

$$\|h\|_{T,x} = \{(h, h)_{T,x}\}^{1/2}.$$

On dit qu'un **opérateur monotone** $T : Q \to E^*$ **est self-concordant** s'il est de classe C^2 et si la relation suivante est satisfaite pour tout $x \in Q$ et pour tout $h_i \in E$, $i = 1, 2, 3$:

$$|T''(x)[h_1, h_2, h_3]| \leq 2 \prod_{i=1}^{3} \|h_i\|_{T,x}. \tag{3.5}$$

L'ensemble des opérateurs self-concordants sur Q est noté $\mathcal{S}(\mathcal{Q})$.

L'opérateur T est fortement self-concordant sur Q si la suite $\hat{T}(x_i)$ converge vers $+\infty$ lorsque la suite $\{x_i\} \in Q$ converge vers un point de la frontière de Q. L'ensemble des opérateurs fortement self-concordants est noté $\mathcal{S}^+(\mathcal{Q})$.

Self-concordance de la dérivée d'une fonction self-concordante

De nouveau, le résultat qui va suivre est extrait du livre de Nesterov et Nemirovskii [2].

Soit f, une fonction self-concordante sur Q. On peut démontrer que l'opérateur $F \equiv f'$ est également self-concordant sur Q.
De plus, si f est une fonction fortement self-concordante sur Q, alors l'opérateur F est fortement self-concordant sur Q.

Notation

Soient $T \in \mathcal{S}(\mathcal{Q})$ et $x \in Q$. Définissons les ensembles suivants :

$E(T,x) = \{h \in E \mid \|h\|_{T,x} = 0\}$,
$W_r(T,x) = \{y \in E \mid \|y-x\|_{T,x} \leq r\}$,
$W_r^0(T,x) = \text{int } W_r(T,x)$.

3.3 Propriétés des opérateurs self-concordants

Proposition 3.2.1

1. *Si T est un opérateur self-concordant sur l'ensemble Q, alors on a :*

   (a) *$E(T,x)$ ne dépend pas de $x \in Q$, c'est-à-dire*
   $$E(T,x) \equiv E(T) \quad \forall x \in Q;$$

   (b) *Si $x \in Q$, alors pour tout $y \in Q \cap W_1(T,x)$ et pour tout $h_1, h_2, h \in E$, les relations suivantes sont vérifiées:*

$$|\langle [T'(y) - T'(x)]h_1, h_2\rangle| \leq \left\{ \frac{1}{[1 - \|y-x\|_{T,x}]^2} - 1 \right\} \|h_1\|_{T,x}\|h_2\|_{T,x}, \tag{3.6}$$

$$\|h\|_{T,x}(1 - \|y-x\|_{T,x}) \leq \|h\|_{T,y} \leq \frac{\|h\|_{T,x}}{(1 - \|y-x\|_{T,x})}. \tag{3.7}$$

2. *Si T est fortement self-concordant sur Q, alors on a également:*
$$W_1^0(T,x) \subset Q, \quad \forall x \in Q. \tag{3.8}$$

preuve

1. Pour démontrer la première partie du théorème, nous aurons besoin du résultat suivant :

   Soit $f : dom f \subseteq I\!R \to I\!R$ une fonction absolument continue telle que

   $$|f'(t)| \leq g(t)|f(t)| \quad \text{où } g \text{ est non nulle.}$$

   Alors, soit $f \equiv 0$ sur domf, soit $f(t) \neq 0 \quad \forall t \in dom f$.

   (a) Démontrons d'abord que $E(T, x)$ est indépendant de x.
   Pour cela, prenons $h \in E(T, x)$ et considérons la fonction suivante :

   $$\phi(y) = \langle \hat{T}(y)h, h\rangle = \langle T'(y)h, h\rangle \ : \ Q \to I\!R.$$

   Cette fonction satisfait la relation $\phi(x) = 0$. Vu que T est self-concordant, on a

   $$|\langle \phi'(y), e\rangle| = |T''(y)[e, h, h]| \leq 2\|e\|_{T,y}\|h\|^2_{T,y} = 2\|e\|_{T,y}\phi(y) \quad \forall y \in Q.$$

   En appliquant le résultat rappelé ci-dessus avec $f(t) = \phi(x + te)$, domf $= [0, 1]$ et $e = y - x$, on peut conclure que $\phi \equiv 0$.
   On a donc montré que $E(T, x) \subset E(T, y)$ $\forall y \in Q$. Vu que x est un élément arbitraire de Q, la dernière relation implique que

   $$E(T, x) \equiv E(T) \quad \forall x \in Q.$$

   (b) Démontrons la relation (3.7).
   Soient $x, y \in Q$ tels que $\|y - x\|_{T,x} < 1$. Considérons la fonction

   $$\phi(t) = \langle T'(x + te)e, e\rangle \ : \ [0, 1] \to I\!R \text{ où } e = y - x.$$

   Cette fonction est de classe C^1 et vérifie la relation suivante :

   $$|\phi'(t)| = |T''(x + te)[e, e, e]| \leq 2\|e\|^3_{T,x+te} \equiv 2\phi^{\frac{3}{2}}(t).$$

   L'inégalité

   $$|\phi'(t)| \leq 2\phi^{3/2}(t), \quad 0 \leq t \leq 1$$

   implique que

   $$\text{soit } \phi(t) \equiv 0, \quad \text{soit } \ |\frac{d}{dt}\phi^{-1/2}(t)| \leq 1, \quad 0 \leq t \leq 1.$$

Supposons que $|\frac{d}{dt}\phi^{-1/2}(t)| \leq 1, \quad 0 \leq t \leq 1$. Dans ce cas, on a que

$$-1 \leq \frac{d}{dt}\phi^{-1/2}(t) \leq 1 \quad 0 \leq t \leq 1.$$

En intégrant cette dernière relation par rapport à t, on obtient :

$$\phi^{-1/2}(t) \geq \phi^{-1/2}(0) - t \tag{3.9}$$

$$\phi^{-1/2}(t) \leq \phi^{-1/2}(0) + t. \tag{3.10}$$

Montrons maintenant que $\frac{\|e\|^2_{T,x}}{(1+t\|e\|_{T,x})^2} \leq \phi(t) \leq \frac{\|e\|^2_{T,x}}{(1-t\|e\|_{T,x})^2} \quad 0 \leq t \leq 1$.

(i) Reprenons la relation (3.9).
Vu que $\phi(0) = \|e\|^2_{T,x} < 1$, on obtient l'inégalité suivante :

$$\phi^{-1/2}(t) \geq \frac{1}{\|e\|_{T,x}} - t = \frac{1 - t\|e\|_{T,x}}{\|e\|_{T,x}}.$$

Par conséquent, on a

$$\phi^{1/2}(t) \leq \frac{\|e\|_{T,x}}{1 - t\|e\|_{T,x}}. \tag{3.11}$$

(ii) Reprenons cette fois-ci la relation (3.10).
Vu que $\phi(0) = \|e\|^2_{T,x}$, on obtient :

$$\phi^{-1/2}(t) \leq t + \frac{1}{\|e\|_{T,x}} = \frac{t\|e\|_{T,x} + 1}{\|e\|_{T,x}}$$

Ce qui nous donne la relation voulue, c'est-à-dire :

$$\frac{\|e\|_{T,x}}{(t\|e\|_{T,x} + 1)} \leq \phi^{1/2}(t).$$

Evidemment, ces deux résultats sont également valables dans le cas où $\phi(t) \equiv 0, \quad 0 \leq t \leq 1$.

Soient $h_i \in E$, $i = 1, 2$ tels que $\|h_i\|_{T,x} \leq 1$.
Définissons la fonction $f_i(t)$ par la relation suivante :

$$f_i(t) = \langle T'(x + te)h_i, h_i\rangle = \|h_i\|^2_{T,x+te} \quad 0 \leq t \leq 1.$$

Si l'on dérive f_i, on obtient :

$$|f_i'(t)| = |T'''(x+te)[e, h_i, h_i]| \leq 2\,\|e\|_{T,x+te}\|h_i\|^2_{T,x+te} = 2\,\|e\|_{T,x+te}f_i(t).$$

De nouveau, en employant, le résultat rappelé au début de la démonstration, on peut conclure que soit $f_i \equiv 0$, soit $f_i > 0$.
Si $f_i > 0$, on a :

$$|\frac{d}{dt}\ln f_i(t)| \leq 2\|e\|_{T,x+te} \leq \frac{2\ \|e\|_{T,x}}{(1-t\|e\|_{T,x})},$$

où la dernière inégalité vient de la relation (3.11). Par conséquent, on obtient

$$-2\frac{\|e\|_{T,x}}{(1-t\|e\|_{T,x})} \leq \frac{d}{dt}\ln f_i(t) \leq 2\frac{\|e\|_{T,x}}{(1-t\|e\|_{T,x})}.$$

En intégrant par rapport à t, on obtient :

$$(1-t\|e\|_{T,x})^2 \leq \frac{f_i(t)}{f_i(0)} \leq \frac{1}{(1-t\|e\|_{T,x})^2} \quad 0 \leq t \leq 1.$$

Vu que $f_i(0) = \|h_i\|^2_{T,x}$, on obtient finalement la relation suivante

$$\|h_i\|_{T,x}(1-t\|e\|_{T,x}) \leq \|h_i\|_{T,x+te} \leq \|h_i\|_{T,x}(1-t\|e\|_{T,x})^{-1} \quad 0 \leq t \leq 1.$$

Lorsque t vaut 1, on obtient la relation (3.7). Cette relation est également valable dans le cas où $f_i \equiv 0$.

(c) Démontrons la relation (3.6).
Pour cela, définissons la fonction f comme suit :

$$f(t) = \langle T'(x+te)h_1, h_2\rangle, \quad 0 \leq t \leq 1.$$

Grâce au résultat obtenu au point (b), on a que

$$\begin{aligned}|f'(t)| &= |T''(x+te)[e,h_1,h_2]| \leq 2\ \|e\|_{T,x+te}\|h_1\|_{T,x+te}\ \|h_2\|_{T,x+te}\\ &\leq 2\ \frac{\|e\|_{T,x}\|h_1\|_{T,x}\|h_2\|_{T,x}}{(1-t\|e\|_{T,x})^3}.\end{aligned}$$

Dès lors, on obtient la relation voulue, c'est-à-dire

$$\begin{aligned}|\langle(T'(y)-T'(x))h_1,h_2\rangle| &= |f(1)-f(0)|\\ &\leq 2\ \|e\|_{T,x}\|h_1\|_{T,x}\|h_2\|_{T,x}\int_0^1 \frac{dt}{(1-t\|e\|_{T,x})^3}\\ &= \left(\frac{1}{(1-\|y-x\|_{T,x})^2}-1\right)\ \|h_1\|_{T,x}\|h_2\|_{T,x}.\end{aligned}$$

2. Démontrons la relation (3.8).

Il suit de la relation (3.7) que l'ensemble des opérateurs $\{\hat{T}(y) \mid y \in W_r^0(T,x) \cap Q\}$ est borné pour tout $r < 1$. Vu que $T \in \mathcal{S}^+(\mathcal{Q})$, on a immédiatement que $W_0^1(T,x) \subset Q$.
En effet, prenons $y \in W_1^0(T,x)$ et montrons que $y \in Q$.
Par l'absurde, supposons que $y \notin Q$. Vu que $x \in Q$ et que Q est un ensemble ouvert, on peut construire une suite de points

$$x_i = x + t_i(y-x) \ \in Q \quad 0 \le t_i < 1$$

qui converge vers un point de la frontière de Q.
Vu que T est fortement self-concordant, la suite $\hat{T}(x_i)$ doit converger vers $+\infty$. Or par la relation (3.7), cette suite est bornée, ce qui nous amène à la contradiction voulue.

□

Chapitre 4

Elaboration d'un algorithme pour résoudre le problème VI(S, Ω)

Comme nous l'avons annoncé précédemment, on étend dans ce chapitre la contrainte de non négativité de l'ensemble Ω à des contraintes convexes et nous introduirons une fonction barrière self-concordante.
Nous établirons, de manière analogue à ce qui a été fait au Chapitre 2, un algorithme pour résoudre le problème d'inéquations variationnelles VI(S, Ω). En effet, nous définirons d'abord une famille de sous-problèmes barrières associée à VI(S, Ω) et nous présenterons deux algorithmes pour résoudre ces sous-problèmes barrières. Ensuite, nous utiliserons l'un de ces algorithmes pour résoudre le problème VI(S, Ω).

4.1 Cadre de travail

4.1.1 Définition du problème d'inéquations variationnelles

Soient $f_i : I\!\!R^n \to I\!\!R$ des fonctions convexes, de classe C^3 pour tout i appartenant à l'ensemble $\{1, ..., m\}$, A une matrice de $I\!\!R^{p\times n}$ $(p < n)$ et b un vecteur de $I\!\!R^p$. Définissons les ensembles G, Γ et Ω comme suit :

$$G = \{x |\ f_i(x) \leq 0,\ i = 1, ..., m\},$$
$$\Gamma = \{x |\ Ax = b\},$$
$$\Omega = G \cap \{x |\ Ax = b\}.$$

Supposons également que S est un opérateur défini sur l'ensemble Ω, à valeurs dans $I\!\!R^n$ et définissons l'ensemble Ω^0 comme étant l'intersection entre Ω et l'intérieur de G, c'est-à-dire,

$$\Omega^0 = \Omega \cap G_{int}.$$

Le problème d'inéquations variationnelles VI(S, Ω) est le suivant :

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \langle S(x^*), x - x^* \rangle \geq 0, \quad \forall x \in \Omega. \tag{4.1}$$

4.1.2 Les hypothèses sur Ω, A, S et F

Les hypothèses ci-dessous seront valables dans tout ce chapitre. La fonction F dont il est question est celle qui va nous permettre de définir la famille de sous-problèmes barrières associée à VI(S, Ω).

(A1) $\Omega^0 \neq \emptyset$;

(A2) $F : G_{int} \to I\!R$ est une fonction barrière ν-self-concordante sur G $(\nu \geq 1)$;

(A3) G est un ensemble compact;

(A4) La matrice A est de rang p;

(A5) L'opérateur S est monotone et continûment différentiable.

Commentaires

1. Grâce aux Propositions 2.2.1 et 2.2.2 du Chapitre 2, on peut facilement démontrer que, sous les hypothèses qui viennent d'être posées, le problème VI(S, Ω) admet au moins une solution.

2. L'hypothèse $\nu \geq 1$ n'est nullement restrictive. En effet, si F est une fonction barrière ν-self-concordante, où $\nu \leq 1$, alors les inégalités suivantes sont vérifiées :

$$|\langle F'(x), h \rangle| \leq \nu^{1/2} \|h\|_{x,F} \leq \|h\|_{x,F} \quad \forall h \in I\!R^n, x \in G_{int}.$$

   Dès lors, on peut prendre $\nu = 1$ et l'hypothèse (A2) est ainsi satisfaite.

4.2 Les sous-problèmes barrières VI$({}_\mu\, \Omega)$

4.2.1 Définition

Considérons la famille d'opérateurs

$$\{S_\mu(x) = S(x) + \mu F'(x)\}_{\mu > 0}. \tag{4.2}$$

Cette famille est composée d'opérateurs strictement monotones.

En effet, par l'hypothèse (A5), on sait déjà que l'opérateur S est monotone. De plus, par l'hypothèse (A3), on sait que E_F est non dégénéré (c'est-à-dire $E_F \neq \{0\}$). Dès lors, pour tout vecteur non nul h de $I\!R^n$ et pour tout vecteur $x \in G_{int}$, on a que $h^t F''(x) h > 0$, ce qui revient à dire que F' est strictement monotone.

Pour certaines valeurs du paramètre μ, nous considérerons le problème d'inéquations variationnelles $\mathrm{VI}(S_\mu, \Omega)$ suivant :

$$\text{Trouver } x(\mu) \in \Omega^0 \quad \text{tel que} \quad \langle S_\mu(x(\mu)), x - x(\mu)\rangle \geq 0 \quad \forall x \in \Omega. \tag{4.3}$$

On appelle l'ensemble $\{x(\mu),\ \mu > 0\}$, s'il existe, la trajectoire centrale.

Par les Propositions 2.2.1 et 2.2.2 et les hypothèses qui viennent d'être posées, on sait que le problème $\mathrm{VI}(S_\mu, \Omega)$ admet au moins une solution. Vu que S_μ est un opérateur strictement monotone, on sait qu'il admet au plus une solution.

4.2.2 Méthode de suivi de chemin et méthode de Newton

Pour résoudre le problème $\mathrm{VI}(S, \Omega)$, nous suivrons la trajectoire centrale dans le sens des μ décroissants. En effet, on montre facilement que lorsque μ tend vers 0, $S_\mu(x)$ converge vers $S(x)$. On espère donc qu'en prenant des valeurs de μ de plus en plus petites, on s'approchera de plus en plus d'une solution du problème $\mathrm{VI}(S, \Omega)$. C'est ce que nous analyserons maintenant.

Remarquons d'abord que les hypothèses (A1) et (A4) sont semblables à l'hypothèse de qualification de contraintes de Slater.

En effet, l'hypothèse de Slater est vérifiée si et seulement si il existe un point $\bar{x}$ tel que $f_i(\bar{x}) < 0,\ i \in \{1, ..., m\},\ A\bar{x} = b$.

Dès lors, on peut montrer qu'il existe un vecteur $y(\mu) \in I\!R^p$ tel que les conditions de KKT suivantes sont satisfaites :

$$S(x(\mu)) + \mu F'(x(\mu)) - A^t y(\mu) = 0, \tag{4.4}$$

$$Ax(\mu) = b, \quad x(\mu) \in G_{int}. \tag{4.5}$$

Inversement, si le couple $(x(\mu), y(\mu)) \in I\!R^n \times I\!R^p$ satisfait ces inégalités, alors $x(\mu)$ est la solution de $\mathrm{VI}(S_\mu, \Omega)$.

Théorème 4.1

Soit $\mu > 0$. Si les hypothèses (A1)-(A5) sont satisfaites, alors l'inégalité suivante est vérifiée:

$$\langle S(x(\mu)), x(\mu) - x\rangle \leq \mu\nu, \quad \forall x \in \Omega. \tag{4.6}$$

preuve

Par la relation (4.3) et l'inégalité (3.4) du Chapitre 3, on a :

$$\begin{aligned}\langle S(x(\mu)), x(\mu) - x\rangle &\leq \mu\langle F'(x(\mu)), x - x(\mu)\rangle \\ &\leq \mu\,\nu.\end{aligned}$$

□

Grâce à ce théorème, on sait qu'il existe une sous-suite de $\{x(\mu)\}_{\mu>0}$ qui converge vers une solution de $\mathrm{VI}(S, \Omega)$.
En effet, vu que G est compact, la suite $\{x(\mu)\}_{\mu>0}$ admet une sous-suite qui converge vers un point $x^* \in \Omega$ lorsque μ tend vers 0 et x^* vérifie la relation suivante :

$$\langle S(x^*), x^* - x\rangle \leq 0, \quad \forall x \in \Omega.$$

Le sous-problème $\mathrm{VI}(S_\mu, \Omega)$ va être résolu de façon inexacte. C'est pourquoi, il va falloir une mesure de précision pour déterminer si la solution approximative d'un sous-problème est proche de la solution exacte. Pour cela, considérons le système des conditions de KKT associé au sous-problème $\mathrm{VI}(S_\mu, \Omega)$:

$$S_\mu(x) - A^t y = 0, \tag{4.7}$$
$$Ax = b, \tag{4.8}$$
$$x \in G_{int} \quad y \in I\!R^p. \tag{4.9}$$

Définissons la fonction $\phi(S_\mu, x, y)$ par la relation suivante :

$$\phi(S_\mu, x, y) = S_\mu(x) - A^t y.$$

Par définition de $x(\mu)$ et de $y(\mu)$, $\phi(S_\mu, x(\mu), y(\mu)) = 0$. Dès lors, on peut naturellement choisir la fonction

$$\|\phi(S_\mu, x, y)\|^*_{F',x} = \sup_{w \neq 0}\{\frac{\phi^t(S_\mu, x, y)w}{\|w\|_{F',x}}\}$$

pour déterminer si un vecteur x est proche de la solution $x(\mu)$.

Réecrivons les équations (4.7) et (4.8) de la façon suivante:

$$H_\mu(x, y) \equiv \begin{bmatrix} S_\mu(x) - A^t y \\ Ax - b \end{bmatrix} = 0.$$

Etant donné $x \in \Omega^0$, $y \in I\!R^p$ et $\mu > 0$, pour trouver la direction de Newton, il faut résoudre le système suivant:

$$H_\mu(x, y) + H'_\mu(x, y) \begin{bmatrix} -\Delta x \\ -\Delta y \end{bmatrix} = 0$$

où $\Delta x = x - x^+$, $\quad \Delta y = y - y^+$. Ce système est équivalent à:

$$S_\mu(x) - A^t y + S'_\mu(x)(-\Delta x) + A^t \Delta y = 0. \tag{4.10}$$

$$A\Delta x = 0. \tag{4.11}$$

Ce système est défini vu que la matrice $S'_\mu(x)$ est définie positive, et par l'hypothèse (A4), la matrice A est de rang plein.
En particulier,

$$y^+ = [A(S'_\mu(x))^{-1}A^t]^{-1}A(S'_\mu(x))^{-1}S_\mu(x) \tag{4.12}$$

$$x^+ = x + (S'_\mu(x))^{-1}[A^t y^+ - S_\mu(x)] \tag{4.13}$$

sont des solutions de ce système.

Le théorème suivant indique la relation entre la mesure de proximité ϕ et la norme de la direction de Newton Δx. En particulier, dans un certain voisinage de $x(\mu)$, le nouvel itéré obtenu à partir d'un élement $x \in \Omega^0$ se trouve également dans l'ensemble Ω^0.

Théorème 4.2

Soient $\mu > 0$, $x \in \Omega^0$, $y \in I\!R^p$. Dès lors la relation suivante est vérifiée:

$$\|\Delta x\|_{F',x} \leq \frac{\|\phi(S_\mu, x, y)\|^*_{F',x}}{\mu}. \tag{4.14}$$

De plus , si $\|\phi(S_\mu, x, y)\|^*_{F',x}/\mu < 1$ alors $x^+ \in \Omega^0$.

preuve

1. Prouvons d'abord la relation (4.14).
   Par la relation (4.10), on a :

$$S'_\mu(x)(\Delta x) - A^t\Delta y = \phi(S_\mu, x, y).$$

   En multipliant les deux membres de cette égalité par $(\Delta x)^t$, on a

$$(\Delta x)^t S'_\mu(x)(\Delta x) - (\Delta x)^t A^t(\Delta y) = (\Delta x)^t\phi(S_\mu, x, y). \qquad (4.15)$$

   De plus, comme S est monotone, on a :

$$\begin{aligned}(\Delta x)^t S'_\mu(x)(\Delta x) &= (\Delta x)^t S'(x)(\Delta x) + \mu(\Delta x)^t F''(x)(\Delta x) \\ &\geq \mu(\Delta x)^t F''(x)(\Delta x).\end{aligned}$$

   Grâce à cette dernière inégalité et à l'égalité $A\Delta x = 0$, on obtient :

$$\begin{aligned}(\Delta x)^t F''(x)(\Delta x) &\leq \frac{(\Delta x)^t S'_\mu(x)(\Delta x)}{\mu} \\ &= \frac{(\Delta x)^t\ \phi(S_\mu, x, y)}{\mu} \\ &\leq \frac{\|\phi(S_\mu, x, y)\|^*_{F',x}\|\Delta x\|_{F',x}}{\mu}.\end{aligned}$$

   Par conséquent, on a que

$$\|\Delta x\|^2_{F',x} \leq \|\phi(S_\mu, x, y)\|^*_{F',x}\|\Delta x\|_{F',x}/\mu,$$

   En divisant les deux membres de cette inégalité par $\|\Delta x\|_{F',x}$, on obtient la relation (4.14).

2. Supposons que $\|\phi(S_\mu, x, y)\|^*_{F',x}/\mu < 1$ et montrons que $x^+ \in \Omega^0$.

   Par le point 1, on sait que $\|\Delta x\|_{F',x} < 1$. Or F' est un opérateur fortement self-concordant sur l'ensemble G_{int} et $x^+ \in W_1^0(F', x)$ avec $x \in G_{int}$. Par conséquent, il suit du résultat (3.8) de la Proposition 3.2.1 que $x^+ \in G_{int}$. De plus, par l'égalité (4.11), on a que $Ax^+ = b$. Dès lors, $x^+ \in \Omega^0$.

□

4.2.3 Lemmes préliminaires

Comme ce qui a été fait au Chapitre 2, nous établirons un premier algorithme pour résoudre le problème VI(S_μ, Ω). Cependant, nous allons pouvoir en établir un second dont la convergence est quadratique. Pour cela, nous devrons ajouter une hypothèse supplémentaire :

(A6) S_μ est un opérateur fortement self-concordant.

Dans ce paragraphe, nous supposerons que cette hypothèse est vérifiée.
Les lemmes suivants vont donc nous permettre d'établir le deuxième algorithme pour résoudre le problème VI(S_μ, Ω). Ils seront également utiles pour trouver une borne supérieure sur une fonction de mérite du problème VI(S, Ω).

Lemme 4.2.1

Soient $\mu > 0$, $x \in \Omega^0$ et $y \in I\!R^p$. Dès lors, l'inégalité suivante est vérifiée

$$\|\Delta x\|_{S_\mu,x} \leq \|\phi(S_\mu, x, y)\|^*_{S_\mu,x}. \tag{4.16}$$

De plus, si $\|\phi(S_\mu, x, y)\|^*_{S_\mu,x} < 1$, alors $x^+ \in \Omega^0$.

preuve

Par la relation (4.15) du Théorème 4.2, on a

$$\|\Delta x\|^2_{S_\mu,x} \leq \|\phi(S_\mu, x, y)\|^*_{S_\mu,x} \|\Delta x\|_{S_\mu,x}.$$

En divisant les deux membres de l'inégalité par $\|\Delta x\|_{S_\mu,x}$, on obtient la relation (4.16).
Le reste de la démonstration est tout à fait similaire à la deuxième partie du Théorème 4.2.

$\square$

Lemme 4.2.2

Soient $\kappa \geq \frac{(7+3\sqrt{5})}{10}$, $x \in \Omega^0$, $y \in I\!R^p$ et $\mu > 0$.

Si $\|\phi(S_\mu, x, y)\|^*_{S_\mu,x} \stackrel{def.}{=} \lambda \leq \frac{1}{5\kappa}$, alors l'inégalité suivante est satisfaite:

$$\|\phi(S_\mu, x^+, y^+)\|^*_{S_\mu,x^+} \leq \kappa\lambda^2. \tag{4.17}$$

preuve
Vu que S_μ est un opérateur fortement self-concordant, par la relation (3.7) de la Proposition 3.2.1 nous avons, pour tout $w \in I\!R^n$

$$\|w\|_{S_\mu,x}(1 - \|\Delta x\|_{S_\mu,x}) \leq \|w\|_{S_\mu,x^+} \leq \frac{\|w\|_{S_\mu,x}}{(1 - \|\Delta x\|_{S_\mu,x})}. \tag{4.18}$$

Par le Lemme 4.2.1, on sait que

$$\|\Delta x\|_{S_\mu,x} \leq \|\phi(S_\mu, x, y)\|^*_{S_\mu,x} \leq \frac{1}{5\kappa} < 1. \tag{4.19}$$

Nous pouvons donc réécrire la relation (4.18) comme suit :

$$\|w\|_{S_\mu,x}(1 - \lambda) \leq \|w\|_{S_\mu,x^+}.$$

Dès lors,

$$\begin{aligned}
\|\phi(S_\mu, x^+, y^+)\|^*_{S_\mu,x^+} &= \sup_{w\neq 0}\left\{\frac{\phi(S_\mu, x^+, y^+)w}{\|w\|_{S_\mu,x^+}}\right\} \\
&\leq \sup_{w\neq 0}\left\{\frac{\phi(S_\mu, x^+, y^+)w}{(1-\lambda)\|w\|_{S_\mu,x}}\right\} \\
&= \frac{\|\phi(S_\mu, x^+, y^+)\|^*_{S_\mu,x}}{1-\lambda}
\end{aligned} \tag{4.20}$$

Choisissons un vecteur $g \in I\!R^n$ tel que $\|g\|_{S_\mu,x} \leq 1$ et définissons la fonction f_g comme suit :

$$f_g(t) = \langle \phi(S_\mu, x - t\Delta x, y - t\Delta y), g\rangle, \quad 0 \leq t \leq 1. \tag{4.21}$$

La dérivée première de f_g vérifie les inégalités suivantes :

$$\begin{aligned}
f_g'(t) &= \left\langle -\phi'(S_\mu, x - t\Delta x, y - t\Delta y)\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix}, g\right\rangle \\
&= \left\langle -\phi'(S_\mu, x, y)\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix}, g\right\rangle \\
&\quad + \left\langle (\phi'(S_\mu, x, y) - \phi'(S_\mu, x - t\Delta x, y - t\Delta y))\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix}, g\right\rangle \\
&= -\langle \phi(S_\mu, x, y), g\rangle + \left\langle \left(\begin{bmatrix} S_\mu'(x) & -A^t\end{bmatrix} - \begin{bmatrix} S_\mu'(x - t\Delta x) & -A^t\end{bmatrix}\right)\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix}, g\right\rangle,
\end{aligned}$$

où la dernière égalité vient de la relation (4.10).

Finalement, on obtient la relation suivante :

$$f_g'(t) + \langle \phi(S_\mu, x, y), g\rangle = \langle [S_\mu'(x) - S_\mu'(x - t\Delta x)]\Delta x, g\rangle. \tag{4.22}$$

Par la relation (3.6) de la Proposition 3.2.1 et par la relation (4.19), on a

$$\begin{aligned} |\langle [S_\mu'(x) - S_\mu'(x - t\Delta x)]\Delta x, g\rangle| &\leq \left(\frac{1}{[1 - t\|\Delta x\|_{S_\mu,x}]^2} - 1\right) \|\Delta x\|_{S_\mu,x}\|g\|_{S_\mu,x}, \\ &\leq \lambda\left\{\frac{1}{(1-t\lambda)^2} - 1\right\} \end{aligned} \tag{4.23}$$

Dès lors, par les relations (4.22) et (4.23), on obtient

$$|f_g'(t) + \langle \phi(S_\mu, x, y), g\rangle| \leq \lambda\left\{\frac{1}{(1-t\lambda)^2} - 1\right\} \tag{4.24}$$

En intégrant l'inégalité (4.24) par rapport à t et en constatant que $\langle \phi(S_\mu, x, y), g\rangle = f_g(0)$, on obtient :

$$|f_g(1)| \leq \lambda \int_0^1 \left\{\frac{1}{(1-t\lambda)^2} - 1\right\} dt. \tag{4.25}$$

Faisons le changement de variable $r = t\lambda$ pour obtenir la relation suivante :

$$|\langle \phi(S_\mu, x^+, y^+), g\rangle| \leq \int_0^\lambda \frac{r(2-r)}{(1-r)^2} dr. \tag{4.26}$$

Vu que g est un vecteur quelconque de $I\!R^n$ tel que $\|g\|_{S_\mu,x} \leq 1$, on a que

$$\|\phi(S_\mu, x^+, y^+)\|_{S_\mu,x}^* \leq \int_0^\lambda \frac{r(2-r)}{(1-r)^2} dr. \tag{4.27}$$

Finalement, par la relation (4.20), on obtient

$$\begin{aligned} \|\phi(S_\mu, x^+, y^+)\|_{S_\mu,x^+}^* &\leq \left\{\int_0^\lambda \frac{r(2-r)}{(1-r)^2} dr\right\} /1 - \lambda \\ &\leq \kappa\lambda^2. \end{aligned}$$

□

Le lemme suivant nous indique que la méthode de Newton est localement et quadratiquement convergente (dans le sens donné par le Lemme 4.2.2) et nous donne une estimation de la distance entre un vecteur x et la solution de VI(S_μ, Ω).

Lemme 4.2.3

Soient $\kappa \geq \frac{(7+3\sqrt{5})}{10}$, $\mu > 0$, $x \in \Omega^0$, *et* $y \in I\!R^p$.
Considérons la suite des itérés $\{(x^i, y^i)\}_{i>0}$ *obtenue par les équations (4.10)-(4.11) avec comme point de départ (x,y).*
Si $\|\phi(S_\mu, x, y)\|^*_{S_\mu,x} < \frac{1}{5\kappa}$, *alors la suite* $\{\|\phi(S_\mu, x^i, y^i)\|^*_{S_\mu,x^i}\}$ *converge vers 0 et la suite* $\{x^i\}$ *converge vers l'unique solution de VI*(S_μ, Ω) *lorsque* i *tend vers* $+\infty$. *De plus, le point limite de la suite* $\{x^i\}$, *noté* x^*, *satisfait la relation suivante:*

$$\|x - x^*\|_{S_\mu,x^*} \leq 5\|\phi(S_\mu, x, y)\|^*_{S_\mu,x}. \tag{4.28}$$

preuve

Afin d'alléger les écritures, posons

$$x^0 = x, \quad y^0 = y, \quad \lambda_i = \|\phi(S_\mu, x^i, y^i)\|^*_{S_\mu,x^i}, \quad \lambda = \lambda_0.$$

1. Montrons d'abord que la suite $\{\|\phi(S_\mu, x^i, y^i)\|^*_{S_\mu,x^i}\}_{i>0}$ converge vers 0.

   Par le Lemme 4.2.2 et par l'hypothèse $\lambda_0 \leq \frac{1}{5\kappa}$, nous avons que

   $$\lambda_1 = \|\phi(S_\mu, x^1, y^1)\|^*_{S_\mu,x^1} \leq \kappa\lambda^2 \leq \frac{\lambda}{5} < 1.$$

   On peut prouver par récurrence que

   $$\lambda_i \leq \frac{\lambda}{5^i}.$$

   Cette inégalité implique que $\|\phi(S_\mu, x^i, y^i)\|^*_{S_\mu,x^i}$ converge vers 0 lorsque i converge vers l'infini.

2. Montrons ensuite que la suite $\{x^i\}_{i>0}$ converge vers l'unique solution de VI(S_μ, Ω).

   Par le Lemme 4.2.1, pour $i > 0$, on a

   $$\|x^i - x^{i-1}\|_{S_\mu,x^{i-1}} \leq \lambda_{i-1} < 1.$$

   L'inégalité (3.7) de la Proposition 3.2.1 implique que pour tout $i > 0$ tel que $x^{i-1} \in W_1^0(S_\mu, x^0)$, on a

   $$\|x^i - x^{i-1}\|_{S_{\mu,x^0}}(1 - \|x^0 - x^{i-1}\|_{S_\mu,x^0}) \leq \|x^i - x^{i-1}\|_{S_\mu,x^{i-1}}$$

Posons

$$\rho_i = \|x^i - x^{i-1}\|_{S_\mu,x^0} \quad \text{et} \quad P_i = \|x^i - x^0\|_{S_\mu,x^0}.$$

On a immédiatement que la relation suivante est vérifiée :

$$P_i \leq P_{i-1} + \rho_i.$$

Nous allons maintenant montrer par récurrence que

$$P_i \leq \frac{1}{2} \quad \forall i.$$

Cas de base $P_0 = 0 \leq 1/2$.
Hypothèse de récurrence Pour tout $j < i$, supposons que $P_j \leq 1/2$.
Montrons que $P_i \leq 1/2$

Comme $x^{i-1} \in W_1^0(S_\mu, x^0)$, on a que

$$\rho_i \leq \frac{\|x^i - x^{i-1}\|_{S_{\mu,x^{i-1}}}}{(1 - \|x^{i-1} - x^0\|_{S_\mu,x^0})} \leq \frac{\lambda_{i-1}}{(1 - \|x^{i-1} - x^0\|_{S_\mu,x^0})} = \frac{\lambda_{i-1}}{(1 - P_{i-1})}.$$

Vu que $\lambda_{i-1} \leq \frac{\lambda}{5^{i-1}}$ et que $P_{i-1} \leq 1/2$, on a que

$$\rho_i \leq \frac{2\lambda}{5^{i-1}}.$$

Par conséquent,

$$P_i \leq P_{i-1} + \rho_i \leq P_0 + \sum_{j=1}^{i} \rho_j = \sum_{j=1}^{i} 2\lambda 5^{1-j} \leq 2,5\lambda < \frac{1}{2}.$$

Remarquons également que $\sum_i \rho_i < +\infty$. En effet, le terme général de la série est plus petit que 1. Dès lors, la suite $\{\rho_i\}$ converge vers 0 et la suite $\{x^i\}$ converge vers un point x^* satisfaisant la relation

$$\|x^* - x^0\|_{S_\mu,x^0} = \|x^* - x\|_{S_\mu,x} \leq 2,5\lambda \leq \frac{1}{2}.$$

Dès lors, $x^* \in W_{1/2}(S_\mu, x^0)$. La relation (3.8) de la Proposition 3.2.1 implique que $x^* \in G_{int}$. De plus, $Ax^i = b$ et donc, en passant à la limite, $x^* \in \Gamma$. Dès lors, $x^* \in \Omega^0$.

3. Finalement, montrons que $\|x - x^*\|_{S_\mu,x^*} \leq 5\|\phi(S_\mu, x, y)\|^*_{S_\mu,x}$.
Etant donné x^{i-1}, prenons $\bar{y}^i$, une solution du système (4.10)-(4.11), choisie selon l'équation (4.12).
La suite $\{\bar{y}^i\}$ converge donc vers un point $\bar{y}^*$. De plus, vu que

$$\lim_{i \to +\infty} \phi(S_\mu, x^i, y^i) = 0,$$

le couple $(x^*, \bar{y}^*)$ satisfait le système d'équations (4.7)-(4.8). Dès lors, x^* est la solution de $\mathrm{VI}(S_\mu, \Omega)$. Finalement, il suit du résultat (3.7) de la Proposition 3.2.1 que

$$\begin{aligned} \|x - x^*\|_{S_\mu,x^*} &\leq \frac{\|x - x^*\|_{S_\mu,x}}{(1 - \|x^* - x\|_{S_\mu,x})} \\ &\leq 2\|x - x^*\|_{S_\mu,x} \leq 5\lambda. \end{aligned}$$

□

4.2.4 Précision des approximations

Dans cette section, nous définirons une fonction de mérite pour le problème $\mathrm{VI}(S, \Omega)$. Ensuite, nous établirons un résultat qui nous fournira une borne supérieure sur cette fonction. Ce sera l'objet du Théorème 4.3.

Considérons la fonction linéaire g_x, $x \in \Omega^0$, définie pour tout z de $I\!R^n$ par la relation suivante :

$$g_x(z) = \langle S(x), x - z \rangle.$$

La fonction de mérite g est définie comme suit :

$$g(x) = \max\{g_x(z) |\ z \in \Omega\}.$$

On peut facilement démontrer que $g(x) \geq 0$ pour tout x appartenant à Ω^0 et $g(x) = 0$ si et seulement si x est une solution de $\mathrm{VI}(S, \Omega)$. Dès lors, nous emploierons la fonction $g(x)$ pour vérifier si un point $x \in \Omega^0$ est proche de la solution de $\mathrm{VI}(S, \Omega)$.

Théorème 4.3

Soient $x \in \Omega^0$, $y \in I\!R^p$, $\mu > 0$, $\kappa \geq \frac{(7+3\sqrt{5})}{10}$. Si la relation suivante est vérifiée

$$\frac{\|\phi(S_\mu, x, y)\|^*_{F',x}}{\mu} \stackrel{def.}{=} \lambda \leq \frac{1}{5\kappa},$$

alors, on a

$$g(x) \leq 2\mu\nu.$$

preuve

Considérons la fonction f définie sur Ω^0 par

$$f(u) = \frac{S(x)}{\mu} + F'(u) \quad \text{où } x \in \Omega. \tag{4.29}$$

En constatant que

1) $\phi(f, x, \frac{y}{\mu}) = f(x) - \frac{A^t y}{\mu} = \frac{S_\mu(x) - A^t y}{\mu} = \frac{\phi(S_\mu, x, y)}{\mu}$;

2) $\| \cdot \|_{f,x} = \| \cdot \|_{F',x}$ car $f' \equiv F'$,

on obtient la relation suivante :

$$\|\phi(f, x, \frac{y}{\mu})\|^*_{f,x} = \frac{\|\phi(S_\mu, x, y)\|^*_{F',x}}{\mu} = \lambda. \tag{4.30}$$

Considérons ensuite le problème d'inéquations variationnelles VI(f, Ω) suivant :

$$\text{Trouver } x^* \in \Omega^0 \quad \text{tel que} \quad \langle f(x^*), z - x^* \rangle \geq 0, \quad \forall z \in \Omega.$$

Grâce aux Propositions 2.2.1 et 2.2.2, on sait qu'il existe au moins une solution au problème VI(f, Ω). L'hypothèse (A2) implique que F' est de classe C^2 sur Ω^0 et par l'hypothèse (A3), F' est strictement monotone sur l'ensemble Ω^0. Dès lors, f est également strictement monotone et de classe C^2 sur Ω^0. Le problème VI(f, Ω) admet donc une et une seule solution. Dès lors, il existe deux vecteurs $x^* \in \Omega^0$ et $y^* \in I\!R^p$ qui vérifient les conditions de KKT associées au problème VI(f, Ω), c'est-à-dire,

$$f(x^*) - A^t y^* = 0, \tag{4.31}$$

$$Ax^* = b, \quad x^* \in G_{int}. \tag{4.32}$$

Par les relations (4.29) et de (4.31), on a

$$S(x) = -\mu F'(x^*) + \mu A^t y^*.$$

Vu que $Ax = Ax^* = b$, on a

$$\begin{aligned} g_x(x^*) &= \langle -S(x), x^* - x \rangle \\ &= \langle \mu F'(x^*), x^* - x \rangle - \mu \langle A^t y^*, x^* - x \rangle \\ &= \langle \mu F'(x^*), x^* - x \rangle \\ &\leq \mu \|F'(x^*)\|^*_{F',x^*} \|x^* - x\|_{F',x^*}. \end{aligned}$$

Comme x est fixé et F' est un opérateur fortement self-concordant, f est également un opérateur fortement self-concordant. Par la relation (4.30), le Lemme 4.2.3 où l'on remplace S_μ par f, on a que

$$\|x - x^*\|_{F',x^*} = \|x - x^*\|_{f,x^*} \leq 5\lambda.$$

Comme F est une fonction barrière fortement self-concordante pour G, on a

$$\begin{aligned}\|F'(x^*)\|^*_{F',x^*} &= \sup_{h\neq 0}\{\frac{\langle F'(x^*), h\rangle}{\|h\|_{F',x^*}}\} \\ &= \sup_{h\neq 0}\{\frac{\langle F'(x^*), h\rangle}{\|h\|_{x^*,F}}\} \leq \nu^{1/2}.\end{aligned}$$

Dès lors, en considérant aussi le fait que $\nu \geq 1$, on a que

$$g_x(x^*) \leq 5\lambda\mu\nu^{1/2} \leq \mu\nu.$$

En constatant que

$$g_x(z) - g_x(x^*) = \langle S(x), x^* - z\rangle,$$

on obtient :

$$\begin{aligned}\max_{z\in\Omega} g_x(z) - g_x(x^*) &= \max_{z\in\Omega}\{\langle S(x), x^* - z\rangle\} \\ &= \max_{z\in\Omega}\{\mu\langle F'(x^*), z - x^*\rangle + \mu\langle A^t y^*, x^* - z\rangle\} \\ &= \max_{z\in\Omega}\{\mu\langle F'(x^*), z - x^*\rangle\} \leq \mu\nu,\end{aligned}$$

où l'inégalité est due à la relation (3.4) de la Proposition 3.2.1. On a donc la thèse vu que

$$\begin{aligned}g(x) &= \max_{z\in\Omega}\{g_x(z)\} - g_x(x^*) + g_x(x^*) \\ &\leq \mu\nu + g_x(x^*) \leq 2\mu\nu.\end{aligned}$$

□

4.2.5 La condition de centrage

Soient $x \in \Omega^0$, $y \in I\!R^p$, $\kappa \geq \frac{(7+3\sqrt{5})}{10}$, $\mu > 0$. On dit que le couple (x, y) satisfait la condition de centrage si

$$\frac{\|\phi(S_\mu, x, y)\|^*_{F',x}}{\mu} \leq \frac{1}{5\kappa}.$$

Dès lors, on peut reformuler le théorème précédent de la façon suivante :

Si le couple (x, y) satisfait la condition de centrage, alors $g(x) \leq 2\mu\nu$.

4.2.6 Reformulation de VI(S_μ, Ω)

A l'instar de ce qui a été fait au Chapitre 2, nous définissons une fonction de mérite associée au problème (S_μ, Ω). Pour avoir des résultats semblables à ceux obtenus précédement, il faut que S_μ soit fortement monotone.

Grâce au lemme suivant, nous savons que si F' est fortement monotone, alors S_μ l'est aussi.

Lemme 4.2.4

Supposons que F' soit fortement monotone sur Ω^0, c'est-à-dire,

$$u^t F''(x) u \geq \gamma \|u\|^2, \quad \forall x \in \Omega^0, u \in I\!R^n.$$

Sous cette hypothèse, S_μ est fortement monotone et la constante vaut $\mu\gamma$.

preuve
Vu que S est monotone sur Ω, on a

$$\langle S'(x)u, u\rangle \geq 0 \quad \forall x \in \Omega^0, \ u \in I\!R^n.$$

Par définition de S_μ,

$$S(x) = S_\mu(x) - \mu F'(x).$$

Dès lors, on obtient

$$\langle S'_\mu(x)u, u\rangle \geq \mu\langle F''(x)u, u\rangle \geq \mu\gamma\|u\|^2, \quad \forall x \in \Omega^0, u \in I\!R^n.$$

□

Dans la suite de ce paragraphe, nous supposerons donc que F est une fonction barrière ν-self-concordante dont la dérivée est fortement monotone.

La fonction de mérite associée à VI(S_μ, Ω)

Soit B, une matrice $n \times n$ symétrique et définie positive.

Pour chaque $x \in \Omega^0$, considérons $\mathbf{P(x)} \in \Omega$ la solution du problème d'inéquations variationnelles suivant :

$$[S_\mu(x) + B(P(x) - x)]^t[y - P(x)] \geq 0, \quad \forall y \in \Omega, \tag{4.33}$$

qui est équivalent au problème suivant :

$$\min_{\hat{y} \in \Omega} \|\hat{y} - (x - B^{-1}S_\mu(x))\|_B^2,$$

dont la solution optimale est précisément

$$P(x) = \text{Proj}_{\Omega,B}(x - B^{-1}S_\mu(x)).$$

Comme Ω est un ensemble fermé et convexe et B est une matrice définie positive, ce problème admet une solution unique.

Nous pouvons alors définir la fonction de mérite $f_\mu : \Omega^0 \to I\!\!R$ comme suit :

$$f_\mu(x) = -S_\mu(x)^t(P(x) - x) - \frac{1}{2}[P(x) - x]^t B[P(x) - x] \tag{4.34}$$

et considérer le problème d'Optimisation suivant :

$$\min_{x \in \Omega^0} f_\mu(x) = \min_{x \in \Omega^0} \left(-S_\mu(x)^t(P(x) - x) - \frac{1}{2}[P(x) - x]^t B[P(x) - x]\right). \tag{4.35}$$

Définissons également, pour chaque $x \in \Omega^0$, le vecteur $\mathbf{Z(x)} \in \Gamma$ comme étant la solution du problème suivant :

$$[S_\mu(x) + S'_\mu(x)^t(Z(x) - x)]^t[y - Z(x)] \geq 0 \quad \forall y \in \Gamma. \tag{4.36}$$

Vu que S_μ est fortement monotone, $S'_\mu(x)$ est une matrice fortement définie positive. Dès lors, on peut montrer que ce problème admet une et une seule solution (ref. [7]).

Remarque

Le système (4.10)-(4.11) représente les conditions de (KKT) du problème (4.36). Dès lors,

$$x^+ = Z(x).$$

Propriétés de la fonction f_μ

La proposition suivante est tout à fait similaire à la Proposition 2.3.1 du Chapitre 2. Bien que l'ensemble Ω ait changé, la preuve faite au chapitre 2 est encore valable dans notre cas. C'est pourquoi, je ne démontrerai pas cette proposition.

Proposition 4.4.1

Soient la fonction f_μ définie plus haut et $x(\mu)$ la solution du problème $VI(S_\mu, \Omega)$. La fonction f_μ vérifie les propriétés suivantes:

1. $f_\mu(x) \geq 0$, $\forall x \in \Omega^0$;
2. $f_\mu(x) = 0$ *si et seulement si* $x \in \Omega^0$ *est une solution du problème* $VI(S_\mu, \Omega)$ *et du problème (4.35);*
3. $f_\mu(x) \geq [\mu\gamma - \frac{1}{2}\|B\|]\ \|x(\mu) - x\|^2 \quad \forall x \in \Omega^0$;
4. *La fonction $f_\mu(x)$ est continûment différentiable et son gradient est donné par la formule suivante:*

$$f'_\mu(x) = S_\mu(x) - [S'_\mu(x) - B][P(x) - x], \quad \forall x \in \Omega^0;$$

5. *Si x n'est pas une solution de $VI(S_\mu, \Omega)$, alors une direction de descente pour f_μ en x est donnée par $d = Z(x) - x$ où $Z(x) \neq x$, c'est-à-dire,*

$$\langle f'_\mu(x), Z(x) - x\rangle < -(\mu\gamma - \frac{1}{2}\|B\|)\|d\|^2, \quad \forall x \in \Omega^0, \quad \text{si } \|B\| \leq 2\mu\gamma.$$

Remarque

Vu que la matrice B dépend de μ, nous la noterons B_μ.
On peut choisir $B_\mu = 2\mu\gamma I$ où I est la matrice identé $n \times n$. La condition $\|B\| \leq 2\mu\gamma$ est ainsi immédiatement vérifiée.

4.2.7 Premier algorithme pour résoudre les sous-problèmes VI(S_μ, Ω) : l'algorithme 4.1

Etablissons maintenant l'algorithme pour résoudre VI(S_μ, Ω). Celui-ci est tout à fait similaire à celui du Chapitre 2. Pour la preuve de la convergence de l'algorithme, nous renvoyons le lecteur à ce chapitre.

Initialisation : Choisir $x^0 \in \Omega^0$, $0 < \rho, \sigma < 1$ et poser $k = 0$.

Itération k : Effectuer les pas 1 à 4 ci-dessous.

Pas 1: Critère d'arrêt

Si un critère de convergence est satisfait, alors on arrête.
Ce critère de convergence peut être

$$f_\mu(x^k) \leq \varepsilon \quad \text{ou} \quad \|Z(x^k) - x^k\| \leq \varepsilon \ \text{avec } \varepsilon > 0$$

Pas 2: La direction de Newton

Résoudre le système suivant

$$\begin{cases} S_\mu(x^k) + S'_\mu(x^+ - x^k) - A^t y^+ = 0 \\ Ax^+ = b \end{cases}$$

On obtient ainsi $d^k = x^+ - x^k$.

Pas 3: Recherche linéaire inexacte

Prendre t^k, le plus grand nombre parmi $\{1, \rho, \rho^2, ...\}$ tel que :

$$\begin{cases} f_\mu(x^k + t^k d^k) \leq f_\mu(x^k) + \sigma t^k f'_\mu(x^k)^t d^k, \\ x^k + t^k d^k \in \Omega^0. \end{cases}$$

Pas 4: Mise à jour

Poser $x^{k+1} = x^k + t^k d^k$, $y^{k+1} = y^+$, $k = k + 1$.

Théorème 4.4 Convergence de l'algorithme 4.1.

Soit B, une matrice telle que $\|B\| \leq 2\mu\gamma$.
Etant donné un vecteur initial $x^0 \in \Omega^0$*, l'algorithme 4.1 engendre une suite d'itérés* $\{x^k\}$ *qui converge vers l'unique solution* $x(\mu)$ *du problème* $VI(S_\mu, \Omega)$.

preuve

De nouveau, cette preuve est tout à fait similaire à celle du Théorème 2.2. En effet, il faut commencer par démontrer que la suite $\{f_\mu(x^k)\}$ est décroissante et que la suite $\{x^k\}$ est bornée. Ensuite, on montre que $\lim_{x \to \bar{\Omega}_{int}} f_\mu(x) = +\infty$. On montre ainsi que la suite $\{x^k\}$ possède une sous-suite convergente. Finalement, il ne reste plus qu'à démontrer que la limite de cette sous-suite est la solution du problème VI(S_μ, Ω).

□

4.2.8 Deuxième algorithme pour résoudre les sous-problèmes VI(S_μ, Ω): l'algorithme 4.2

Suposons que S_μ soit un opérateur fortement self-concordant. En nous basant sur le résultat du Lemme 4.2.3, nous établirons un nouvel algorithme pour résoudre VI(S_μ, Ω) selon l'idée suivante: lorsque le couple (x^k, y^k) vérifiera la condition de centrage, nous poserons $t^k = 1$. Sinon, nous effectuerons une recherche linéaire le long de la direction de Newton.

Initialisation :

Choisir $x^0 \in \Omega^0,\ y \in I\!R^p,\ 0 < \rho, \sigma < 1,\ \kappa \geq \frac{(7+3\sqrt{5})}{10}$ et poser $k = 0$.

Itération k : Effectuer les pas 1 à 4 ci-dessous.

Pas 1: Critère d'arrêt

Si un critère de convergence est satisfait, alors arrêter.

Pas 2: La direction de Newton

Trouver (x^+, y^+) en résolvant le systèm suivant

$$\begin{cases} S_\mu(x^k) + S'_\mu(x^+ - x^k) - A^t y^+ = 0 \\ Ax^+ = b \end{cases}$$

On obtient alors $d^k = x^+ - x^k$.

Pas 3: Recherche linéaire

1. Si $\|\phi(S_\mu, x^+, y^+)\| < \frac{1}{5\kappa}$ alors poser $t^k = 1,\ x^{k+1} = x^+$;

2. Sinon, prendre t^k, le plus grand nombre parmi $\{1, \rho, \rho^2, ...\}$ tel que:

$$\begin{cases} f_\mu(x^k + t^k d^k) \leq f_\mu(x^k) + \sigma t^k f'_\mu(x^k)^t d^k, \\ x^k + t^k d^k \in \Omega^0. \end{cases}$$

   Poser $x^{k+1} = x^k + t^k d^k$.

Pas 4: Mise à jour

Poser $y^{k+1} = y^+,\ \ k = k + 1$.

Nous pouvons montrer que cet algorithme converge quadratiquement vers la solution du problème VI(S_μ, Ω). Ce sera l'objet du théorème suivant. La preuve est tout à fait semblable à celle du Théorème 2.3.

Théorème 4.5 Convergence de l'algorithme 4.2.

Il existe un entier $\bar{k}$ dépendant de μ et de n tel que pour tout $k \geq \bar{k}$, la suite $\{x^k\}$ engendrée par l'algorithme 4.2 converge localement et quadratiquement vers la solution de VI(S_μ, Ω) avec $t^k = 1$.

4.3 Algorithme 4.3 pour résoudre VI(Ω)

Dans cette section, nous allons développer un algorithme selon l'idée suivante : si l'itéré courant satisfait la condition de centrage, alors le paramètre de pénalité est multiplié par un facteur θ, $0 < \theta < 1$. Sinon, on applique l'algorithme 4.1 pour que l'itéré vérifie de nouveau la condition de centrage. Cet algorithme sera convergent.

Avant de présenter cet algorithme, nous expliquons certaines notations.
K est l'ensemble des indices de tous les itérés engendrés par l'algorithme 4.3. K^1 est l'ensemble des indices des itérés qui vérifient la condition de centrage. K^2 est le complémentaire de K^1. Pour une valeur de k^1 fixée, l'ensemble $K^2_{k^1}$ est composé des indices des itérés qui ne vérifient pas la condition de centrage pour une certaine valeur de μ.

<u>**Initialisation**</u>

Choisir $x^0 \in \Omega^0$, $y^0 \in I\!R^p$, $0 < \theta, \sigma, \rho < 1$, $\mu^0 > 0$;
$\kappa \geq (7 + 3\sqrt{5})/10$, $\varepsilon > 0$, B^0 tel que $\|B^0\| \leq 2\mu\gamma$;
Poser $k = 0$, $k^1 = 0$, $K = \{0\}$, $K^1 = \emptyset$, $K^2 = \emptyset$, $K^2_0 = \emptyset$.

<u>**Itération k**</u> Effectuer les pas 1 à 4 ci-dessous.

<u>**Pas 1**</u>**: Critère d'arrêt**

Si un critère de convergence est satisfait, alors arrêter. Ce critère de convergence peut être

$$\text{soit } g(x^k) \leq \varepsilon \quad \text{soit } \mu^k \leq \frac{\varepsilon}{2\nu}.$$

<u>**Pas 2**</u>**: Vérification de la condition de centrage**

Si $\|\phi(S_{\mu^k}, x^k, y^k)\|^*_{F', x^k}/\mu^k \leq \frac{1}{5\kappa}$,
alors poser $\mu^{k+1} = \theta\mu^k$, $K^1 = K^1 \cup \{k\}$, $k^1 = k^1 + 1$, $K^2_{k^1} = \emptyset$.
Sinon, poser $\mu^{k+1} = \mu^k$, $K^2_{k^1} = K^2_{k^1} \cup \{k\}$, et $K^2 = K^2 \cup \{k\}$.

Pas 3: Recentrage

Résoudre le système suivant :

$$\begin{cases} S_\mu(x^k) + S'_\mu(x^k)(x^+ - x^k) - A^t y^+ = 0 \\ Ax^+ = b \end{cases}$$

pour trouver la direction de descente $d^k = x^+ - x^k$.

Pas 3a: Recherche linéaire

Prendre t^k la plus grande valeur parmi $\{1, \rho, \rho^2, ...\}$ telle que

$$\begin{cases} f_{\mu^{k+1}}(x^k + t^k d^k) \leq f_{\mu^{k+1}}(x^k) + \sigma t^k f'_{\mu^{k+1}}(x^k)^t d^k, \\ x^k + t^k d^k \in \Omega^0 \end{cases}$$

Poser $x^{k+1} = x^k + t^k d^k$.

Pas 3b: Mise à jour

Poser $y^{k+1} = y^+$, $K = K \cup \{k+1\}$, $k = k+1$ et retourner au premier pas.

Théorème 4.6 : convergence de l'algorithme 4.3

Toute sous-suite convergente de $\{x^k\}_{k \in K^1}$ engendrée par l'algorithme 4.3 converge vers une solution du problème VI(S, Ω), pour tout itéré initial $x^0 \in \Omega^0$.

preuve

1. Tout d'abord, il faut montrer que l'ensemble $K^2_{k^1}$ est fini.
   Cela se fait de manière analogue à la première partie de la preuve du Théorème 2.4.

2. Montrons ensuite que toute sous-suite de $\{x^k\}_{k \in K^1}$ converge vers une solution du problème VI(S, Ω).
   Remarquons d'abord que pour tout $k \in K^1$, le couple (x^k, y^k) vérifie la condition de centrage. Dès lors, par le Théorème 4.3, on a

   $$g(x^k) \leq 2\mu^k \nu.$$

   Or, μ^k converge vers zéro lorsque k converge vers $+\infty$. Par conséquent,

   $$\lim_{\substack{k \in K^1 \\ k \to +\infty}} g(x^k) = 0.$$

De plus, $x^k \in \Omega^0$. Vu que Ω est un ensemble compact, il existe une sous-suite de $\{x^k\}_{k \in K^1}$ qui converge vers x^*. Comme la fonction g est continue, $g(x^*) = 0$. Dès lors, x^* est une solution du problème VI(S, Ω).

$\square$

Remarque

1. Pour "recentrer" l'itéré, nous appliquons l'algorithme 4.1. Cependant, si l'hypothèse (A6) est vérifiée, on peut remplacer l'algorithme 4.1 par l'algorithme 4.2 et on obtient des résultats analogues à ceux énoncés ci-dessus.

2. Lors de l'exécution de cet algorithme, la fonction de mérite doit être évaluée. Cependant, cela peut avoir un coût assez élevé. Pratiquement, on peut éviter cela en choisissant ρ assez petit de sorte que la fonction de mérite diminue suffisamment.

3. La complexité de cet algorithme dépend du comportement de la méthode de Newton qui se situe au pas 3. Vu que l'opérateur S_μ n'est pas forcément self-concordant, on ne sait pas combien d'itérations seront nécéssaires pour "centrer" l'itéré courant. La complexité de cet algorithme reste donc inconnue.

Conclusion

Dans un premier temps, nous nous sommes intéressés aux conditions d'existence et d'unicité associées au problème d'inéquations variationnelles suivant :

$$\text{Trouver } x^* \in \Omega \quad \text{tel que} \quad \forall x \in \Omega \quad F(x^*)^t(x - x^*) \geq 0,$$

où Ω est un sous-ensemble de $\mathbb{R}^n$ et $F : \Omega \to \mathbb{R}^n$. Nous avons alors démontré qu'il suffisait que Ω soit un ensemble fermé, convexe, à intérieur non vide et que F soit monotone et continue pour que ce problème admette au moins une solution. De plus, si l'ensemble Ω est borné, cette solution est unique. Dès lors, grâce aux hypothèses posées au début des Chapitres 2 et 4, on sait qu'il existe une et une seule solution à chaque sous-problème barrière et qu'il en existe au moins une au problème VI(F, Ω) ainsi qu'au problème VI(S, Ω).

Ensuite, on a considéré le problème VI(F, Ω). On a d'abord établi l'algorithme 2.1 pour résoudre les sous-problèmes barrières. On a vu que cet algorithme était convergent et que si la fonction F vérifiait la propriété (SLC) alors, à partir d'un certain moment, on pouvait fixer le pas de la méthode de Newton à 1. L'algorithme ainsi défini converge quadratiquement vers l'unique solution du problème VI($\bar{F}, \Omega_{int}$). A partir de l'algorithme 2.1, nous avons élaboré l'algorithme 2.4. Grâce aux hypothèses posées au début du deuxième chapitre, nous avons montré que l'algorithme 2.4 engendrait une suite d'itérés $\{x^k\}$ convergeant vers une solution du problème VI(F, Ω).

Finalement, nous avons étudié le problème d'inéquations variationnelles précédent dans un cadre plus général. Cela nous a amené à considérer le problème VI(S, Ω). Nous avons montré que si nous ajoutions à la fonction objectif la dérivée d'une fonction barrière self-concordante multipliée par un facteur μ, alors nous avions des résultats de convergence analogues à ceux du Chapitre 2. En effet, les deux algorithmes pour résoudre les sous-problèmes barrières ainsi que l'algorithme pour résoudre VI(S, Ω) se sont avérés convergents.

Bibliographie

[1] F. SHARIFI-MOKHTARIAN et J.L. GOFFIN. *Long-Step Interior-Point Algorithms for a Class of Variational Inequalities with Monotone Operators. Journal of Optimization Theory and Applications*, 97:181–210, Avril 1998.

[2] NESTEROV Y. et NEMIROVSKII A. *Interior-Point Polynomial Algorithms in Convex Programming.* SIAM Studies in Applied Mathematics, Philadelphia, 1994.

[3] NASH S.G. et SOFER A. *Linear and Nonlinear Programming. Mc Graw-Hill International Series*, 1996.

[4] WU J.H. *A Modified Path-Following Scheme for the Monotone Variational Inequality Problem. Journal of Optimization Theory and Applications*, 95:189–208, Octobre 1997.

[5] WU J.H. *Long-Step Primal Path-Following Algorithms for Monotone Variational Inequalities Problems. Journal of Optimization Theory and Applications*, 99:509–531, Novembre 1998.

[6] FUKUSHIMA M. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming*, 53:99–110, 1995.

[7] TSENG P. *Global Linear Convergence of a Path-Following Algorithm for Some Monotone Variational Inequalities Problems. Journal of Optimization Theory and Applications*, 75:265–279, 1992.

[8] LANSER S. *Barrier functions for interior point methods for Nonlinear Programming and their relation to the characteristic function.* 1995.