University of Namur

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Méthodes primales et duales de type Dinkelbach en programmation fractionnelle généralisée

Simons, Emilie

Award date:
2006

FUNDP
NAMUR

Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B - 5000 Namur (Belgique)

Méthodes primales et duales de type Dinkelbach en programmation fractionnelle généralisée

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

SIMONS Emilie

Promoteur : STRODIOT Jean-Jacques

Année Académique 2005-2006

Je souhaite adresser ici tous mes remerciements aux personnes qui m'ont apporté leur aide et qui ont ainsi contribué à l'élaboration de ce mémoire.

Tout d'abord, Monsieur Jean-Jacques Strodiot pour l'aide et le temps qu'il a bien voulu me consacrer. Enfin, j'adresse mes plus sincères remerciements à tous mes proches et amis qui m'ont toujours soutenue et encouragée au cours de la réalisation de ce mémoire.

Méthodes primales et duales de type Dinkelbach en programmation fractionnelle généralisée

Résumé

Dans ce mémoire, nous présentons plusieurs méthodes implémentables pour résoudre des problèmes de programmation fractionnelle généralisée. Nous développons les méthodes de type Dinkelbach traditionnelles ainsi que les méthodes de type Dinkelbach où un terme de régularisation proximale est ajouté pour éviter les difficultés numériques qui peuvent survenir. Ensuite, dans le cas de la programmation fractionnelle convexe, nous développons une méthode qui peut être considérée comme l'algorithme "dual" de l'algorithme de type Dinkelbach.

Dinkelbach Primal and Dual Methods in Generalized Fractional Programming

Abstract

In this thesis, we present several implementable methods for solving generalized fractional programming problems. We develop the traditional Dinkelbach-type methods as well as the Dinkelbach-type methods in which a prox-regularization term is added in order to prevent the numerical difficulties from arising. Subsequently, in the framework of convex fractional programming, we develop a method that can be seen as the "dual" algorithm to the Dinkelbach-type one.

Table des matières

Introduction

Dans ce mémoire, nous nous intéressons à une famille d'algorithmes permettant de résoudre les problèmes de programmation fractionnelle généralisée du type suivant :

$$(P) \qquad \bar{\theta} = \inf_{x \in S} \left(\max_{1 \leq i \leq p} \frac{f_i(x)}{g_i(x)} \right)$$

où $S \subseteq \mathbb{R}^n$ est un ensemble non vide, les fonctions f_i sont continues sur S et les fonctions g_i sont continues et positives sur S.

Le cas où $p = 1$ correspond à la programmation fractionnelle traditionnelle et a été énormément développé dans les années 1970. La programmation fractionnelle généralisée ($p > 1$) a, quant à elle, été développée un peu plus tard.

Nous commençons par développer un algorithme qui n'est autre qu'une généralisation de l'algorithme de Dinkelbach. Cet algorithme nécessite la résolution de l'équation suivante

$$F(\theta) = 0$$

où $F(\theta)$ est la valeur optimale du problème paramétrique suivant

$$(P_\theta) \qquad F(\theta) = \inf_{x \in S} \left(\max_{1 \leq i \leq p} [f_i(x) - \theta g_i(x)] \right).$$

Lorsque les fonctions f_i sont convexes et les fonctions g_i concaves, nous ajoutons, à la fonction objectif du problème (P_θ), un terme de régularisation proximale qui permet de supprimer les difficultés numériques survenant lorsque la solution du problème n'est pas unique. Néanmoins, les sous-problèmes restent difficiles à résoudre puisqu'ils sont non différentiables. Une stratégie possible est d'utiliser la stratégie faisceaux pour construire une approximation convexe linéaire par morceaux de la fonction objectif non différentiable des sous-problèmes. Nous introduisons également, pour la programmation fractionnelle généralisée, un problème dual sans saut de dualité. Dans le cas linéaire, ce problème a la même forme que le problème primal et par conséquent, les algorithmes que nous avons développés sont également applicables au problème dual. Dans le cas non linéaire, ce problème dual peut être résolu en utilisant également une approche paramétrique. L'algorithme développé peut être considéré, en quelque sorte, comme le "dual" de l'algorithme de type Dinkelbach que nous avons développé auparavant. Comme dans le cas linéaire, il est également possible, sous certaines conditions supplémentaires, de retrouver la solution optimale du problème primal.

Ce mémoire est composée de sept chapitres. Le premier chapitre est consacré aux définitions et aux propriétés des fonctions convexes et quasiconvexes. Le second est dédié à la description des méthodes de point proximal approché. La programmation fractionnelle généralisée est développée dans le chapitre trois et plus particulièrement dans les cas compact et linéaire dans les chapitres quatre et cinq respectivement. Le chapitre six présente les méthodes de Dinkelbach de régularisation proximale et, enfin, nous terminons par la dualité en programmation fractionnelle convexe dans le chapitre huit.

Ce mémoire est basé sur les articles de Crouzeix, Ferland et Schaible, *An Algorithm for Generalized Fractional Programs*, [1], et *A note on an Algorithm for Generalized Fractional Programs*, [2] mais également sur l'article de Strodiot, Crouzeix, Ferland et Nguyen, *An inexact Proximal Point Method for Solving Generalized Fractional Programs*, [3], ainsi que sur base de l'article de Barros, Frenk, Schaible et Zhang, *A New Algorithm for Generalized Fractional Programs*, [4].

Chapitre 1

Convexité et quasiconvexité

Ce chapitre est consacré à la description des fonctions quasiconvexes, quasiconcaves, strictement quasiconvexes et strictement quasiconcaves. Commençons par définir la notion d'ensemble convexe.

Définition 1.0.1. Soit $X \subseteq \mathbb{R}^n$. L'ensemble X est convexe si

$$\forall x, y \in X,\ \forall t \in [0,1],\ tx + (1-t)y \in X.$$

De manière équivalente, en définissant le segment de droite $[x, y]$ comme étant l'ensemble $\{(1-t)x + ty \mid 0 \leq t \leq 1\}$, un ensemble $X \subseteq \mathbb{R}^n$ est *convexe* si le segment de droite joignant toute paire de points de X est dans X.

1.1 Fonctions quasiconvexes et fonctions quasiconcaves

1.1.1 Fonctions quasiconvexes

Définition 1.1.1. Soient $X \subseteq \mathbb{R}^n$ un ensemble convexe et f une fonction définie sur X. La fonction f est quasiconvexe sur X si

$$\forall x_1, x_2 \in X, \forall t \in [0,1],\ f(x_2) \leq f(x_1) \Rightarrow f((1-t)x_1 + tx_2) \leq f(x_1).$$

Exemple 1.1.1.

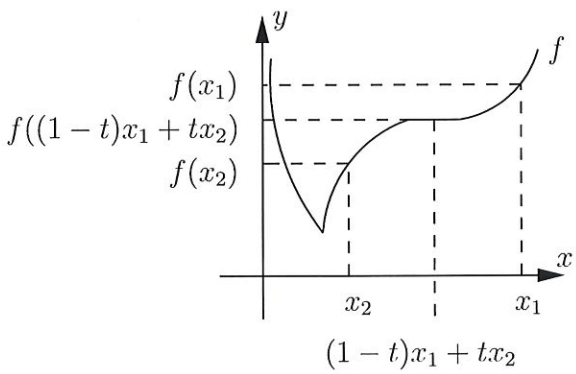

FIG. 1.1 – Illustration d'une fonction quasiconvexe

1.1.2 Fonctions quasiconcaves

Définition 1.1.2. Soient $X \subseteq \mathbb{R}^n$ un ensemble convexe et f une fonction définie sur X. La fonction f est quasiconcave sur X si

$$\forall x_1, x_2 \in X, \forall t \in [0,1],\ f(x_1) \leq f(x_2) \Rightarrow f(x_1) \leq f((1-t)x_1 + tx_2).$$

Exemple 1.1.2.

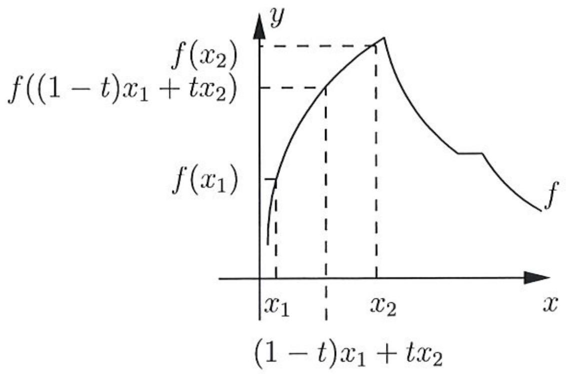

FIG. 1.2 – Illustration d'une fonction quasiconcave

1.1.3 Propriétés

Enonçons, à présent, trois propriétés [6] :

1. Soient $\mathcal{O}$ un ensemble ouvert de $\mathbb{R}^n$ et X un sous-ensemble convexe de $\mathcal{O}$. La fonction $f : \mathcal{O} \to \mathbb{R}$ est quasiconcave sur l'ensemble X si $-f$ est quasiconvexe sur X.

2. Soit f une fonction définie sur un ensemble convexe $X \subseteq \mathbb{R}^n$, soit $L(\alpha) = \{x \in X \mid f(x) \leq \alpha\}$ l'ensemble niveau inférieur et soit $U(\alpha) = \{x \in X \mid f(x) \geq \alpha\}$ l'ensemble niveau supérieur.

   Alors,
$$\begin{aligned} f \text{ quasiconvexe sur } X \quad &\Leftrightarrow \quad L(\alpha) \text{ convexe } \forall \alpha \in \mathbb{R}, \\ f \text{ quasiconcave sur } X \quad &\Leftrightarrow \quad U(\alpha) \text{ convexe } \forall \alpha \in \mathbb{R}. \end{aligned}$$

3. Un rapport de fonctions non linéaires $\frac{\varphi}{\psi}$ est quasiconvexe sur un ensemble convexe S si

$$\varphi \text{ est convexe sur } S,\ \psi > 0 \text{ sur } S$$
$$\text{et}$$
$$[\psi \text{ linéaire sur } \mathbb{R}^n] \text{ ou } [\psi \text{ concave sur } S,\ \varphi \geq 0 \text{ sur } S].$$

1.2 Fonctions strictement quasiconvexes et fonctions strictement quasiconcaves

1.2.1 Fonctions strictement quasiconvexes

Définition 1.2.1. Soient $X \subseteq \mathbb{R}^n$ un ensemble convexe et f une fonction définie sur X. La fonction f est strictement quasiconvexe sur X si

$$\forall x_1, x_2 \in X, \forall t \in (0,1), \; f(x_2) < f(x_1) \Rightarrow f((1-t)x_1 + tx_2) < f(x_1).$$

Exemple 1.2.1.

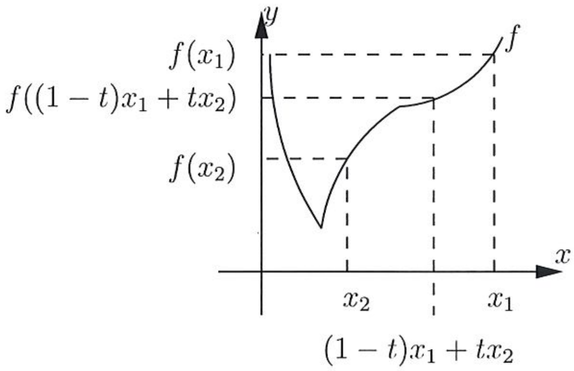

FIG. 1.3 – Illustration d'une fonction strictement quasiconvexe

1.2.2 Fonctions strictement quasiconcaves

Définition 1.2.2. Soient $X \subseteq \mathbb{R}^n$ un ensemble convexe et f une fonction définie sur X. La fonction f est strictement quasiconcave sur X si

$$\forall x_1, x_2 \in X, \forall t \in (0,1), \; f(x_1) < f(x_2) \Rightarrow f(x_1) < f((1-t)x_1 + tx_2).$$

Exemple 1.2.2.

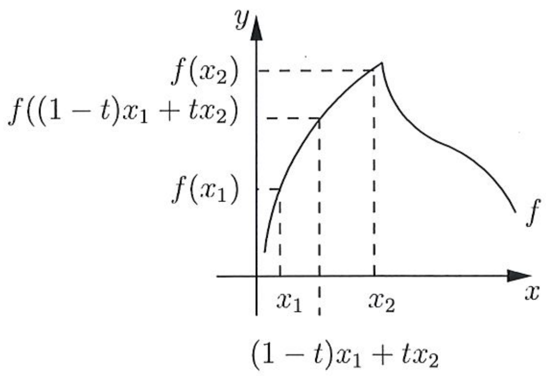

FIG. 1.4 – Illustration d'une fonction strictement quasiconcave

1.2.3 Propriétés

Enonçons, à présent, quatre propriétés [6] :

1. Soient $\mathcal{O}$ un ensemble ouvert de $\mathbb{R}^n$ et X un sous-ensemble convexe de $\mathcal{O}$. La fonction $f : \mathcal{O} \to \mathbb{R}$ est strictement quasiconcave sur l'ensemble X si $-f$ est strictement quasiconvexe sur X.
2. Soit f une fonction semi-continue inférieurement (supérieurement) définie sur un ensemble convexe $X \subseteq \mathbb{R}^n$. Si f est strictement quasiconvexe (strictement quasiconcave) sur X, alors f est quasiconvexe (quasiconcave) sur X, mais pas réciproquement.
3. Soit f une fonction définie sur un convexe $X \subseteq \mathbb{R}^n$ et soit $\bar{x} \in X$ un minimum (maximum) local. Si f est strictement quasiconvexe (strictement quasiconcave) en $\bar{x}$, alors $f(\bar{x})$ est un minimum (maximum) global de f sur X.
   Par conséquent, une fonction strictement quasiconvexe n'est pas toujours quasiconvexe. Par exemple, considérons la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par :
   $$f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon} \end{cases}$$
   Cette fonction est strictement quasiconvexe sur $\mathbb{R}$ mais n'est pas quasiconvexe sur $\mathbb{R}$. En effet, en prenant $x_1 = -1$, $x_2 = 1$ et $\lambda = \frac{1}{2}$, nous avons que

$f(x_2) = 0 = f(x_1)$ mais $f[(1-t)x_1 + tx_2] = 1 > 0 = f(x_1)$. D'où, cette fonction n'est pas quasiconvexe sur $\mathbb{R}$.

4. Une rapport de fonctions linéaires est strictement quasiconvexe et strictement quasiconcave (et donc quasiconvexe et quasiconcave).

Chapitre 2

Méthodes de point proximal approché

L'objectif de ce chapitre est de ramener le problème de la minimisation d'une fonction convexe non différentiable à un problème de minimisation différentiable que nous pourrions résoudre à l'aide des méthodes classiques comme la méthode du gradient ou la méthode BFGS ("Broyden-Fletcher-Goldfarb-Shanno"). Pour cela, nous allons construire une fonction différentiable convexe qui approxime la fonction f convexe non nécessairement différentiable de telle sorte que les minima de f et de son approximation coïncident. Dès lors, la fonction approximante, appelée régularisation de Moreau-Yosida de f, étant convexe et différentiable, nous pourrons utiliser les méthodes d'optimisation classiques.

2.1 Définitions

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe. Considérons la fonction de perturbation quadratique $\tilde{f}_M : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par

$$\tilde{f}_M(x, y) = f(y) + \frac{1}{2} \|y - x\|_M^2$$

où M est une matrice symétrique définie positive de dimension n et $\|z\|_M^2 = z^T M z$.

Définition 2.1.1. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et soit M une matrice symétrique définie positive. La fonction $f_M : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par

$$f_M(x) = \min_{y \in \mathbb{R}^n} \tilde{f}_M(x, y) \tag{2.1}$$

est appelée *régularisation de Moreau-Yosida* de f associée à la métrique M.
L'unique minimum de (2.1) noté par $p_M^f(x)$ est appelé le point proximal de x associé à f et M.

Afin de définir la notion de point stationnaire à ϵ près, définissons les notions de sous-différentiel et de ϵ-sous-différentiel.

Définition 2.1.2. Le sous-différentiel d'une fonction convexe $f : \mathbb{R}^n \rightarrow \mathbb{R}$ au point $x \in \mathbb{R}^n$ est l'ensemble

$$\partial f(x) = \left\{ x^* \in \mathbb{R}^n \mid f(y) \geq f(x) + x^{*T}(y - x) \quad \forall y \in \mathbb{R}^n \right\}.$$

Les éléments de $\partial f(x)$ sont appelés sous-gradients de f en x.

Géométriquement,

$$x^* \in \partial f(x) \Leftrightarrow \forall y \in \mathbb{R}^n \quad f(y) \geq f(x) + x^{*T}(y - x)$$

signifie que x^* est la pente d'une fonction affine qui minimise f et qui passe par le point $(x, f(x))$.
Illustrons la notion de sous-différentiel par un exemple : soit $f(x) = |x|$.

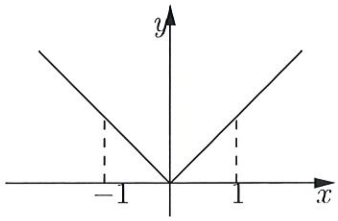

FIG. 2.1 – $f(x) = |x|$

En observant ce graphique, nous pouvons facilement voir que

$$\partial f(0) = [-1, 1], \quad \partial f(x) = \{1\} \text{ si } x > 0 \quad \text{et} \quad \partial f(x) = \{-1\} \text{ si } x < 0.$$

Définition 2.1.3. Soit $\epsilon \geq 0$. Alors, le ϵ-sous-différentiel d'une fonction convexe $f : \mathbb{R}^n \to \mathbb{R}$ au point $x \in \mathbb{R}^n$ est l'ensemble

$$\partial_\epsilon f(x) = \left\{x^* \in \mathbb{R}^n | f(y) \geq f(x) + x^{*T}(y - x) - \epsilon \quad \forall y \in \mathbb{R}^n\right\}.$$

Tout élément $x^* \in \partial_\epsilon f(x)$ est appelé un ϵ-sous-gradient de f en x.

Illustrons la notion de ϵ-sous-différentiel par un exemple : soit $f(x) = |x|$.

$$\partial_\epsilon f(0) = \partial f(0) = [-1, 1].$$

Plus généralement, pour $\epsilon > 0$, nous avons

$$\partial_\epsilon f(x) = \begin{cases} [-1, -1 - \frac{\epsilon}{x}], & \text{si } x < -\frac{\epsilon}{2}, \\ [-1, 1], & \text{si } -\frac{\epsilon}{2} \leq x \leq \frac{\epsilon}{2}, \\ [1 - \frac{\epsilon}{x}, 1], & \text{si } x > \frac{\epsilon}{2}. \end{cases}$$

Définition 2.1.4. Un point $\bar{x}$ est ϵ-stationnaire s'il existe $s \in \partial_\epsilon f(\bar{x})$ avec $\|s\| \leq \epsilon$.

2.2 Méthodes de point proximal

Enonçons tout d'abord le théorème suivant :

Théorème 2.2.1. *Les affirmations suivantes sont équivalentes*

1. x *minimise* f ;
2. $p^f_M(x) = x$;
3. $s^f_M(x) = 0$;
4. x *minimise* f_M ;
5. $f(p^f_M(x)) = f(x)$;
6. $f_M(x) = f(x)$.

Dès lors, nous pouvons constater que minimiser f revient à trouver le point fixe de l'opérateur p^f_M. Par conséquent, l'itération $x^{k+1} = p^f_M(x^k)$ permettra de trouver le minimum de f. Cet algorithme porte le nom d'**Algorithme de point proximal.**

Algorithme de point proximal

1. Choisir $x^0 \in \mathbb{R}^n$ et $M_0 \succ 0$. Prendre $k = 0$.
2. Calculer $x^{k+1} = p^f_{M_k}(x^k)$ en résolvant le problème suivant
$$\min_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{1}{2} \left\| y - x^k \right\|^2_{M_k} \right\}.$$
3. Si $x^{k+1} = x^k$ alors STOP. x^{k+1} est un minimum de f.
4. Choisir $M_{k+1} \succ 0$, remplacer k par $k+1$ et retourner au pas 2.

Afin d'observer la convergence de ce type d'algorithme, supposons que les matrices M_k sont de la forme
$$M_k = \frac{1}{t_k} I \quad \text{où} \quad t_k > 0 \quad \forall k.$$
Notons $\{x^k\}$ la suite générée par l'algorithme de point proximal. Par la définition

de $x^{k+1} = \text{argmin}_y \left\{ f(y) + \frac{1}{2t_k} \left\| y - x^k \right\|^2 \right\}$, nous avons

$$\gamma^k \equiv \frac{1}{t_k}(x^k - x^{k+1}) \in \partial f(x^{k+1}).$$

Par conséquent, l'itération proximale peut s'écrire

$$x^{k+1} = x^k - t_k \gamma^k \quad \text{où} \quad \gamma^k \in \partial f(x^{k+1}).$$

Théorème 2.2.2. *Soit $\{x^k\}$ la suite générée par l'algorithme de point proximal. Si $\sum_{k=0}^{+\infty} t_k = +\infty$, alors*

1. $\lim_{k\to\infty} f(x^k) = f^* = \inf_{x\in I\!R^n} f(x)$.
2. *la suite $\{x^k\}$ converge vers un minimum de f (s'il en existe un).*

2.3 Méthodes de point proximal approché

L'algorithme de point proximal nécessite la résolution du problème de minimisation suivant :

$$\min_{y\in \mathbb{R}^n} \left\{ f(y) + \frac{1}{2} \left\| y - x^k \right\|^2_{M_k} \right\}. \tag{2.2}$$

Celui-ci est généralement aussi complexe à résoudre que le problème de départ et se résout à l'aide d'une méthode itérative. Deux types d'itérations interviendront dans l'algorithme utilisé pour résoudre ce type de problème :

- les itérations dites internes à savoir les itérations nécessaires lors de la résolution approximative du sous-problème (2.2) ;
- les itérations dites externes à savoir les itérations donnant x^{k+1} à partir de x^k.

L'élaboration d'un tel algorithme nécessite la présence d'un critère d'arrêt qui empêchera l'algorithme de boucler dans les itérations internes sans détruire la convergence des itérations externes.

2.3.1 Critères d'arrêt

Avant d'aborder l'algorithme proprement dit, déterminons tout d'abord le critère d'arrêt. Le premier critère que nous allons présenter ne sera que purement théorique tandis que le second pourra être implémenté.

Critère d'arrêt conceptuel

Supposons qu'à l'itération k, nous disposons du point x^k et d'une matrice symétrique définie positive M_k, la boucle interne de l'algorithme qui calcule $f_{M_k}(x^k)$ et $p^f_{M_k}(x^k)$ s'arrêtera au point x^{k+1} lorsqu'il vérifiera la condition suivante :

$$f(x^{k+1}) \leq f(x^k) - m\delta^k, \tag{2.3}$$

où $m \in (0, 1)$ est une constante et

$$\delta^k = f(x^k) - f_{M_k}(x^k). \tag{2.4}$$

Notons que $\delta^k \geq 0$. En effet, $f_M \leq f$ par le premier point du théorème suivant :

Théorème 2.3.1. *La régularisation de Moreau-Yosida f_M possède les propriétés suivantes :*

1. *f_M est convexe et $f_M \leq f$ sur $I\!R^n$.*
2. *f_M est différentiable et son gradient est donné par*

$$\nabla f_M(x) = s^f_M(x) = M(x - p^f_M(x)) \in \partial f(p^f_M(x)). \tag{2.5}$$

3. *∇f_M est Lipschitzienne sur $I\!R^n$ de constante $\lambda_{max}(M)$, i.e. ,*

$$\forall x, y \in I\!R^n \qquad \|\nabla f_M(x) - \nabla f_M(y)\| \leq \lambda_{max}(M)\, \|x - y\|\,.$$

Notons également que, par le théorème 2.2.1, δ^k sera nul si et seulement si x^k minimise f puisque x^k minimise f est équivalent à $f_M(x^k) = f(x^k)$.

Avec ce critère d'arrêt (2.3), énonçons à présent deux résultats :

Théorème 2.3.2. *Soit $M_k \succ 0$ et supposons que x^k n'est pas un minimum de f. Soit $\{y^j\} \subseteq I\!R^n$ une suite convergeant vers $p^f_{M_k}(x^k)$. Alors, il existe $j \in I\!N$ tel que $x^{k+1} = y^j$ satisfait la condition suivante :*

$$f(x^{k+1}) \leq f(x^k) - m\delta^k.$$

Dès lors, la boucle interne de l'algorithme se terminera après un nombre fini d'itérations.
Le résultat suivant nous permettra de conclure que la boucle externe de l'algorithme convergera.

Théorème 2.3.3. *Supposons que la suite $\{M_k\}$ de matrices symétriques définies positives satisfait*

$$\sum_{k=0}^{+\infty} \frac{1}{\lambda_{max}(M_k)} = +\infty. \tag{2.6}$$

Supposons aussi que la suite $\{x^k\}$ est bornée et satisfait la condition suivante pour tout k :

$$f(x^{k+1}) \leq f(x^k) - m\delta^k$$

où $\delta_k = f(x^k) - f_{M_k}(x^k)$.
Alors, tout point limite de $\{x^k\}$ minimise f et $f(x^k) \to \min_x f(x)$.

L'égalité (2.6) nous permet de dire que la suite des valeurs propres de $\{M_k\}$ peut croître mais pas trop rapidement. Illustrons cela par l'exemple suivant : si $M_k = \frac{1}{t_k} I$ alors

$$\sum_{k=0}^{+\infty} \frac{1}{\lambda_{max}(M_k)} = +\infty$$

est équivalent à

$$\sum_k t_k = +\infty$$

à condition que $t_k = t > 0$ pour tout k.

Ce critère d'arrêt est uniquement conceptuel car la fonction $f_{M_k}(x^k)$ ne peut pas être calculée exactement. Il est alors nécessaire de déterminer un critère d'arrêt implémentable qui impliquerait la condition (2.3).

Critère d'arrêt pratique

La stratégie utilisée pour déterminer le critère d'arrêt pratique est de remplacer la fonction f par une fonction $\varphi \leq f$ pour laquelle le calcul du point proximal est aisée. Si nous construisons une fonction φ proche assez de f au point proximal π, alors le calcul du point proximal associé au modèle φ est suffisant pour obtenir la convergence décrite dans le théorème 2.3.3.

Théorème 2.3.4. *Soient x^k, M_k et $\varphi : I\!R^n \to I\!R$ une fonction convexe minimisant f. Définissons $\pi = p^{\varphi}_{M_k}(x^k)$. Alors*

$$f(\pi) \leq f(x^k) - m(f(x^k) - \varphi(\pi)) \tag{2.7}$$

implique la condition suivante avec $x^{k+1} = \pi$

$$f(x^{k+1}) \leq f(x^k) - m\delta^k.$$

De plus, si x^k ne minimise pas f alors il existe $\rho^k > 0$ indépendant de φ et π tel que

$$f(\pi) - \varphi(\pi) \leq \rho^k \Rightarrow (2.7).$$

Lien avec les méthodes de faisceaux

Choisissons une métrique M de la forme $\gamma_k I$ et la fonction modèle φ^k construite à partir de la linéarisation de f aux points proximaux approchés. Cette fonction est alors de la forme

$$\varphi^k(x) = \max_{j \in J_k} \{ f(x^k) + (s^j)^T(x - x^k) - \alpha_j^k \} \quad \forall x \in \mathbb{R}^n.$$

Dès lors, le calcul du point proximal

$$p^{\varphi^k}_{\gamma^k I}(x^k) = \text{argmin}_x \left\{ \varphi^k(x) + \frac{\gamma_k}{2} \left\| x - x^k \right\|^2 \right\}$$

est équivalent à la résolution du problème de programmation quadratique suivant :

$$QP(x^k, \gamma_k) \left\{ \begin{array}{ll} \min_{v,x} & v + \frac{\gamma_k}{2} \left\| x - x^k \right\|^2 \\ s.c. & -\alpha_j^k + (s^j)^T (x - x^k) \leq v, \ j \in J_k \end{array} \right.$$

2.3.2 Description de la méthode

Soient donnés x^k et M_k. Notons les itérations de la boucle interne par j et supprimons l'index k de la boucle externe afin de faciliter les notations. Dès lors, utilisons les notations suivantes :

$$\begin{aligned} y^j &:= p^{\varphi^j}_M(x) \\ g^j &:= s^{\varphi^j}_M(x) = M(x - y^j) \in \partial \varphi^j(y^j). \end{aligned} \tag{2.8}$$

A l'étape k, la boucle interne aura pour objectif de construire une fonction φ telle que la différence $f(\pi) - \varphi(\pi)$ soit suffisamment petite (pour rappel, $\pi = p^{\varphi}_{M_k}(x^k)$).

La perturbation quadratique de toute fonction convexe φ^j peut s'écrire comme suit :

$$\tilde{\varphi}^j(y) := \tilde{\varphi}^j_M(x, y) = \varphi^j(y) + \frac{1}{2} \left\| y - x \right\|^2_M \quad \forall y \in \mathbb{R}^n.$$

Notre but est de construire une suite de modèles $\{\varphi^j\}$ telle que $f(y^j) - \varphi^j(y^j) \to 0$. Par le théorème 2.3.4, la condition d'arrêt pratique (2.7) sera satisfaite après un nombre fini de pas si x ne minimise pas f. Si x minimise f alors nous aurons que $y^j \to x$.

Introduisons, à présent, la fonction aggrégée $l^j : \mathbb{R}^n \to \mathbb{R}$ définie par

$$l^j(y) := \varphi^j(y^j) + \left\langle g^j, y - y^j \right\rangle. \tag{2.9}$$

Comme $g^j \in \partial \varphi^j(y^j)$, nous avons que $\varphi^j(y) \geq \varphi^j(y^j) + g^{jT}(y - y^j)$ pour tout $y \in \mathbb{R}^n$ et donc que $l^j \leq \varphi^j$.

Imposons trois conditions sur le modèle φ^j pour tout j :

$$\begin{aligned} &\varphi^{j+1}(y) \leq f(y) && \forall y \in \mathbb{R}^n, \\ &\varphi^{j+1}(y) \geq f(y^j) + \left\langle s(y^j), y - y^j \right\rangle && \forall y \in \mathbb{R}^n, \\ &\varphi^{j+1}(y) \geq l^j(y) && \forall y \in \mathbb{R}^n, \end{aligned} \tag{2.10}$$

où $s(y^j)$ est le sous-gradient de f au point y^j. La première condition est la propriété habituelle de minimisation, les deux autres conditions montrent comment le nouveau modèle doit être construit une fois que le point y^j a été calculé.
Imposons la condition suivante sur la matrice M :

$$0 \prec M \preceq \nu I \quad \text{où} \quad \nu > 0 \tag{2.11}$$

pour que la suite $\{M_k\}$ vérifie la condition

$$\sum_{k=0}^{+\infty} \frac{1}{\lambda_{max}(M_k)} = +\infty$$

afin d'obtenir la convergence de la suite x_k vers le minimum de f.

Théorème 2.3.5. *Soit $x \in I\!R^n$ et supposons que les fonctions convexes $\{\varphi^j\}$, la métrique M et les points tests $\{y^j\}$ satisfont les propriétés suivantes*

$$\begin{aligned}
&y^j := p_M^{\varphi^j}(x), g^j := s_M^{\varphi^j}(x) = M(x - y^j) \in \partial\varphi^j(y^j)\\
&l^j(y) = \varphi^j(y^j) + \langle g^j, y - y^j\rangle\\
&\varphi^{j+1}(y) \leq f(y) \quad \forall y \in I\!R^n\\
&\varphi^{j+1}(y) \geq f(y^j) + \langle s(y^j), y - y^j\rangle \quad \forall y \in I\!R^n\\
&\varphi^{j+1}(y) \geq l^j(y) \quad \forall y \in I\!R^n\\
&0 \prec M \preceq \nu I \quad \text{où} \quad \nu > 0
\end{aligned}$$

Alors

$$f(y^j) - \varphi^j(y^j) \to 0, \tag{2.12}$$

$$y^j \to p_M^f(x). \tag{2.13}$$

Enonçons, à présent, une condition nécessaire et suffisante pour que x minimise f.

Corollaire 2.3.1. *Sous les conditions du théorème précédent,*

$$x \text{ minimise } f$$

$$\Leftrightarrow$$

$$y^j \to x.$$

2.3.3 Algorithme de point proximal approché

Soient $\nu_{max} > 0$, $\epsilon_{tol} \geq 0$, $m \in (0, 1)$ et un point de départ $x^0 \in \mathbb{R}^n$.

1. **Pour** $k = 0, 1, \ldots$ **faire**
2. Choisir M_k tel que $0 \prec M_k \prec \nu_{max} I$.
3. Choisir une fonction modèle φ^1 minimisant f telle que $\varphi^1(x^k) = f(x^k)$.
4. Prendre $j = 0$.
5. **Répéter**
6. $j \leftarrow j + 1$
7. Calculer $y^j = \operatorname{argmin}_y \left\{ \varphi^j(y) + \frac{1}{2} \left\| y - x^k \right\|_{M_k}^2 \right\}$
   et le sous-gradient $g^j = s_{M_k}^{\varphi^j}(x^k)$.
8. **Si** $f(y^j) - \varphi^j(y^j) \leq \epsilon_{tol}$ et $\|g^j\| \leq \epsilon_{tol}$ **alors**
9. Prendre $\bar{x} = y^j$ et **sortir** $\Rightarrow \bar{x}$ est un point stationnaire à ϵ_{tol} près.
10. **Fin du si**
11. Choisir une fonction modèle φ^{j+1} satisfaisant les conditions suivantes

$$\varphi^{j+1}(y) \leq f(y) \qquad \forall y \in \mathbb{R}^n$$
$$\varphi^{j+1}(y) \geq f(y^j) + \langle s(y^j), y - y^j \rangle \qquad \forall y \in \mathbb{R}^n$$
$$\varphi^{j+1} \geq l^j(y) \qquad \forall y \in \mathbb{R}^n$$

12. **jusqu'à ce que** $f(y^j) \leq f(x^k) - m\left[f(x^k) - \varphi^j(y^j)\right]$
13. Prendre $x^{k+1} = y^j$
14. **Fin du pour**

2.3.4 Convergence

Théorème 2.3.6. *Soit $\{x^k\}$ une suite générée par l'algorithme avec $\epsilon_{tol} = 0$. Si la suite $\{x^k\}$ est finie et si x^K est le dernier élément de la suite, alors x^K minimise f. Sinon, si $\{x^k\}$ est bornée, alors tous ses points limites minimisent f et $f(x^k) \to \min\limits_x f(x)$.*

Démonstration.

- Supposons d'abord que la suite $\{x^k\}$ est finie. Considérons alors la suite $\{y^j\}$ associée au dernier élément x^K de la suite $\{x^k\}$. Cette suite $\{y^j\}$ doit être infinie. En effet, lorsque j est fini, la condition de sortie de l'algorithme $f(y^j) - \varphi^j(x^j) \leq \epsilon_{tol} = 0$ n'est pas vérifiée car $\varphi \leq f$ et $f(y^j) - \varphi^j(y^j) \to 0$ uniquement lorsque $j \to 0$. Par conséquent, la boucle "répéter" se termine à l'étape 12 de l'algorithme et nous passons à l'étape 13. Ce qui est impossible car x^K est le dernier élément de la suite $\{x^k\}$.

  Par le théorème 2.3.5, nous avons que
  $$f(y^j) - \varphi^j(y^j) \to 0$$
  et que x^K est un point stationnaire.

  Supposons maintenant que x^K ne minimise pas f. Dès lors, par le théorème 2.3.4, il existe $\rho^K > 0$ indépendant de φ et de π tel que
  $$f(\pi) - \varphi(\pi) \leq \rho^K \Rightarrow f(\pi) \leq f(x^K) - m(f(x^K) - \varphi(\pi)).$$
  Par conséquent, le critère d'arrêt donné à la ligne 12 de l'algorithme est satisfait pour un certain j. Ce qui contredit le fait que la suite $\{y^j\}$ est infinie. x^K minimise donc f lorsque la suite $\{x^k\}$ est finie.

- Supposons ensuite que la suite $\{x^k\}$ est infinie et bornée. Par construction, nous avons que
  $$f(x^{k+1}) \leq f(x^k) - m\delta^k.$$
  Comme $M_k \preceq \nu_{max} I$ par l'étape 2 de l'algorithme, nous avons également que
  $$\sum_{k=0}^{\infty} \frac{1}{\lambda_{max}(M_k)} = +\infty.$$

Dès lors, en appliquant le théorème 2.3.3, tout point limite de $\{x^k\}$ minimise f et $f(x^k) \to \min_x f(x)$.

□

Théorème 2.3.7. *Supposons que f est bornée inférieurement et soit $\epsilon_{tol} > 0$. Alors, l'algorithme de point proximal approché s'arrête après un nombre fini d'itérations avec un point stationnaire $\bar{x}$ à ϵ_{tol} près.*

Démonstration.
Pour tout $y \in \mathbb{R}^n$, nous avons par (2.10)

$$\varphi^j(y) \leq f(y). \tag{2.14}$$

Par (2.8), nous savons que

$$g^j \in \partial\varphi^j(y^j).$$

Dès lors, par la définition de sous-différentiel,

$$\varphi^j(y) \geq \varphi^j(y^j) + \langle g^j, y - y^j \rangle \qquad \forall y \in \mathbb{R}^n. \tag{2.15}$$

En combinant (2.14) et (2.15), nous obtenons donc

$$f(y) \geq \varphi^j(y^j) + \langle g^j, y - y^j \rangle \qquad \forall y \in \mathbb{R}^n.$$

En utilisant un artifice de calcul, nous pouvons écrire

$$f(y) \geq f(y^j) + \langle g^j, y - y^j \rangle - \underbrace{(f(y^j) - \varphi^j(y^j))}_{\overset{not}{=}\eta_j}.$$

En utilisant la définition de ϵ-sous-différentiel,

$$g^j \in \partial_{\eta_j} f(y^j) \quad \text{où} \quad \eta_j = f(y^j) - \varphi^j(y^j). \tag{2.16}$$

Par conséquent, si l'algorithme s'arrête à l'étape 8 c'est-à-dire si $f(y^j) - \varphi^j(y^j) \leq \epsilon_{tol}$ et si $\|g^j\| \leq \epsilon_{tol}$, alors nous prendrons $\bar{x} = y^j$ et $\bar{x}$ sera un point stationnaire à ϵ_{tol} près puisque $\|g^j\| \leq \epsilon_{tol}$ et que

$$\begin{aligned} g^j \in \partial_{\eta_j} f(y^j) \quad &\Leftrightarrow f(y) \geq f(y^j) + \langle g^j, y - y^j \rangle - \eta_j \\ &\Leftrightarrow f(y) \geq f(y^j) + \langle g^j, y - y^j \rangle - \epsilon_{tol} \qquad \text{car } \eta_j \leq \epsilon_{tol} \\ &\Leftrightarrow g^j \in \partial_{\epsilon_{tol}} f(y^j). \end{aligned}$$

Il nous reste donc à montrer que l'algorithme s'arrête après un nombre fini d'itérations. Pour cela, supposons que le critère d'arrêt de l'étape 8 n'est jamais vérifié c'est-à-dire que $f(y^j) - \varphi^j(y^j) > \epsilon_{tol}$ et $\|g^j\| > \epsilon_{tol}$ (HA). Séparons le cas où la suite $\{x^k\}$ est finie et le cas où elle est infinie :

- Supposons d'abord que la suite $\{x^k\}$ est finie et considérons la suite infinie $\{y^j\}$ associée à la dernière itération externe K. Cette suite est infinie car sinon l'algorithme ne s'arrêterait pas à l'itération K.

  Par le théorème 2.3.5, nous pouvons dire que

  $$f(y^j) - \varphi^j(y^j) = \eta_j \to 0.$$

  Supposons à présent que x^K ne minimise pas f. Par le théorème 2.3.4, il existe $\rho^K > 0$ indépendant de φ et de π tel que

  $$f(\pi) - \varphi(\pi) \leq \rho^K \Rightarrow f(\pi) \leq f(x^K) - m(f(x^K) - \varphi(\pi)).$$

  Par conséquent, le critère d'arrêt donné à la ligne 12 de l'algorithme est satisfait pour un certain j. Ce qui contredit le fait que la suite $\{y^j\}$ est infinie. x^K minimise donc f.

  Comme x^K minimise f, par le corollaire 2.3.1,

  $$y^j \to x^K \tag{2.17}$$

  Dès lors, comme $y^j \to x^K$ et $\|g^j\| = \left\|M_K(x^K - y^j)\right\| \leq \nu_{max} \left\|x^K - y^j\right\|$,

  $$g^j \to 0. \tag{2.18}$$

  Par (2.17) et (2.18), $\eta_j \leq \epsilon_{tol}$ et $\|g^j\| \leq \epsilon_{tol}$, ce qui contredit l'hypothèse de l'absurde (HA). Le critère d'arrêt de la ligne 8 de l'algorithme sera donc satisfait pour un certain j.

- Supposons ensuite que la suite $\{x^k\}$ est infinie et notons par $\bar{g}^k$, $\bar{\eta}^k$, et $\bar{\varphi}^k$, les derniers g^j, η^j et φ^j de l'itération externe k. La condition d'arrêt de la ligne 12 de l'algorithme étant satisfaite (c'est-à-dire la condition (2.7) avec $\pi = x^{k+1}$), nous pouvons écrire

  $$f(x^{k+1}) - f(x^k) \leq m(\bar{\varphi}(x^{k+1}) - f(x^k)) \leq -m\delta^k \leq 0$$

Dès lors, la suite $\{f(x^k)\}$ est une suite décroissante et bornée inférieurement (puisque f est bornée inférieurement). Cette suite est donc convergente. D'où

$$\bar{\varphi}^k(x^{k+1}) - f(x^k) \to 0. \tag{2.19}$$

La suite $\{f(x^k)\}$ étant décroissante,

$$f(x^k) \geq f(x^{k+1}).$$

En soustrayant $\bar{\varphi}^k(x^{k+1})$ des deux côtés de l'inégalité,

$$\underbrace{f(x^k) - \bar{\varphi}^k(x^{k+1})}_{\overset{(2.12)}{\to 0}} \geq f(x^{k+1}) - \bar{\varphi}^k(x^{k+1}) \overset{def}{=} \bar{\eta}^k \geq 0.$$

Par le théorème de l'étau, nous pouvons conclure que $\bar{\eta}^k \to 0$.

D'un autre côté,

$$f(x^k) \geq \bar{\varphi}^k(x^{k+1})$$

En soustrayant $\bar{\varphi}^k(x^k)$ des deux côtés de l'inégalité,

$$\begin{aligned}
\underbrace{f(x^k) - \bar{\varphi}^k(x^{k+1})}_{\overset{(2.12)}{\to 0}} &\geq \bar{\varphi}^k(x^k) - \bar{\varphi}^k(x^{k+1}) \\
&\geq \langle \bar{g}^k, x^k - x^{k+1} \rangle && (\text{car } \bar{g}^k \in \partial\bar{\varphi}^k(x^{k+1})) \\
&= \left\|\bar{g}^k\right\|^2_{M_k^{-1}} \geq \nu_{max}^{-1} \left\|\bar{g}^k\right\|^2 \geq 0 && (\text{car } M_k^{-1} \succeq \nu_{max}^{-1} I)
\end{aligned}$$

Par (2.19) et les inégalités précédentes,

$$\bar{g}^k \to 0.$$

Comme dans le cas précédent, la condition d'arrêt est satisfaite et l'hypothèse de l'absurde (HA) est donc contredite. Par conséquent, le critère d'arrêt de la ligne 8 de l'algorithme est donc satisfait pour un certain j.

□

Chapitre 3

Programmation fractionnelle généralisée

3.1 Description du problème

Soient $S \subseteq \mathbb{R}^n$ un ensemble non vide, des fonctions f_i continues sur S et des fonctions g_i continues et positives sur S. Le problème fractionnel généralisé s'écrit comme suit :

$$(P) \qquad \bar{\theta} = \inf_{x \in S} \left(\max_{1 \leq i \leq p} \frac{f_i(x)}{g_i(x)} \right).$$

Notons que, lorsque p vaut 1, le problème devient

$$\bar{\theta} = \inf_{x \in S} \frac{f(x)}{g(x)}$$

et correspond à la programmation fractionnelle.

La méthode que nous allons décrire pour résoudre ce genre de problème nécessite de trouver les racines de l'équation suivante :

$$F(\theta) = 0$$

où $F(\theta)$ est la valeur optimale du problème paramétrique décrit par

$$(P_\theta) \qquad \inf_{x\in S}\left(\max_{1\leq i\leq p}[f_i(x)-\theta g_i(x)]\right).$$

L'utilité de résoudre ce problème (P_θ) plutôt que le problème (P) réside dans le fait que la structure de (P_θ) est plus simple que celle de (P). Citons, par exemple, le cas où les fonctions f_i sont non négatives et convexes sur S, où les fonctions g_i sont concaves sur S et où S est un ensemble convexe ; nous aurons alors affaire à un problème (P_θ) convexe pour toute valeur positive de θ alors que (P) n'est que quasiconvexe.

L'application de l'algorithme, que nous décrirons par la suite, fournit les solutions optimales et les valeurs optimales du problème paramétrique (P_θ).

3.2 Notations

Pour $\theta \in \mathbb{R}$, notons $M(\theta)$ l'ensemble des solutions optimales (qui peut être vide) de (P_θ) c'est-à-dire

$$M(\theta)=\left\{x\in S \;\middle|\; F(\theta)=\max_{1\leq i\leq p}[f_i(x)-\theta g_i(x)]\right\}.$$

Notons également, pour $x \in S$,

$$\begin{aligned}
\underline{g}(x) &= \min_{1\leq i\leq p} g_i(x),\\
\bar{g}(x) &= \max_{1\leq i\leq p} g_i(x),\\
I(x,\theta) &= \{i \mid f_i(x)-\theta g_i(x)=F(\theta)\},\\
J(x) &= \left\{j \;\middle|\; \frac{f_j(x)}{g_j(x)}=\max_{1\leq i\leq p}\frac{f_i(x)}{g_i(x)}\right\}.
\end{aligned}$$

3.3 Propriétés de la fonction F

Etudions les propriétés de la fonction F dans le cas le plus général c'est-à-dire lorsque les fonctions f_i et g_i sont des fonctions arbitraires et continues et lorsque S est un ensemble arbitraire de $\mathbb{R}^n$.

Propriété 3.3.1. *F est décroissante et semi-continue supérieurement.*

Démonstration.

Les fonctions g_i étant positives, la fonction F sera monotone.

La fonction $\max_i[f_i(x) - \theta g_(x)]$ est continue en (x, θ). Par conséquent, la fonction F sera semi-continue supérieurement en θ. □

Propriété 3.3.2. *$F(\theta) < 0 \Leftrightarrow \theta > \bar{\theta}$; d'où $F(\bar{\theta}) \geq 0$.*

Démonstration.

Montrons d'abord que

$$F(\theta) < 0 \Rightarrow \theta > \bar{\theta}.$$

Supposons alors que $F(\theta) < 0$. Dès lors,

$$\begin{aligned}
&\inf_{x \in S}\left(\max_{1 \leq i \leq p}[f_i(x) - \theta g_i(x)]\right) < 0 \\
\Leftrightarrow\ &\exists\ \hat{x} \in S \text{ t.q. } \max_{1 \leq i \leq p}[f_i(\hat{x}) - \theta g_i(\hat{x})] < 0 \\
\Leftrightarrow\ &\exists\ \hat{x} \in S \text{ t.q. } f_i(\hat{x}) - \theta g_i(\hat{x}) < 0 \qquad i = 1, \ldots, p.
\end{aligned}$$

Par conséquent, comme les fonctions g_i sont positives,

$$\begin{aligned}
&\exists\ \hat{x} \in S \text{ t.q. } \frac{f_i(\hat{x})}{g_i(\hat{x})} < \theta \qquad i = 1, \ldots, p \\
\Leftrightarrow\ &\exists\ \hat{x} \in S \text{ t.q. } \max_{1 \leq i \leq p} \frac{f_i(\hat{x})}{g_i(\hat{x})} < \theta.
\end{aligned}$$

$\bar{\theta}$ étant la valeur qui minimise $\max_{1 \leq i \leq p} \dfrac{f_i(\hat{x})}{g_i(\hat{x})}$,

$$\bar{\theta} \leq \max_{1 \leq i \leq p} \frac{f_i(\hat{x})}{g_i(\hat{x})} < \theta.$$

Montrons maintenant que

$$\bar{\theta} < \theta \Rightarrow F(\theta) < 0.$$

Supposons donc que $\bar{\theta} < \theta$. Par conséquent, il existe une valeur $\hat{x} \in S$ qui minimise la fonction $F(\theta)$:

$$\begin{aligned} & \exists\, \hat{x} \in S \text{ t.q. } \max_{1\leq i\leq p} \frac{f_i(\hat{x})}{g_i(\hat{x})} < \theta \\ \Leftrightarrow\ & \exists\, \hat{x} \in S \text{ t.q. } \frac{f_i(\hat{x})}{g_i(\hat{x})} < \theta \qquad i = 1, \ldots, p \\ \Leftrightarrow\ & \exists\, \hat{x} \in S \text{ t.q. } f_i(\hat{x}) - \theta g_i(\hat{x}) < 0 \qquad i = 1, \ldots, p \qquad (\text{car } g_i \text{ positives}) \\ \Leftrightarrow\ & \exists\, \hat{x} \in S \text{ t.q. } \max_{1\leq i\leq p} [f_i(\hat{x}) - \theta g_i(\hat{x})] < 0. \end{aligned}$$

Dès lors,

$$F(\theta) = \inf_{x\in S} \left(\max_{1\leq i\leq p} [f_i(\hat{x}) - \theta g_i(\hat{x})] \right) < 0.$$

Il nous reste à montrer que

$$F(\bar{\theta}) \geq 0.$$

Supposons par l'absurde que $F(\bar{\theta}) < 0$. Par la condition nécessaire de la propriété 3.3.2 démontrée ci-dessus, nous concluons que

$$\bar{\theta} > \bar{\theta}.$$

Cette dernière inégalité est absurde. Par conséquent, $F(\bar{\theta}) \geq 0$. □

Propriété 3.3.3. *Si (P) a une solution optimale, alors $F(\bar{\theta}) = 0$.*

Démonstration.

Désignons par $\bar{x}$ une solution optimale de (P). Dès lors, $\bar{x} \in S$ et nous pouvons écrire

$$\bar{\theta} = \max_{1\leq i\leq p} \frac{f_i(\bar{x})}{g_i(\bar{x})}. \tag{3.1}$$

D'une part, notons par î l'indice pour lequel le maximum est atteint. Dès lors,

$$\begin{aligned} & \bar{\theta} = \tfrac{f_{\hat{\imath}}(\bar{x})}{g_{\hat{\imath}}(\bar{x})} \\ \Leftrightarrow\quad & f_{\hat{\imath}}(\bar{x}) - \bar{\theta} g_{\hat{\imath}}(\bar{x}) = 0. \end{aligned} \tag{3.2}$$

D'autre part, par (3.1),

$$\begin{aligned} &\bar{\theta} \geq \tfrac{f_i(\bar{x})}{g_i(\bar{x})} && i = 1, \ldots, p \\ \Leftrightarrow\ & f_i(\bar{x}) - \bar{\theta} g_i(\bar{x}) \leq 0 && i = 1, \ldots, p \quad \text{(car } g_i \text{ positives).} \end{aligned} \tag{3.3}$$

Par conséquent, par (3.2) et (3.3), nous pouvons conclure

$$\max_{1 \leq i \leq p} \left[f_i(\bar{x}) - \bar{\theta} g_i(\bar{x}) \right] = 0.$$

Comme, par la propriété 3.3.2, $F(\bar{\theta}) \geq 0$, nous obtenons

$$0 \leq F(\bar{\theta}) = \inf_{x \in S} \left(\max_{1 \leq i \leq p} [f_i(x) - \bar{\theta} g_i(x)] \right) = 0.$$

D'où, $F(\bar{\theta}) = 0$. □

Propriété 3.3.4. *Si $F(\bar{\theta}) = 0$, alors les problèmes (P) et $(P_{\bar{\theta}})$ possèdent le même ensemble de solutions optimales (qui peut être vide).*

Démonstration.

Notons par $\bar{x}$ une solution optimale de (P) et montrons que $\bar{x}$ est également une solution optimale de $(P_{\bar{\theta}})$. Par la propriété 3.3.3, le problème (P) ayant une solution optimale $\bar{x}$, nous avons que $F(\bar{\theta}) = 0$. Par conséquent, $\bar{x}$ est également une solution optimale de $(P_{\bar{\theta}})$.

Montrons maintenant que toute solution optimale de $(P_{\bar{\theta}})$ est solution de (P). Pour ce faire, supposons que $F(\bar{\theta}) = 0$ et désignons par $\bar{x}$ une solution optimale de $(P_{\bar{\theta}})$. Alors,

$$\begin{aligned} & F(\bar{\theta}) = 0 \\ \Leftrightarrow\ & \inf_{x \in S} \left(\max_{1 \leq i \leq p} [f_i(x) - \bar{\theta} g_i(x)] \right) = 0 \\ \Leftrightarrow\ & \exists\, \bar{x} \in S \text{ t.q. } \max_{1 \leq i \leq p} [f_i(\bar{x}) - \bar{\theta} g_i(\bar{x})] = 0 \\ \Leftrightarrow\ & \exists\, \bar{x} \in S \text{ t.q. } f_i(\bar{x}) - \bar{\theta} g_i(\bar{x}) \leq 0 \qquad i = 1, \ldots, p \\ \Leftrightarrow\ & \exists\, \bar{x} \in S \text{ t.q. } \frac{f_i(\bar{x})}{g_i(\bar{x})} \leq \bar{\theta} \qquad i = 1, \ldots, p \quad \text{(car } g_i \text{ positives).} \end{aligned}$$

Par conséquent,

$$\exists\, \bar{x} \in S \text{ t.q. } \max_{1\leq i\leq p} \frac{f_i(\bar{x})}{g_i(\bar{x})} \leq \bar{\theta}.$$

Or,

$$\bar{\theta} = \inf_{x\in S}\left(\max_{1\leq i\leq p} \frac{f_i(x)}{g_i(x)}\right) \leq \max_{1\leq i\leq p} \frac{f_i(\bar{x})}{g_i(\bar{x})} \leq \bar{\theta}.$$

Alors,

$$\bar{\theta} = \max_{1\leq i\leq p} \frac{f_i(\bar{x})}{g_i(\bar{x})}.$$

Nous pouvons alors conclure que $\bar{x}$ est également une solution optimale du problème (P). □

Proposition 3.3.1. *Soit θ tel que $F(\theta)$ est finie et $M(\theta)$ est non vide. Soit $\bar{x} \in M(\theta)$. Alors,*

$$F(\mu) \leq F(\theta) + (\theta-\mu)\underline{g}(\bar{x}), \qquad si \quad \mu > \theta, \tag{3.4}$$

$$F(\mu) \leq F(\theta) + (\theta-\mu)\bar{g}(\bar{x}), \qquad si \quad \mu < \theta. \tag{3.5}$$

Démonstration.

Puisque $\bar{x} \in M(\theta)$ c'est-à-dire puisque $\bar{x}$ est une solution optimale de (P_θ), nous avons

$$\begin{aligned} F(\theta) &= \max_{1\leq i\leq p}[f_i(\bar{x}) - \theta g_i(\bar{x})] \\ &\geq f_i(\bar{x}) - \theta g_i(\bar{x}) \qquad i = 1\ldots,p \\ &= f_i(\bar{x}) - \theta g_i(\bar{x}) - \mu g_i(\bar{x}) + \mu g_i(\bar{x}) \\ &= -\mu g_i(\bar{x}) + f_i(\bar{x}) - (\theta-\mu) g_i(\bar{x}) \qquad i = 1,\ldots,p. \end{aligned}$$

Dès lors,

$$F(\theta) + (\theta-\mu)g_i(\bar{x}) \geq f_i(\bar{x}) - \mu g_i(\bar{x}) \qquad i = 1,\ldots,p.$$

Supposons tout d'abord que $\mu > \theta$. Par conséquent, $\theta - \mu < 0$ et puisque les fonctions g_i sont positives, nous obtenons

$$F(\theta) + (\theta-\mu)\underline{g}(\bar{x}) \geq F(\theta) + (\theta-\mu)g_i(\bar{x}) \geq f_i(\bar{x}) - \mu g_i(\bar{x}) \qquad i = 1,\ldots,p.$$

Donc,

$$\begin{aligned} F(\theta) + (\theta - \mu)\underline{g}(\bar{x}) &\geq \max_{1\leq i\leq p}[f_i(\bar{x}) - \mu g_i(\bar{x})] \\ &\geq F(\mu). \end{aligned}$$

Supposons maintenant que $\mu < \theta$. Par conséquent, $\theta - \mu > 0$ et puisque les fonctions g_i sont positives, nous pouvons écrire

$$F(\theta) + (\theta - \mu)\bar{g}(\bar{x}) \geq F(\theta) + (\theta - \mu)g_i(\bar{x}) \geq f_i(\bar{x}) - \mu g_i(\bar{x}) \qquad i = 1, \dots, p.$$

Donc,

$$\begin{aligned} F(\theta) + (\theta - \mu)\bar{g}(\bar{x}) &\geq \max_{1\leq i\leq p}[f_i(\bar{x}) - \mu g_i(\bar{x})] \\ &\geq F(\mu). \end{aligned}$$

□

Remarque 3.3.1. Lorsque $p = 1$, nous avons

$$\underline{g}(x) = \bar{g}(x) = g_1(x).$$

Par la définition du sous-gradient et par (3.4) et (3.5), nous pouvons observer que $-g_1(x)$ est le sous-gradient de F en θ pour tout $x \in M(\theta)$. En effet,

$$\begin{aligned} -g_1 \in \partial F(\theta) &\Leftrightarrow F(\mu) \leq F(\theta) + \langle -g_1, \mu - \theta\rangle \quad \forall\, \mu \\ &\Leftrightarrow F(\mu) \leq F(\theta) - \langle g_1, \mu - \theta\rangle \quad \forall\, \mu \\ &\Leftrightarrow F(\mu) \leq F(\theta) + \langle g_1, \theta - \mu\rangle \quad \forall\, \mu \\ &\Leftrightarrow (3.4) \quad \text{et} \quad (3.5). \end{aligned}$$

Proposition 3.3.2. *Supposons qu'il existe $m > 0$ tel que $g_i(x) \geq m$, pour tout $x \in S$ et $i = 1, \dots, p$. Alors, pour $\mu > \theta$, nous avons*

$$F(\mu) + (\mu - \theta)m \leq F(\theta).$$

D'où, $F(\theta)$ est décroissante sur l'intervalle où elle est finie.

Démonstration.

Pour tout $x \in S$ et $i = 1, \ldots, p$,

$$\begin{aligned} f_i(x) - \theta g_i(x) &= f_i(x) - \mu g_i(x) + \underbrace{(\mu - \theta)}_{>0} \underbrace{g_i(x)}_{\geq m > 0} \\ &\geq f_i(x) - \mu g_i(x) + (\mu - \theta)m. \end{aligned}$$

Ce qui implique que, pour tout $x \in S$,

$$\begin{aligned} &\max_{1\leq i\leq p} [f_i(x) - \theta g_i(x)] \geq \max_{1\leq i\leq p} [f_i(x) - \mu g_i(x)] + (\mu - \theta)m \\ \Rightarrow\ &\underbrace{\inf_{x\in S} \left(\max_{1\leq i\leq p} [f_i(x) - \theta g_i(x)] \right)}_{F(\theta)} \geq \underbrace{\inf_{x\in S} \left(\max_{1\leq i\leq p} [f_i(x) - \mu g_i(x)] \right)}_{F(\mu)} + (\mu - \theta)m \\ \Leftrightarrow\ &F(\theta) \geq F(\mu) + (\mu - \theta)m. \end{aligned}$$

□

La propriété 3.3.2 et la proposition 3.3.2 nous permettent d'énoncer la propriété suivante :

Propriété 3.3.5. $F(\theta) = 0 \Rightarrow \theta = \bar{\theta}$

Démonstration.

Par la propriété 3.3.2 et la proposition 3.3.2, nous savons que

$$F(\mu) + (\mu - \theta)m \leq F(\theta) \qquad \forall \mu > \theta \tag{3.6}$$

et que

$$F(\theta) < 0 \Leftrightarrow \theta > \bar{\theta}. \tag{3.7}$$

Par la contraposée de (3.7), nous obtenons

$$F(\theta) = 0 \Rightarrow \theta \leq \bar{\theta}.$$

Montrons que $\theta < \bar{\theta}$ conduit à une absurdité. En effet, par (3.6),

$$\text{si} \quad \theta < \bar{\theta}, \qquad \underbrace{F(\bar{\theta})}_{=0} + \underbrace{(\bar{\theta} - \theta)}_{>0} \underbrace{m}_{>0} \leq \underbrace{F(\theta)}_{=0}.$$

Ce qui est impossible. Par conséquent,

$$F(\theta) = 0 \Rightarrow \theta = \bar{\theta}.$$

$\square$

3.4 Exemples

L'objectif de cette section est d'illustrer par des exemples les différentes propriétés de la fonction F citées précédemment.

3.4.1 Exemple 1

Soient

$$n = 1, \quad p = 1, \quad f_1(x) = 1 + x, \quad g_1(x) = x, \quad S = \{x | x \geq 1\}.$$

D'une part, le problème (P) peut s'écrire comme suit :

$$(P) \qquad \bar{\theta} = \inf_{x \geq 1} \left\{ \frac{1+x}{x} \right\}.$$

Résolvons ce dernier :

$$\begin{aligned} \bar{\theta} &= \inf_{x \geq 1} \left\{ \frac{1+x}{x} \right\} \\ &= \inf_{x \geq 1} \left\{ \frac{1}{x} \right\} + 1 \\ &= 1. \end{aligned}$$

D'autre part, le problème (P_θ) peut s'écrire comme suit :

$$(P_\theta) \qquad F(\theta) = \inf_{x \geq 1} \{(1 + x) - \theta x\}.$$

Résolvons ce dernier :

$$\begin{aligned} F(\theta) &= \inf_{x \geq 1} \{(1 + x) - \theta x\} \\ &= \inf_{x \geq 1} \underbrace{\{1 + (1 - \theta)x\}}_{\overset{not}{=} h(x)}. \end{aligned}$$

$$\begin{aligned}
\text{Si } \theta \leq 1,\ \tfrac{dh}{dx} = (1-\theta) \geq 0 \ &\Rightarrow h \text{ est croissante} \\
&\Rightarrow \text{l'infimum est atteint en } x = 1 \\
&\Rightarrow F(\theta) = 2 - \theta.
\end{aligned}$$

$$\begin{aligned}
\text{Si } \theta > 1,\ \tfrac{dh}{dx} = (1-\theta) < 0 \ &\Rightarrow h \text{ est décroissante} \\
&\Rightarrow \text{l'infimum est atteint en } x = +\infty \\
&\Rightarrow F(\theta) = -\infty.
\end{aligned}$$

Par conséquent,

$$F(\theta) = \begin{cases} 2-\theta, & \text{si} \quad \theta \leq 1, \\ -\infty, & \text{si} \quad \theta > 1. \end{cases}$$

Cet exemple illustre le fait que la fonction F n'est pas nécessairement finie sur $\mathbb{R}$. De plus, nous observons que l'existence d'une solution optimale de (P_θ) n'implique pas que $F(\bar{\theta}) = 0$.

3.4.2 Exemple 2

Soient

$$n = 1, \quad p = 1, \quad f_1(x) = e^x, \quad g_1(x) = e^{2x}, \quad S = \mathbb{R}.$$

D'une part, le problème (P) peut s'écrire comme suit :

$$(P) \qquad \bar{\theta} = \inf_{x \in \mathbb{R}} \left\{ \frac{e^x}{e^{2x}} \right\}.$$

Résolvons ce dernier :

$$\begin{aligned}
\bar{\theta} &= \inf_{x \in \mathbb{R}} \left\{ \frac{e^x}{e^{2x}} \right\} \\
&= \inf_{x \in \mathbb{R}} \left\{ e^{-x} \right\} \\
&= 0.
\end{aligned}$$

D'autre part, le problème (P_θ) s'écrit :

$$(P_\theta) \qquad F(\theta) = \inf_{x \in \mathbb{R}} \left\{ e^x - \theta e^{2x} \right\}.$$

Par un raisonnement semblable à l'exercice précédent, la solution du problème (P_θ) est la suivante :

$$F(\theta) = \begin{cases} 0, & \text{si} \quad \theta \leq 0, \\ -\infty, & \text{si} \quad \theta > 0. \end{cases}$$

Cet exemple illustre le fait que la fonction F n'est pas forcément décroissante sur l'intervalle sur lequel F est finie. De plus, $F(\bar{\theta}) = 0$ n'implique pas l'existence d'une solution optimale de (P), et ce même si S est fermé.

3.4.3 Exemple 3

Soient

$$n = 1, \quad p = 2, \quad f_1(x) = 2x, \quad g_1(x) = 2, \quad f_2(x) = -x, \quad g_2(x) = 1, \quad S = [0, 1].$$

D'une part, le problème (P) peut s'écrire comme suit :

$$(P) \qquad \bar{\theta} = \inf_{x \in [0,1]} \left\{ \max_{i=1,2} \left\{ \frac{2x}{2}, \frac{-x}{1} \right\} \right\}.$$

La solution du problème (P) sera alors :

$$\begin{aligned} \bar{\theta} &= \inf_{x \in [0,1]} \left\{ \max_{i=1,2} \left\{ \frac{2x}{2}, \frac{-x}{1} \right\} \right\} \\ &= \inf_{x \in [0,1]} \left\{ \max_{i=1,2} \{x, -x\} \right\} \\ &= 0. \end{aligned}$$

D'autre part, le problème (P_θ) s'écrit :

$$(P_\theta) \qquad F(\theta) = \inf_{x \in [0,1]} \left\{ \max_{i=1,2} \{2x - 2\theta, -x - \theta\} \right\}.$$

Par un raisonnement semblable à l'exercice 1, la solution du problème (P_θ) est la suivante :

$$F(\theta) = \begin{cases} -2\theta, & \text{si} \quad \theta \leq 0, \\ -\left(\frac{4}{3}\right)\theta, & \text{si} \quad 0 \leq \theta \leq 3, \\ -1 - \theta, & \text{si} \quad \theta \geq 3. \end{cases}$$

Notons tout d'abord que, lorsque $p = 1$, la fonction F est concave puisqu'il s'agit d'un infimum de fonctions concaves $f_1(x) - \theta g_1(x)$ en θ et il suit que la fonction F est continue sur l'intérieur de son domaine de finitude. Cet exemple illustre le fait que, lorsque $p > 1$, la concavité de la fonction F disparaît.

Remarque 3.4.1. Remarquons que, dans l'exemple 3, la fonction F est continue en $\bar{\theta}$ et que, dans les exemples 1 et 2, la fonction F est discontinue en $\bar{\theta}$ avec un saut infini. Notons, néanmoins, qu'il existe également des fonctions F avec des sauts finis en $\bar{\theta}$.

3.5 Description et analyse de l'algorithme de Dinkelbach

Introduisons, à présent, un algorithme afin de résoudre le problème (P) via (P_θ). Cet algorithme porte le nom d'*Algorithme de Dinkelbach.*

Algorithme de Dinckelbach

1. Soit $x^0 \in S$. Soit
$$\theta_1 = \max_i \frac{f_i(x^0)}{g_i(x^0)}.$$
Soit $k = 1$.
2. Résoudre (P_{θ_k}). Prendre $x^k \in M(\theta_k)$.
3. Si $F(\theta_k) = 0$, alors STOP. Les points x^k et x^{k-1} sont des solutions optimales du problème (P), et $\theta_k = \bar{\theta}$.
4. Si $F(\theta_k) \neq 0$, calculer
$$\theta_{k+1} = \max_i \frac{f_i(x^k)}{g_i(x^k)}.$$
Prendre $k = k + 1$ et retourner au point 2.

3.5.1 Analyse de l'algorithme

L'application de cet algorithme nécessite la détermination, en plus de la condition initiale, des solutions x^k de (P_{θ_k}), $k = 1, 2, \ldots$.

Par la construction de θ_k, il est évident que

$$\theta_k \geq \bar{\theta}.$$

Nous pouvons également observer que

$$\max_{1\leq i\leq p}[f_i(x^{k-1}) - \theta_k g_i(x^{k-1})] = 0 \qquad k \geq 1. \tag{3.8}$$

En effet,

$$\begin{aligned} & \theta_k = \max_{1\leq i\leq p} \frac{f_i(x^{k-1})}{g_i(x^{k-1})} \\ \Leftrightarrow \quad & \exists\, i_k \in \{1,\ldots,p\} \text{ t.q. } \theta_k = \tfrac{f_{i_k}(x^{k-1})}{g_{i_k}(x^{k-1})} \\ \Leftrightarrow \quad & \exists\, i_k \in \{1,\ldots,p\} \text{ t.q. } f_{i_k}(x^{k-1}) - \theta_k g_{i_k}(x^{k-1}) = 0. \end{aligned} \tag{3.9}$$

De plus,

$$\begin{aligned} \theta_k &= \max_{1\leq i\leq p} \frac{f_i(x^{k-1})}{g_i(x^{k-1})} \\ &\geq \frac{f_i(x^{k-1})}{g_i(x^{k-1})} \qquad i = 1,\ldots,p \\ \Leftrightarrow\ & f_i(x^{k-1}) - \theta_k g_i(x^{k-1}) \leq 0 \qquad (\text{car } g_i \text{ positives}). \end{aligned} \tag{3.10}$$

Alors, par (3.9) et (3.10), nous obtenons bien (3.8). Par conséquent, $F(\theta_k) \leq 0$.
Notons également qu'à l'étape 3, par la propriété 3.3.5,

$$F(\theta_k) = 0 \Rightarrow \theta_k = \bar{\theta}.$$

Remarque 3.5.1. Remarquons que les problèmes (P_{θ_k}) sont des problèmes convexes si les fonctions f_i sont non négatives et convexes, si les fonctions g_i sont concaves et si S est un ensemble convexe.
Au lieu de résoudre le problème (P_{θ_k}), nous pouvons résoudre le problème convexe équivalent suivant :

$$\inf\left\{\lambda | f_i(x) - \theta_k g_i(x) - \lambda \leq 0, i = 1,\ldots,p,\ x \in S\right\}. \tag{3.11}$$

Ces sous-problèmes sont des problèmes linéaires si les fonctions f_i et g_i sont affines et si S est un polyhèdre convexe.

Proposition 3.5.1. *Nous avons*

1. $-\frac{F(\theta_k)}{\bar{g}(x^k)} \leq -\frac{F(\theta_k)}{g_j(x^k)} \leq \theta_k - \theta_{k+1} \leq -\frac{F(\theta_k)}{g_i(x^k)} \leq -\frac{F(\theta_k)}{\underline{g}(x^k)} \quad \forall j \in J(x^k)$ *et* $i \in I(x^k, \theta_k)$;
2. $\theta_k \geq \bar{\theta} \quad \forall k$; *et, si* $\theta_k > \bar{\theta}$, *alors* $\theta_k > \theta_{k+1} \geq \bar{\theta}$.

Démonstration.

Soient $j \in J(x^k)$ et $x^k \in M(\theta_k)$.

1. (a) Montrons d'abord que $-\frac{F(\theta_k)}{\bar{g}(x^k)} \leq -\frac{F(\theta_k)}{g_j(x^k)}$ ·

   Par la définition de $\bar{g}(x^k)$,

$$\begin{aligned}\bar{g}(x^k) &\stackrel{\text{def}}{=} \max_{1\leq i\leq p} g_i(x^k) \\ &\geq g_i(x^k) \qquad i = 1, \ldots, p.\end{aligned}$$

   En particulier, pour $i = j \in J(x^k)$,

$$\bar{g}(x^k) \geq g_j(x^k).$$

   Par conséquent,

$$-\frac{1}{g_j(x^k)} \leq -\frac{1}{\bar{g}(x^k)}.$$

   Comme dit précédemment $\theta_k \geq \bar{\theta}$, pour tout $k = 1, 2, \ldots$. Si $\theta_k > \bar{\theta}$, alors, par la propriété 3.3.2, $F(\theta_k) < 0$. D'où

$$-\frac{F(\theta_k)}{\bar{g}(x^k)} \leq -\frac{F(\theta_k)}{g_j(x^k)}.$$

   (b) Montrons ensuite que $-\frac{F(\theta_k)}{g_j(x^k)} \leq \theta_k - \theta_{k+1}$.

   Puisque $x^k \in M(\theta_k)$,

$$\begin{aligned}F(\theta_k) &= \max_{1\leq i\leq p}[f_i(x^k) - \theta_k g_i(x^k)] \\ &\geq f_i(x^k) - \theta_k g_i(x^k) \qquad i = 1, \ldots, p.\end{aligned}$$

   En particulier, pour $i = j \in J(x^k)$,

$$F(\theta_k) \geq f_j(x^k) - \theta_k g_j(x^k). \tag{3.12}$$

Notons que, puisque $j \in J(x^k)$,

$$f_j(x^k) - \theta_{k+1}\, g_j(x^k) = f_j(x^k) - \underbrace{\max_{1\le i\le p} \frac{f_i(x^k)}{g_i(x^k)}}_{=\frac{f_j(x^k)}{g_j(x^k)}} g_j(x^k) = 0. \tag{3.13}$$

Alors, par (3.12) et (3.13),

$$\begin{aligned} F(\theta_k) &\ge \underbrace{f_j(x^k) - \theta_{k+1}g_j(x^k)}_{=0} + \theta_{k+1}g_j(x^k) - \theta_k g_j(x^k) \\ &= g_j(x^k)(\theta_{k+1} - \theta_k). \end{aligned}$$

Finalement,

$$-\frac{F(\theta_k)}{g_j(x^k)} \le \theta_k - \theta_{k+1}.$$

(c) Montrons également que $\theta_k - \theta_{k+1} \le -\frac{F(\theta_k)}{g_i(x^k)}$.
Soit $i \in I(x^k, \theta_k)$. Alors

$$\begin{aligned} F(\theta_k) &= f_i(x^k) - \theta_k g_i(x^k) \\ &= g_i(x^k)\left(\frac{f_i(x^k)}{g_i(x^k)} - \theta_k\right) \\ &\le g_i(x^k)(\theta_{k+1} - \theta_k) \quad \text{car} \quad \theta_{k+1} = \max_{1\le i\le p} \frac{f_i(x^k)}{g_i(x^k)}. \end{aligned}$$

D'où,

$$\theta_k - \theta_{k+1} \le -\frac{F(\theta_k)}{g_i(x^k)}.$$

(d) Montrons finalement que $-\frac{F(\theta_k)}{g_i(x^k)} \le -\frac{F(\theta_k)}{\underline{g}(x^k)}$.
Par la définition de $\underline{g}(x^k)$,

$$\begin{aligned} \underline{g}(x^k) &= \min_{1\le i\le p} g_i(x^k) \\ &\le g_i(x^k) \qquad i = 1, \ldots, p. \end{aligned}$$

Par conséquent,

$$-\frac{1}{\underline{g}(x^k)} \le -\frac{1}{g_i(x^k)}.$$

Comme dit précédemment $\theta_k \ge \bar{\theta}$, pour tout $k = 1, 2, \ldots$ Si $\theta_k > \bar{\theta}$, alors, par la propriété 3.3.2, $F(\theta_k) < 0$. D'où

$$-\frac{F(\theta_k)}{g_i(x^k)} \le -\frac{F(\theta_k)}{\underline{g}(x^k)}.$$

2. Par la construction de θ_k,
$$\theta_k \geq \bar{\theta} \qquad \forall\, k.$$
De plus, par la propriété 3.3.2, nous savons que si $\theta_k > \bar{\theta}$, alors $F(\theta_k) < 0$. Par conséquent, par le point 1 de cette proposition,
$$0 < -\frac{F(\theta_k)}{g_j(x^k)} \leq \theta_k - \theta_{k+1}.$$
Finalement,
$$\theta_k > \theta_{k+1}.$$

□

Remarque 3.5.2. Notons que, lorsque $p = 1$,
$$\theta_{k+1} = \theta_k - \frac{F(\theta_k)}{-g_1(x^k)}.$$
Comme dit précédemment, dans le cas où $p = 1$, $-g_1(x^k)$ est le sous-gradient de la fonction F en θ_k. Par conséquent, la méthode proposée ci-dessus coïncide avec la méthode de Newton lorsque $p = 1$. Par contre, lorsque $p > 1$, cette similitude disparaît. Illustrons cela par un exemple.

Exemple 3.5.1.
Soient
$$n = 2,\ p = 2,\ f_1(x) = x_1 - 1,\ g_1(x) = 1,\ f_2(x) = 2x_1 + 1,\ g_2(x) = x_1 + x_2 + 1$$
$$S = \left\{x \in \mathbb{R}^2 \mid x_1 \geq 0,\ x_2 \geq 0\right\}.$$
et soit le point initial
$$x^0 = (1, 1).$$
Notons, tout d'abord,
$$\begin{aligned} F(\theta) &= \inf_{x\in S}\left\{\max_{i=1,2}[(x_1 - 1) - \theta; (2x_1 + 1) - \theta(x_1 + x_2 + 1)]\right\} \\ &= \begin{cases} 1 - \theta, & \text{si} \quad \theta \leq 0, \\ -1 - \theta, & \text{si} \quad \theta > 0. \end{cases} \end{aligned}$$

Appliquons alors l'algorithme défini précédemment :

$\boxed{\text{Pas 1}}$ Soit $x^0 = (1,1)$. Alors

$$\begin{aligned}\theta_1 &= \max_{i=1,2}\left\{x_1^0 - 1, \frac{2x_1^0+1}{x_1^0+x_2^0+1}\right\}\\ &= \max_{i=1,2}\{0,1\}\\ &= 1.\end{aligned}$$

Soit $k=1$.

$\boxed{\text{Pas 2}}$ Résoudre

$$(P_{\theta_1}) \qquad \begin{aligned}F(\theta_1) &= \inf_{x\in S}\left\{\max_{i=1,2}[(x_1-1)-1;(2x_1+1)-(x_1+x_2+1)]\right\}\\ &= \inf_{x\in S}\left\{\max_{i=1,2}[x_1-2;x_1-x_2]\right\}\\ &= -1-\theta_1 \quad \text{car} \quad \theta_1 = 1 > 0\\ &= -2.\end{aligned}$$

Prendre $x^1 \in M(\theta_1)$. Choisissons

$$x^1 = (0,2).$$

$\boxed{\text{Pas 4}}$ $F(\theta_1) \neq 0$, prendre

$$\begin{aligned}\theta_2 &= \max_{i=1,2}\left\{x_1^1 - 1, \frac{2x_1^1+1}{x_1^1+x_2^1+1}\right\}\\ &= \max_{i=1,2}\left\{-1, \frac{1}{3}\right\}\\ &= \frac{1}{3}.\end{aligned}$$

Soit $k=2$.

$\boxed{\text{Pas 2}}$ Résoudre

$$(P_{\theta_2}) \qquad \begin{aligned}F(\theta_2) &= \inf_{x\in S}\left\{\max_{i=1,2}[(x_1)-\frac{1}{3};(2x_1+1)-\frac{1}{3}(x_1+x_2+1)]\right\}\\ &= -1-\theta_2 \quad \text{car} \quad \theta_2 = \frac{1}{3} > 0\\ &= -\frac{4}{3}.\end{aligned}$$

Prendre $x^2 \in M(\theta_2)$. Choisissons

$$x^2 = (0, 6).$$

$\boxed{\text{Pas 4}}$ $F(\theta_2) \neq 0$, prendre

$$\begin{aligned}
\theta_3 &= \max_{i=1,2} \left\{ x_1^2 - 1, \frac{2x_1^2 + 1}{x_1^2 + x_2^2 + 1} \right\} \\
&= \max_{i=1,2} \left\{ -1, \frac{1}{7} \right\} \\
&= \frac{1}{7}.
\end{aligned}$$

En faisant de même pour $k = 1, 2, \ldots$, nous obtenons finalement

$$\theta_{k+1} = \frac{1}{1 + \frac{2}{\theta_k}} \quad \text{et} \quad x^k = \left(0, \frac{2}{\theta_k}\right).$$

Par conséquent,

$$\theta_k = \left(\sum_{l=0}^{k-1} 2^l \right)^{-1} \qquad k = 1, 2, \ldots.$$

En effet, pour $k = 1, 2, \ldots$,

$$\begin{aligned}
\theta_k &= \frac{1}{1 + \frac{2}{\theta_{k-1}}} \\
&= \frac{1}{1 + 2\left(1 + \frac{2}{\theta_{k-2}}\right)} \\
&= \frac{1}{1 + 2 + \frac{2^2}{\theta_{k-2}}} \\
&= \frac{1}{1 + 2 + 2^2 + \frac{2^3}{\theta_{k-3}}} \\
&= \frac{1}{1 + 2 + \cdots + \frac{2^{k-1}}{\theta_1}} \\
&= \left(\sum_{l=0}^{k-1} 2^l \right)^{-1}.
\end{aligned}$$

Par conséquent, la suite θ_k converge vers $\bar{\theta} = 0$.

Appliquons, à présent, la méthode de Newton à cet exemple :

Méthode de Newton
Soit x un point de départ
Répéter
1. Calculer la direction d : $\nabla^2 f(x)d = -\nabla f(x)$;
2. $x := x + d$
jusqu'à ce que $\nabla f(x) \approx 0$.

Soient

$$x^0 = (1, 1) \quad \text{et} \quad \theta_1 = 1.$$

Déterminons la suite θ_k :

$$\begin{aligned}
\theta_1 &= 1 \\
\theta_2 &= \theta_1 + \frac{F(\theta_1)}{g_1(x^1)} = 1 + \frac{-1-\theta_1}{1} = -1 \\
\theta_3 &= \theta_2 + \frac{F(\theta_2)}{g_1(x^2)} = -1 + 1 + 1 = 1 \\
&\vdots \\
\theta_k &= (-1)^{k+1}.
\end{aligned}$$

Cette suite θ_k ne converge pas.

Remarquons également que le problème (P) ne possède pas de solution optimale.

3.5.2 Convergence de la méthode

Notons, tout d'abord, qu'il est nécessaire de supposer que $M(\theta) \neq \emptyset$, pour $\theta \in (\bar{\theta}, \theta_1]$ puisqu'il faut que $F(\theta) > -\infty$ pour un tel θ et qu'il existe une solution que nous noterons $\bar{x}(\theta)$. Supposons également que le problème (P) possède une solution optimale $\bar{x}$ (i.e. , $\bar{\theta} > -\infty$) et que cette valeur est atteinte.

Remarque 3.5.3. Notons que, dans le cas linéaire de la progammation fractionnelle généralisée, l'ensemble des solutions optimales de (P_θ) sera non vide à condition que

$F(\theta_1) > -\infty$ pour au moins un $\theta_1 > \bar{\theta}$, puisque le problème linéaire équivalent (3.11) atteint toujours la valeur optimale $F(\theta)$. Remarquons, néanmoins, que, même dans le cas linéaire, $F(\theta) = -\infty$, pour tout $\theta > \bar{\theta}$ est possible.

Proposition 3.5.2. *Si (P) possède une solution optimale $\bar{x}$ et si $M(\theta_k) \neq \emptyset$, alors*

$$(\theta_{k+1} - \bar{\theta}) \leq \left(1 - \frac{\underline{g}(\bar{x})}{\bar{g}(x^k)}\right)(\theta_k - \bar{\theta}), \qquad k = 1, 2, \ldots.$$

Démonstration.

Par la proposition 3.5.1,

$$\begin{aligned}
& & -\tfrac{F(\theta_k)}{\bar{g}(x^k)} \leq \theta_k - \theta_{k+1} \\
\Leftrightarrow & & \theta_{k+1} \leq \theta_k + \tfrac{F(\theta_k)}{\bar{g}(x^k)} \\
\Leftrightarrow & & \theta_{k+1} - \bar{\theta} \leq \theta_k - \bar{\theta} + \tfrac{F(\theta_k)}{\bar{g}(x^k)}.
\end{aligned} \tag{3.14}$$

Or, (P) possède une solution optimale $\bar{x}$. Par conséquent,

$$\begin{aligned}
&\overset{\text{propri } 3.3.3}{\Rightarrow} & & F(\bar{\theta}) = 0 \\
&\overset{\text{propri } 3.3.4}{\Rightarrow} & & (P) \quad \text{et} \quad (P_{\bar{\theta}}) \text{ ont le même ensemble de solutions optimales} \\
&\Rightarrow & & M(\bar{\theta}) \neq \emptyset.
\end{aligned}$$

Dès lors, nous pouvons appliquer la proposition 3.3.1 avec $\theta_k > \bar{\theta}$:

$$\begin{aligned}
F(\theta_k) &\leq \underbrace{F(\bar{\theta})}_{=0} + (\bar{\theta} - \theta_k)\underline{g}(\bar{x}) \\
&= (\bar{\theta} - \theta_k)\underline{g}(\bar{x}).
\end{aligned}$$

La relation (3.14) devient

$$\begin{aligned}
& & \theta_{k+1} - \bar{\theta} \leq \theta_k - \bar{\theta} + \frac{(\bar{\theta} - \theta_k)\underline{g}(\bar{x})}{\bar{g}(x^k)} \\
\Leftrightarrow & & \theta_{k+1} - \bar{\theta} \leq \left(1 - \frac{\underline{g}(\bar{x})}{\bar{g}(x^k)}\right)(\theta_k - \bar{\theta}).
\end{aligned}$$

$\square$

Corollaire 3.5.1. *Si (P) possède une solution $\bar{x}$, si $M(\theta_k) \neq \emptyset$, $k = 1, 2, \ldots$, et si $\sup_k \bar{g}(x^k) < \infty$, alors $\{\theta_k\}$ converge vers $\bar{\theta}$ et ce, linéairement.*

Démonstration.

Rappelons, tout d'abord, qu'une suite u_k converge linéairement vers u^* si

$$\exists\, 0 < \gamma < 1,\ \exists\, k_0,\ \forall\, k \geq k_0 \quad \|u_{k+1} - u^*\| \leq \gamma \|u_k - u^*\|\,.$$

Par la proposition 3.5.2 et puisque $\sup\limits_k \bar{g}(x^k) < \infty$, nous savons que

$$\begin{aligned}\|\theta_{k+1} - \bar{\theta}\| &\leq \left\|(\theta_k - \bar{\theta})\left(1 - \frac{\underline{g}(\bar{x})}{\bar{g}(x^k)}\right)\right\| \\ &= \|\theta_k - \bar{\theta}\| \left\|\left(1 - \frac{\underline{g}(\bar{x})}{\bar{g}(x^k)}\right)\right\|.\end{aligned}$$

Prendre $\gamma = \left\|1 - \frac{\underline{g}(\bar{x})}{\bar{g}(x^k)}\right\|$. Vérifions que ce γ est compris entre les bornes 0 et 1 :

- $\gamma > 0$: toujours vérifié par propriété de la norme
- $\gamma < 1$: pour ce faire, montrons, en développant cette inégalité, que nous arrivons à une expression qui est toujours satisfaite :

$$\begin{aligned}&\gamma < 1 \\ \Leftrightarrow\ & 1 - \frac{\underline{g}(\bar{x})}{\bar{g}(x^k)} < 1 \\ \Leftrightarrow\ & \bar{g}(x^k) - \underline{g}(\bar{x}) < \bar{g}(x^k) \\ \Leftrightarrow\ & \underline{g}(\bar{x}) > 0 : \ \text{toujours vérifié par hypothèse sur } g_i \text{ pour tout } i.\end{aligned}$$

□

Remarque 3.5.4. Notons que l'algorithme peut encore converger linéairement même si certaines hypothèses du corollaire 3.5.1 ne sont pas satisfaites. Illustrons cela en reprenant l'exemple 3.5.1. Soit $x^0 = (1, 1)$ le point initial et soient les expressions suivantes calculées précédemment

$$\bar{\theta} = 0, \ \theta_{k+1} = \frac{\theta_k}{2 + \theta_k}.$$

Par conséquent, la suite $\{\theta_k\}$ satisfait

$$\theta_{k+1} - \bar{\theta} = \left(\frac{1}{2+\theta_k}\right)(\theta_k - \bar{\theta}).$$

Puisque

$$\frac{1}{3} \leq \frac{1}{2+\theta_k} \leq \frac{1}{2}, \qquad k = 1, 2, \ldots,$$

la suite $\{\theta_k\}$ converge linéairement vers $\bar{\theta}$ et ce, malgré le fait que toutes les hypothèses du corollaire 3.5.1 ne sont pas vérifiées. En effet, le problème (P) ne possède pas de solution optimale et

$$\sup_k \bar{g}(x^k) = \sup_k \left(1 + \frac{2}{\theta_k}\right) = \infty.$$

Chapitre 4

Programmation fractionnelle généralisée dans le cas compact

Considérons, dans ce chapitre, le problème suivant

$$(P) \qquad \bar{\theta} = \inf_{x \in S} \left(\max_{1 \leq i \leq p} \frac{f_i(x)}{g_i(x)} \right)$$

où $S \subseteq \mathbb{R}^n$ est un ensemble non vide **compact**, les fonctions f_i sont continues sur S et les fonctions g_i sont continues et positives sur S.

4.1 Propriétés

Théorème 4.1.1. *Soit S un ensemble non vide compact de $\mathbb{R}^n$. Alors, les affirmations suivantes sont vérifiées.*

1. *Les problèmes (P) et (P_θ) possèdent toujours une solution optimale, $\bar{\theta}$ est fini, et $F(\bar{\theta}) = 0$. D'où, $F(\theta) = 0$ implique que $\theta = \bar{\theta}$.*
2. *F est finie, continue et décroissante sur $\mathbb{R}$.*
3. *La suite $\{\theta_k\}$, si elle n'est pas finie, converge linéairement vers $\bar{\theta}$, et toute sous-suite convergente de $\{x^k\}$ converge vers une solution optimale de (P).*

Démonstration.

1. Montrons d'abord que (P) et (P_θ) ont toujours une solution $\bar{\theta}$ finie.

   L'existence d'une solution $\bar{\theta}$ résulte de la compacité de S et de la continuité des fonctions f_i et g_i sur cet ensemble. Par conséquent,

   $$\exists\, \bar{\theta} \text{ t.q. } \bar{\theta} = \inf_{x \in S} \left(\max_{1 \leq i \leq p} \frac{f_i(x)}{g_i(x)} \right).$$

   Cette solution $\bar{\theta}$ sera finie car l'ensemble S est borné.

   Il nous reste à montrer que $F(\bar{\theta}) = 0$ et que, par conséquent, $F(\theta) = 0 \Rightarrow \theta = \bar{\theta}$.

   Par le point précédent, le problème (P) possède une solution optimale. Dès lors, par la propriété 3.3.3,

   $$F(\bar{\theta}) = 0.$$

   L'implication a déjà été prouvée par la propriété 3.3.5 du chapitre 3.

2. Montrons ensuite que F est finie, continue et décroissante sur $\mathbb{R}$.

   La fonction F est continue puisqu'il s'agit d'un infimum sur un ensemble compact de fonctions continues.

   Par la proposition 3.3.2, nous pouvons affirmer que la fonction F est décroissante. En effet, pour $\mu > \theta$,

   $$F(\theta) \geq F(\mu) + \underbrace{(\mu - \theta)\, m}_{>0} > F(\mu).$$

3. Montrons également que $\{\theta_k\}$ tend vers $\bar{\theta}$ linéairement si la suite $\{\theta_k\}$ n'est pas finie.

   Pour ce faire, vérifions que toutes les hypothèses du corollaire 3.5.1 sont vérifiées.

   - (P) admet une solution optimale par le point 1.
   - $M(\theta_k) \neq \emptyset$ car S est compact.
   - La dernière hypothèse à vérifier est $\sup_k \bar{g}(x^k) < +\infty$:

     La fonction $\bar{g}$ étant le supremum de fonctions continues est continue. Soit $\{x^l\}$ une suite de S. Alors, l'ensemble défini par $\{\bar{g}(x^l) \mid l = 1, 2, \ldots\}$ est compact. Par conséquent, les fonctions $\bar{g}(x^l)$ sont bornées.

   Dès lors, en appliquant ce corollaire, nous pouvons conclure que la suite $\{\theta_k\}$ converge linéairement vers $\bar{\theta}$.

Montrons finalement que toute sous-suite convergente de $\{x^k\}$ converge vers une solution optimale de (P).
Soit $\{x^{k_l}\}$ une sous-suite convergente de la suite $\{x^k\}$ telle que $\{x^{k_l}\}$ converge vers un certain $\hat{x} \in S$. Puisque x^{k_l} appartient à S compact, $\hat{x} \in S$. Ainsi,

$$F(\theta_{k_l}) = \max_i \left(f_i(x^{k_l}) - \theta_{k_l} g_i(x^{k_l})\right).$$

Par continuité,

$$\underbrace{F(\bar{\theta})}_{=0} = \max_i \left(f_i(\hat{x}) - \bar{\theta} g_i(\hat{x})\right).$$

Par conséquent, $\hat{x}$ est une solution optimale de (P).

□

Remarque 4.1.1. Par la proposition 3.5.2, avec $p = 1$ et $\bar{x}$ une solution optimale de (P),

$$(\theta_{k+1} - \bar{\theta}) \leq \underbrace{\left(1 - \frac{g_1(\bar{x})}{g_1(x^k)}\right)}_{(*)}(\theta_k - \bar{\theta}), \qquad k = 1, 2, \ldots. \tag{4.1}$$

Montrons que (*) converge vers zéro afin de pouvoir affirmer la convergence superlinéaire de la méthode :

$$\begin{aligned} & 1 - \frac{g_1(\bar{x})}{g_1(x^k)} \to 0 \\ \Leftrightarrow\ & \frac{g_1(\bar{x})}{g_1(x^k)} \to 1 \\ \Leftrightarrow\ & g_1(x^k) \to g_1(\bar{x}). \end{aligned}$$

La fonction g_1 étant continue, il suffit de montrer que la suite $\{x^k\}$ converge vers $\bar{x}$; ce que nous venons de démontrer au théorème 4.1.1.
Malheureusement, cette convergence superlinéaire disparaît lorsque $p > 1$. De plus, par (4.1), nous remarquons que la méthode devient de plus en plus lente lorsque p augmente. Plus le nombre de rapports concernés est élevé, plus la méthode sera lente.

Remarque 4.1.2. Dans le chapitre précédent, nous constations que, lorsque $p = 1$, notre méthode coïncide avec l'algorithme de Newton et que, lorsque $p > 1$, la méthode de Newton est assez différente de notre méthode. Notons, ici, que notre algorithme peut être plus lent que la méthode de Newton en général. Néanmoins, ce dernier présente deux avantages :

1. un sous-gradient de F, qui n'est pas facilement disponible, ne doit pas être calculé,
2. notre méthode converge même pour des fonctions non concaves.

4.2 Algorithme de Dinkelbach modifié

Afin d'introduire un autre algorithme, considérons la fonction paramétrique suivante :

$$F(x,w,\theta) = \max_{1\leq i\leq p} \left\{ \frac{f_i(x) - \theta g_i(x)}{w_i} \right\}$$

où $w \in \mathbb{R}^p$, $w_i > 0$ pour tout i et θ est un paramètre réel.

Dans ce chapitre, nous choisirons $w_i^k = g_i(x^k)$.

L'algorithme de Dinkelbach, décrit dans le chapitre précédent, peut alors être modifié comme suit :

Algorithme de Dinkelbach modifié

1. Soit $x^0 \in S$. Soit
$$\theta_1 = \max_i \left(\frac{f_i(x^0)}{g_i(x^0)} \right).$$
Soit $k = 1$.
2. Choisir $w_i^{k-1} = g_i(x^{k-1})$ et résoudre le problème suivant :
$$(P_{w^k,\theta_k}) \qquad F_{w^k}(\theta_k) = \inf_{x\in S} \left\{ \max_{1\leq i\leq p} \frac{f_i(x) - \theta_k g_i(x)}{w_i^{k-1}} \right\}.$$
Prendre x^k une solution optimale de (P_{w^k,θ_k}).
3. Si $F_{w^k}(\theta_k) = 0$. Alors STOP.
4. Si $F_{w^k}(\theta_k) \neq 0$, calculer
$$\theta_{k+1} = \max_i \left(\frac{f_i(x^k)}{g_i(x^k)} \right).$$
Prendre $k := k+1$ et retourner au pas 2.

L'unique différence avec l'algorithme décrit auparavant réside dans le problème à résoudre au pas 2. En effet, ici, il faut résoudre le problème (P_{w^k,θ_k}) à la place du problème (P_{θ_k}) qui était le suivant :

$$(P_{\theta_k}) \qquad F(\theta_k) = \inf_{x \in S} \left\{ \max_{1 \leq i \leq p} [f_i(x) - \theta_k g_i(x)] \right\}.$$

Etudions, à présent, la convergence de l'algorithme de Dinkelbach modifié dans le cas où l'ensemble S est compact. Dans ce cas, les problèmes (P_{w^k,θ_k}) possèdent une solution optimale et l'algorithme peut être appliqué.

Théorème 4.2.1. *Supposons que l'ensemble S est compact.*

1. *Si $F_{w^k}(\theta_k) = 0$, alors $\theta_k = \bar{\theta}$ et x^k est une solution optimale du problème (P).*
2. *La suite $\{\theta_k\}$, si elle n'est pas finie, converge au moins linéairement vers $\bar{\theta}$, et toute sous-suite convergente de $\{x^k\}$ converge vers une solution optimale du problème (P).*

Démonstration.

Notons, tout d'abord, que le problème (P) peut être reformulé comme suit :

$$(P) \qquad \bar{\theta} = \inf_{x \in S} \left\{ \max_{1 \leq i \leq p} \left[\frac{f_i(x)/g_i(x^{k-1})}{g_i(x)/g_i(x^{k-1})} \right] \right\},$$

pour tout $k = 1, 2, \ldots$.

Définissons, à présent, la fonction F_{w^k} (où nous choisissons $w^k = g_i(x^k)$) de la manière suivante :

$$F_{w^k}(\theta) = \inf_{x \in S} \left\{ \max_{1 \leq i \leq p} \left[\left(\frac{1}{g_i(x^{k-1})} \right) (f_i(x) - \theta g_i(x)) \right] \right\}.$$

1. Par construction, nous savons que

$$\theta_k \geq \bar{\theta}.$$

En appliquant la propriété 3.3.5 à notre fonction F_{w^k}, nous obtenons :

$$F_{w^k}(\theta_k) = 0 \Rightarrow \theta_k = \bar{\theta}.$$

Par la propriété 3.3.4, le problème (P) possède une solution optimale x^k. Notons, pour la suite, que, par la propriété 3.3.4, x^k est également une solution optimale de (P_{w^k,θ_k}).

2. Soit $\bar{x}$ une solution optimale du problème (P). Par la propriété 3.3.4, $\bar{x}$ est également solution du problème (P_{w^k,θ_k}) et puisque $\theta_k \geq \bar{\theta}$, par la proposition 3.3.1, nous savons que

$$F_{w^k}(\theta_k) \leq F_{w^k}(\bar{\theta}) + (\bar{\theta} - \theta_k) \min_{1\leq i\leq p} \left[\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right].$$

Notons que $F_{w^k}(\bar{\theta}) = 0$ par la propriété 3.3.3. Par conséquent,

$$F_{w^k}(\theta_k) \leq (\bar{\theta} - \theta_k) \min_{1\leq i\leq p} \left[\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right]. \tag{4.2}$$

x^k étant une solution optimale de (P_{w^k,θ_k}),

$$\begin{aligned} F_{w^k}(\theta_k) &= \max_{1\leq i\leq p} \left[\frac{g_i(x^k)}{g_i(x^{k-1})}\left(\frac{f_i(x^k)}{g_i(x^k)} - \theta_k\right)\right] \\ &\geq \tfrac{g_j(x^k)}{g_j(x^{k-1})}(\theta_{k+1} - \theta_k) \end{aligned} \tag{4.3}$$

pour tout $j \in J(x^k) = \left\{j \mid \frac{f_j(x^k)}{g_j(x^k)} = \theta_{k+1}\right\}$. En combinant (4.2) et (4.3), nous obtenons

$$\theta_{k+1} - \bar{\theta} \leq (\theta_k - \bar{\theta})\left(1 - \left(\frac{g_j(x^{k-1})}{g_j(x^k)}\right) \min_{1\leq i\leq p}\left[\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right]\right). \tag{4.4}$$

En effet,

$$\begin{aligned} &(\bar{\theta} - \theta_k) \min_{1\leq i\leq p}\left[\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right] \geq \frac{g_j(x^k)}{g_j(x^{k-1})}(\theta_{k+1} - \theta_k) \\ &\Leftrightarrow \theta_{k+1} - \theta_k \leq (\bar{\theta} - \theta_k)\frac{g_j(x^{k-1})}{g_j(x^k)} \min_{1\leq i\leq p}\left(\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right) \\ &\Leftrightarrow \theta_{k+1} - \theta_k + \theta_k - \bar{\theta} \leq (\bar{\theta} - \theta_k)\frac{g_j(x^{k-1})}{g_j(x^k)} \min_{1\leq i\leq p}\left(\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right) + \theta_k - \bar{\theta} \\ &\Leftrightarrow \theta_{k+1} - \bar{\theta} \leq (\theta_k - \bar{\theta})\left(1 - \frac{g_j(x^{k-1})}{g_j(x^k)} \min_{1\leq i\leq p}\left(\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right)\right). \end{aligned}$$

Définissons

$$\alpha = \inf_{x,y\in S}\left(\min_{1\leq l\leq p}\left[\frac{g_l(x)}{g_l(y)}\right] \min_{1\leq i\leq p}\left[\frac{g_i(\bar{x})}{g_i(x)}\right]\right).$$

Alors, $\alpha > 0$ puisque S est un ensemble compact et que les fonctions g_i sont positives et continues sur ce compact. D'où, (4.4) peut se reformuler comme suit :

$$0 \leq \theta_{k+1} - \bar{\theta} \leq (\theta_k - \bar{\theta})(1 - \alpha).$$

Par conséquent, la suite $\{\theta^k\}$ converge linéairement vers $\bar{\theta}$.

Soit $\tilde{x}$ la limite de toute sous-suite convergente de la suite $\{x^k\}$. Alors, $\tilde{x} \in S$ et il existe $\hat{x} \in S$ et une suite croissante d'entiers positifs $\{n_k\}$ telle que la suite (x_{n_k-1}, x_{n_k}) converge vers $(\hat{x}, \tilde{x})$.

Pour $y \in S$ et $\mu \in \mathbb{R}$, définissons

$$r(y, \mu) = \inf_{x \in S} \left(\max_{1 \le i \le p} \left(\left[\frac{1}{g_i(y)} \right] [f_i(x) - \mu g_i(x)] \right) \right).$$

Notons que r est continue sur $S \times \mathbb{R}$. Nous obtenons alors

$$r(x_{n_k-1}, \theta_{n_k}) = \max_{1 \le i \le p} \left(\left[\frac{1}{g_i(x_{n_k-1})} \right] [f_i(x_{n_k}) - \theta_{n_k} g_i(x_{n_k})] \right).$$

Lorsque k tend vers $+\infty$, nous obtenons

$$r(\hat{x}, \bar{\theta}) = \max_{1 \le i \le p} \left(\left[\frac{1}{g_i(\hat{x})} \right] [f_i(\tilde{x}) - \bar{\theta} g_i(\tilde{x})] \right).$$

Par la propriété 3.3.3, $F(\bar{\theta}) = 0$ car (P) admet une solution.

Dès lors, $\max\limits_{1 \le i \le p} f_i(\tilde{x}) - \bar{\theta} g_i(\tilde{x}) = 0$.

D'où $r(\hat{x}, \bar{\theta}) = 0$ et $\tilde{x}$ est une solution optimale du problème (P).

□

Si le problème (P) possède une solution optimale unique, alors la suite $\{x^k\}$ générée par l'algorithme est convergente. Dans ce cas, la vitesse de convergence de la suite $\{\theta_k\}$ vers $\bar{\theta}$ est au moins superlinéaire comme le montre le théorème suivant :

Théorème 4.2.2. *Supposons que l'ensemble S est compact et que la suite $\{x^k\}$ converge vers $\bar{x}$. Alors, $\{\theta_k\}$ converge superlinéairement vers $\bar{\theta}$. Si, de plus, les fonctions g_i satisfont la condition de Lipschitz sur S, alors il existe une constante M telle que*

$$0 \le \theta_{k+1} - \bar{\theta} \le M(\theta_k - \bar{\theta}) \left[\left\| x^k - x^{k-1} \right\| + \left\| x^{k-1} - \bar{x} \right\| \right], \tag{4.5}$$

pour k suffisamment grand.

Démonstration.

Remarquons, tout d'abord, que $\bar{x}$ est une solution optimale du problème (P). De l'inégalité (4.4), nous savons que

$$\theta_{k+1} - \bar{\theta} \leq (\theta_k - \bar{\theta})(1 - \alpha_k), \tag{4.6}$$

où

$$\alpha_k = \min_{1 \leq l \leq p} \left[\frac{g_l(x^{k-1})}{g_l(x^k)} \right] \min_{1 \leq i \leq p} \left[\frac{g_i(\bar{x})}{g_i(x^{k-1})} \right]. \tag{4.7}$$

Lorsque k tend vers $+\infty$, la suite $\{x^k\}$ converge vers $\bar{x}$. Par conséquent, lorsque k tend vers $+\infty$, α_k tend vers 1. Dès lors, la suite $\{\theta_k\}$ converge superlinéairement vers $\bar{\theta}$.

Montrons, à présent, l'inégalité (4.5). Par hypothèse, nous savons que les fonctions g_i sont Lipschitziennes sur S. Par conséquent, il existe une constante N positive telle que, pour tout $x, y \in S$,

$$|g_i(x) - g_i(y)| \leq N \|x - y\|, \qquad i = 1, 2, \ldots, p.$$

D'un autre côté, il existe un nombre réel positif β tel que, pour tout $x \in S$,

$$0 < g_i(x) \leq \beta, \qquad i = 1, 2, \ldots, p.$$

D'où

$$\min_{1 \leq i \leq p} \left(\frac{g_i(x)}{g_i(y)} \right) \geq \max \left[0, 1 - \left(\frac{N}{\beta} \right) \|x - y\| \right],$$

pour tout $x, y \in S$. En effet,

$$\begin{aligned}
& |g_i(x) - g_i(y)| \leq N \|x - y\| \quad \text{et} \quad 0 < g_i(x) \leq \beta \quad \forall x \in S \\
\Rightarrow\ & |g_i(x) - \beta| \leq N \|x - y\| \\
\Leftrightarrow\ & \beta - g_i(x) \leq N \|x - y\| \\
\Leftrightarrow\ & g_i(x) \geq \beta - N \|x - y\| \\
\Leftrightarrow\ & \frac{g_i(x)}{g_i(y)} \geq \frac{\beta - N \|x - y\|}{g_i(y)} \\
\Leftrightarrow\ & \frac{g_i(x)}{g_i(y)} \geq \frac{\beta - N \|x - y\|}{\beta} = 1 - \left(\frac{N}{\beta} \right) \|x - y\|.
\end{aligned}$$

En injectant cette dernière inégalité dans l'expression de α_k, (4.7), nous obtenons

$$\alpha_k \geq \max \left(0, 1 - \left(\frac{N}{\beta} \right) [\|x^k - x^{k-1}\| + \|x^{k-1} - \bar{x}\|] \right).$$

En effet,

$$\begin{aligned}
\alpha_k &= \min_{1\leq l\leq p}\left[\frac{g_l(x^{k-1})}{g_l(x^k)}\right]\min_{1\leq i\leq p}\left[\frac{g_i(\bar{x})}{g_i(x^{k-1})}\right]\\
&\geq \max\left\{0, 1-\left(\frac{N}{\beta}\right)\left\|x^{k-1}-x^k\right\|\right\}\max\left\{0, 1-\left(\frac{N}{\beta}\right)\left\|\bar{x}-x^{k-1}\right\|\right\}\\
&\geq \max\left\{0, \underbrace{\left(1-\left(\frac{N}{\beta}\right)\left\|x^{k-1}-x^k\right\|\right)\left(1-\left(\frac{N}{\beta}\right)\left\|\bar{x}-x^{k-1}\right\|\right)}_{(*)}\right\}.
\end{aligned}$$

Remarquons que

- si $(*) \leq 0$ alors le maximum sera zéro,
- si $(*) > 0$ alors le maximum sera $(*)$.

Réécrivons l'expression (*) d'une autre façon :

$$\begin{aligned}
(*) &= 1-\left(\frac{N}{\beta}\right)\left\|\bar{x}-x^{k-1}\right\|-\left(\frac{N}{\beta}\right)\left\|x^{k-1}-x^k\right\|+\underbrace{\left(\frac{N}{\beta}\right)^2\left\|\bar{x}-x^{k-1}\right\|\left\|x^{k-1}-x^k\right\|}_{>0}\\
&\geq 1-\left(\frac{N}{\beta}\right)\left\|\bar{x}-x^{k-1}\right\|-\left(\frac{N}{\beta}\right)\left\|x^{k-1}-x^k\right\|\\
&= 1-\left(\frac{N}{\beta}\right)\left(\left\|\bar{x}-x^{k-1}\right\|+\left\|x^{k-1}-x^k\right\|\right).
\end{aligned}$$

Dès lors,

$$\alpha_k \geq \max\left\{0, 1-\left(\frac{N}{\beta}\right)\left(\left\|\bar{x}-x^{k-1}\right\|+\left\|x^{k-1}-x^k\right\|\right)\right\}.$$

Alors, pour k suffisamment grand,

$$0 \leq 1-\alpha_k \leq \left(\frac{N}{\beta}\right)\left(\left\|\bar{x}-x^{k-1}\right\|+\left\|x^{k-1}-x^k\right\|\right).$$

Afin de démontrer l'inégalité (4.5), il suffit alors de prendre

$$M = \frac{N}{\beta}.$$

□

Chapitre 5

Programmation fractionnelle généralisée dans le cas linéaire

Dans ce chapitre, nous allons supposer que l'ensemble S n'est pas obligatoirement borné. Considérons donc un ensemble S tel que

$$S = \{x \in \mathbb{R}^n \mid Cx \leq \gamma,\ x \geq 0\}.$$

Le problème fractionnel généralisé linéaire (P_L) peut alors s'écrire :

$$(P_L) \qquad \bar{\theta} = \inf \left\{ \max_{1 \leq i \leq p} \frac{a_i^T x + \alpha_i}{b_i^T x + \beta_i} \;\middle|\; Cx \leq \gamma,\ x \geq 0 \right\}$$

où a_i^T et b_i^T représentent respectivement la ligne i de la matrice A de taille $p \times n$ et la ligne i de la matrice B de taille $p \times n$. Notons a_j (respectivement b_j) la colonne j de la matrice A (respectivement B).
Notons également

$$\alpha = (\alpha_1, \cdots, \alpha_p)^T \quad \text{et} \quad \beta = (\beta_1, \cdots, \beta_p)^T.$$

De plus, C est une matrice $m \times n$ et $\gamma \in \mathbb{R}^m$.

Dans ce chapitre, nous nous intéresserons au cas où $\bar{\theta}$ est fini et nous supposerons également que

(H1) L'ensemble admissible $S = \left\{x \in \mathbb{R}^n \middle| Cx \leq \gamma,\ x \geq 0\right\}$ est non vide,

(H2) $B > 0$ (i.e. tous les éléments de la matrice B sont positifs) et $\beta > 0$.

5.1 Problème dual associé

Commençons par établir le problème dual associé au problème fractionnel généralisé linaire (P_L). Pour ce faire, nous aurons besoin du lemme suivant :

Lemme 5.1.1. Lemme de Farkas
Soient $D \in I\!R^{r \times n}$ et $d \in I\!R^r$. Alors, il existe $x \in I\!R^n$ tel que

$$Dx \leq d, \; x \geq 0$$

ou il existe $y \in I\!R^r$ tel que

$$D^T y \geq 0, \; y \geq 0, \; d^T y < 0$$

mais pas les deux.

Remarquons d'abord que

$$(P_L) \qquad \bar{\theta} = \inf \left\{ \max_{1 \leq i \leq p} \frac{a_i^T x + \alpha_i}{b_i^T x + \beta_i} \;\middle|\; Cx \leq \gamma, \, x \geq 0 \right\}$$

$$\Leftrightarrow$$

$$\bar{\theta} = \inf \left\{ t \;\middle|\; \frac{a_i^T x + \alpha_i}{b_i^T x + \beta_i} \leq t \mid Cx \leq \gamma, \; x \geq 0, \; t \in \mathbb{R} \right\}.$$

Les dénominateurs du problème (P_L) étant positifs, ce dernier peut s'écrire comme suit :

$$\inf \left\{ t \mid (A - tB)x \leq (-\alpha + t\beta), \; Cx \leq \gamma, x \geq 0, \; t \in \mathbb{R} \right\}.$$

En effet,

$$\begin{aligned}
& \frac{a_i^T x + \alpha_i}{b_i^T x + \beta_i} \leq t \qquad \forall i = 1, \cdots, p \\
\Leftrightarrow\ & a_i^T x + \alpha_i \leq t(b_i^T x + \beta_i) \qquad \forall i = 1, \cdots, p \quad \text{car} \quad b_i^T x + \beta_i > 0 \\
\Leftrightarrow\ & Ax + \alpha \leq t(Bx + \beta) \\
\Leftrightarrow\ & Ax - tBx \leq t\beta - \alpha \\
\Leftrightarrow\ & (A - tB)x \leq (-\alpha + t\beta).
\end{aligned}$$

Pour t fixé, définissons l'ensemble

$$S_t = \left\{x \in \mathbb{R}^n \mid (A - tB)x \leq (-\alpha + t\beta),\ Cx \leq \gamma, x \geq 0\right\}.$$

Remarquons que, lorsque $t \in \mathbb{R}$ est suffisamment grand, cet ensemble S_t est non vide puisque l'ensemble S est non vide (H1). Dès lors, la valeur optimale $\bar{\theta}$ du problème linéarisé (P_L) sera

$$\bar{\theta} = \inf\left\{t \mid S_t \neq \emptyset\right\}.$$

L'ensemble S_t étant défini par un système d'inégalités linéaires, nous pouvons appliquer le lemme de Farkas. Par conséquent, soit il existe $x \in \mathbb{R}^n$ tel que

$$(A - tB)x \leq (-\alpha + t\beta),\ Cx \leq \gamma,\ x \geq 0$$

soit il existe $y \in \mathbb{R}^{p+m}$ tel que

$$\left((A - tB)^T, C^T\right) y \geq 0,\ y \geq 0,\ \left((-\alpha + t\beta)^T, \gamma^T\right) y < 0.$$

Dès lors, par ce même lemme, nous pouvons conclure que l'ensemble S_t sera non vide lorsque l'ensemble

$$T_t = \left\{y \in \mathbb{R}^{p+m} \mid \left((A - tB)^T, C^T\right) y \geq 0,\ y \geq 0,\ \left((-\alpha + t\beta)^T, \gamma^T\right) y < 0\right\}$$

est vide. En prenant

$$y = \begin{pmatrix} u \\ w \end{pmatrix}, \qquad u \in \mathbb{R}^p, \qquad w \in \mathbb{R}^m,$$

nous pouvons écrire

$$T_t = \left\{\begin{pmatrix} u \\ w \end{pmatrix} \middle| (A - tB)^T u + C^T w \geq 0,\ u \geq 0,\ w \geq 0,\ (-\alpha + t\beta)^T u + \gamma^T w < 0\right\}.$$

Donc, puisque $\bar{\theta} = \inf\left\{t \mid S_t \neq \emptyset\right\}$, nous avons que

$$\bar{\theta} = \inf\left\{t \mid T_t = \emptyset\right\}.$$

Remarquons que

$$T_t = \emptyset \Rightarrow T_{t'} = \emptyset, \text{ pour tout } t' \geq t. \tag{5.1}$$

Montrons cela par contraposition :

$$\exists\, t' \geq t \;\text{ t.q. }\; T_{t'} \neq \emptyset \Rightarrow T_t \neq \emptyset.$$

Par hypothèse, nous savons qu'il existe $t' = t + \epsilon \geq t$ tel que $T_{t'} \neq \emptyset$ c'est-à-dire tel que

$$\exists\, y = \begin{pmatrix} u \\ w \end{pmatrix} \;\text{ t.q. }\; (A - t'B)^T u + C^T w \geq 0,\; u \geq 0,\; w \geq 0,\; (-\alpha + t'\beta)^T u + \gamma^T w < 0.$$

Montrons alors que $T_t \neq \emptyset$ c'est-à-dire

$$\exists\, y = \begin{pmatrix} u \\ w \end{pmatrix} \;\text{ t.q. }\; (A - tB)^T u + C^T w \overset{1}{\geq} 0,\; u \geq 0,\; w \geq 0,\; (-\alpha + t\beta)^T u + \gamma^T w \overset{2}{<} 0.$$

Pour ce faire, notons que les inégalités $u \geq 0$ et $w \geq 0$ sont satisfaites par l'hypothèse de contraposition. Les inégalités 1 et 2 sont également satisfaites. En effet,

1. $$\begin{aligned} (A - tB)^T u + C^T w &= (A - (t' - \epsilon)B)^T u + C^T w \\ &= \underbrace{(A - t'B)^T u + C^T w}_{\geq 0 \text{ par hyp}} + \underbrace{\epsilon B^T u}_{\geq 0} \geq 0, \end{aligned}$$
2. $$\begin{aligned} (-\alpha + t\beta)^T u + \gamma^T w &= (-\alpha + (t' - \epsilon)\beta)^T u + \gamma^T w \\ &= \underbrace{(-\alpha + t'\beta)^T u + \gamma^T w}_{<0} - \underbrace{\epsilon\beta^T u}_{\geq 0} < 0. \end{aligned}$$

Dès lors, par (5.1), nous avons

$$\begin{aligned} \bar{\theta} &= \sup \left\{ t \mid T_t \neq \emptyset \right\} \\ &= \sup \left\{ t \mid \exists\, u \geq 0, w \geq 0 \text{ t.q. } (A - tB)^T u + C^T w \geq 0, \right. \\ &\qquad \left. (-\alpha + t\beta)^T u + \gamma^T w < 0,\; t \in \mathbb{R} \right\}. \end{aligned}$$

Remarquons que, dans cette dernière égalité, $u \neq 0$, puisque sinon

$$C^T w \geq 0,\; w \geq 0,\; \gamma^T w < 0.$$

Ce qui implique, par le lemme de Farkas, qu'il n'existe pas de $x \in \mathbb{R}^n$ tel que $Cx \leq \gamma$, $x \geq 0$ et donc contredit l'hypothèse $(H1)$ à savoir qu'il existe $\hat{x} \geq 0$ tel que $C\hat{x} \leq \gamma$. Dès lors, nous obtenons le problème suivant :

$$\bar{\theta} = \sup \left\{ t \mid t\beta^T u < \alpha^T u - \gamma^T w,\; tB^T u \leq A^T u + C^T w,\; u \geq 0,\; u \neq 0,\; w \geq 0,\; t \in \mathbb{R} \right\}.$$

De plus, puisque $\beta^T u > 0$ et $B^T u > 0$, ce problème de maximisation peut s'écrire comme suit :

$$(D_L) \quad \bar{\theta} = \sup_{\substack{u \geq 0 \\ u \neq 0 \\ w \geq 0}} \left\{ \min \left[\frac{\alpha^T u - \gamma^T w}{\beta^T u}, \min_{1 \leq j \leq n} \frac{a_j^T u + c_j^T w}{b_j^T u} \right] \right\}.$$

Par conséquent, $\bar{\theta}$ étant la valeur optimale du problème (D_L), celui-ci peut être considéré comme le problème dual de (P_L) avec aucun saut de dualité entre les deux problèmes. Donc, si nous notons par $\bar{\lambda}$ la valeur optimale du problème dual associé à (P_L), nous avons $\bar{\theta} = \bar{\lambda}$.

Afin d'écrire ce problème autrement, effectuons le changement de variables suivant :

$$\bar{u} \leftarrow \frac{u}{\sum_i u_i}.$$

Le problème (D_L) s'écrit donc

$$\bar{\theta} = \sup_{\substack{u \geq 0 \\ u \neq 0 \\ w \geq 0}} \left\{ \min \left[\frac{\frac{\alpha^T u}{\sum u_i} - \frac{\gamma^T w}{\sum u_i}}{\frac{\beta^T u}{\sum u_i}}, \min_{1 \leq j \leq n} \frac{\frac{a_j^T u}{\sum u_i} + \frac{c_j^T w}{\sum u_i}}{\frac{b_j^T u}{\sum u_i}} \right] \right\}.$$

Or,

$$\begin{aligned} u \geq 0 \quad \text{et} \quad u \neq 0 &\Rightarrow \sum_i u_i \neq 0 \\ &\Rightarrow \sum_i \bar{u}_i \neq 0 \\ &\Rightarrow \sum_i \bar{u}_i = \frac{\sum_i u_i}{\sum_i u_i} = 1. \end{aligned}$$

D'où,

$$\bar{\theta} = \sup_{\substack{\bar{u} \geq 0 \\ \sum_i \bar{u}_i = 1}} \left\{ \sup_{w \geq 0} \left[\min \left(\frac{\alpha^T \bar{u} - \gamma^T w}{\beta^T \bar{u}}, \min_{1 \leq j \leq n} \frac{a_j^T \bar{u} + c_j^T w}{b_j^T \bar{u}} \right) \right] \right\}.$$

Ou encore,

$$\bar{\theta} = \sup_{\substack{u \geq 0 \\ e^T u = 1}} \left\{ \sup_{w \geq 0} \left[\min \left(\frac{\alpha^T u - \gamma^T w}{\beta^T u}, \min_{1 \leq j \leq n} \frac{a_j^T u + c_j^T w}{b_j^T u} \right) \right] \right\}.$$

Grâce à la dernière formulation du dual (D_L), nous pouvons observer que même si l'ensemble admissible S du problème primal (P_L) n'est pas borné, l'ensemble admissible du problème dual (D_L) est au moins borné en u.

5.2 Application de l'algorithme de Dinkelbach au problème dual

Dans cette section, nous allons observer les résultats obtenus lorsque nous appliquons l'algorithme de Dinkelbach décrit dans le chapitre 3 au problème dual (D_L) plutôt qu'au problème primal (P_L).

Afin de pouvoir appliquer l'algorithme, réécrivons le problème dual sous la forme d'un problème de minimisation

$$\bar{\theta} = -\inf_{\substack{u\geq 0\\ e^T u=1}} \left\{ \inf_{w\geq 0} \left[\max\left(\frac{-\alpha^T u + \gamma^T w}{\beta^T u}, \max_{1\leq j\leq n} \frac{-a_j^T u - c_j^T w}{b_j^T u} \right) \right] \right\}.$$

Soit $\bar{\mu} = -\bar{\theta}$. Ce problème de minimisation est un problème de programmation fractionnelle généralisée avec un ensemble admissible non borné. Sa F-fonction est donnée par

$$F_D(\mu) = \inf_{\substack{u\geq 0\\ e^T u=1}} \left\{ \inf_{w\geq 0} \left[\max\left(-\alpha^T u + \gamma^T w - \mu\beta^T u, \max_{1\leq j\leq n} (-a_j^T u - c_j^T w - \mu b_j^T u) \right) \right] \right\}.$$

D'où

$$F_D(\mu) = \inf_{\substack{u\geq 0\\ e^T u=1}} h(u;\mu),$$

où

$$h(u,\mu) = \inf_{w\geq 0} \left\{ \max\left[-\alpha^T u + \gamma^T w - \mu\beta^T u, \max_{1\leq j\leq n} (-a_j^T u - c_j^T w - \mu b_j^T u) \right] \right\}. \quad (5.2)$$

Propriété 5.2.1. *Supposons que l'ensemble admissible* $S = \{x \in I\!R^n | Cx \leq \gamma, x \geq 0\}$ *est non vide, que* $B > 0$ *et que* $\beta > 0$. *Supposons également que* $\bar{\theta}$ *est fini. Alors, la fonction* $F_D(\mu)$ *est finie, continue et décroissante pour tout* $\mu \in I\!R$.

Démonstration.

- Montrons que $F_D(u,\mu)$ est finie.

  Par (5.2), nous pouvons voir que $h(u;\mu) < \infty$, pour tout u, μ. De plus, $h(u;\mu) > -\infty$, pour tout u, μ puisque, par (H1) et le lemme de Farkas, nous savons qu'il existe $\hat{w} \geq 0$ tel que $C^T\hat{w} \geq 0$ et $\gamma^T\hat{w} < 0$. Par conséquent, $F_D(\mu)$ est fini.

- Montrons également que $F_D(\mu)$ est continue.

  Remarquons que la fonction $h(u;\mu)$ peut également s'écrire comme suit :

$$\begin{aligned} h(u,\mu) &= \inf_{\substack{w\geq 0\\ t\in I\!R}} \left\{ t \middle| t \geq -\alpha^T u + \gamma^T w - \mu\beta^T u, t \geq -a_j^T u - c_j^T w - \mu b_j^T u, j = 1,\cdots,n \right\} \\ &= \inf_{\substack{w\geq 0\\ t\in I\!R}} \left\{ 0^T w + t \middle| -\gamma^T w + t \geq (-\alpha - \mu\beta)^T u, \right. \qquad (5.3) \\ &\qquad\qquad \left. c_j^T w + t \geq (-a_j^T - \mu b_j^T)u, j = 1,\cdots,n \right\}. \end{aligned}$$

  Ainsi, $h(u;\mu)$ étant la valeur optimale d'un problème linéaire en u,μ, $h(u;\mu)$ est continue en u,μ. Dès lors, la fonction $F_D(\mu)$ est continue en u,μ puisqu'elle est l'infimum d'une fonction continue en u,μ.

- Montrons aussi que $F_D(\mu)$ est décroissante pour tout μ.

  Définissons l'ensemble $\Sigma = \left\{ u \mid u \geq 0, e^T u = 1 \right\}$. Cet ensemble est compact. De plus, $\beta^T u > 0$ et $b_j^T u > 0$ pour tout $u \in \Sigma$ et $j = 1,\cdots,p$. Dès lors, toutes les hypothèses de la proposition 3.3.2 étant satisfaites, nous pouvons conclure que $F_D(\mu)$ est décroissante.

□

Corollaire 5.2.1. *Si l'ensemble admissible $S = \left\{ x \in I\!R^n \middle| Cx \leq \gamma, x \geq 0 \right\}$ est non vide, si $B > 0$, si $\beta > 0$ et si $\bar{\theta}$ est finie, alors $\bar{\mu} = -\bar{\theta}$ est l'unique zéro de la fonction $F_D(\mu)$: i.e. $F_D(\bar{\mu}) = 0$.*

Démonstration.

Par la propriété 3.3.2, nous savons que $F_D(\bar{\mu}) \geq 0$ et que $F_D(\mu) < 0$ pour tout $\mu > \bar{\mu}$. Par conséquent, $F_D(\bar{\mu}) = 0$ car F_D est continue et finie. De plus, $\bar{\mu}$ est l'unique zéro de F_D car F_D est décroissante. □

Etudions, à présent, la convergence de cet algorithme.

Théorème 5.2.1. *La suite $\{\mu_k\}$, si elle n'est pas finie, converge linéairement vers $\bar{\mu} = -\bar{\theta}$.*

Démonstration.
Afin de démontrer cette convergence linéaire, vérifions toutes les hypothèses du corollaire 3.5.1 :

- Montrons que $M(\mu_k) \neq \emptyset,\ k = 1, 2, \cdots$.

  Par la propriété 3.3.4, une solution optimale $\bar{u}, \bar{w}$ de $F_D(\bar{\mu}) = 0$ est également une solution du problème dual (D_L). De plus, $F_D(\mu)$ est atteint pour tout μ en $\bar{u}(\mu), \bar{w}(\mu)$ où $\bar{w}(\mu)$ est la solution du problème linéaire (5.3).

- Nous devons également prouver que

$$\sup_{\substack{u \geq 0 \\ e^T u = 1 \\ w \geq 0}} \max\left(\beta^T u, \max_{1 \leq j \leq n} b_j^T u\right) < \infty.$$

  Cette hypothèse est vérifiée car on maximise sur l'ensemble $\{u \in \mathrm{IR} \mid u \geq 0 \quad \text{et} \quad e^T u = 1\}$ qui est borné.

  Par conséquent, en appliquant le corollaire 3.5.1, la suite $\{\mu_k\}$ converge linéairement vers $\bar{\mu} = -\bar{\theta}$.

$\square$

Exemple 5.2.1.
Soient

$$n = 1,\ p = 1,\ f(x) = 1,\ g(x) = 1 + x,\ S = \{x \mid -x \leq -1,\ x \geq 0\}.$$

Alors, le problème est le suivant :

$$(P_L) \qquad \inf_{x \in S} \left\{\frac{1}{(x+1)}\right\} = 0.$$

Par conséquent,

$$\begin{aligned} F(\theta) &= \inf_{x\in S}\{1-\theta(x+1)\} \\ &= \begin{cases} 1-2\theta, & \text{si} \quad \theta \leq 0, \\ -\infty, & \text{si} \quad \theta > 0. \end{cases} \end{aligned}$$

Puisque la fonction $F(\theta)$ n'est pas finie pour tout $\theta > \bar{\theta}$, nous ne pouvons pas appliquer l'algorithme au problème (P_L). Par ailleurs, ce problème satisfait toutes les conditions du corollaire 5.2.1. Par conséquent, la méthode convergera si nous l'appliquons au problème dual associé au problème (P_L).

Le problème dual (D_L) est le suivant :

$$\begin{aligned} (D_L) \qquad & \sup_{\substack{u\geq 0\\ u\neq 0\\ w\geq 0}} \left\{ \min \left[\frac{\alpha^T u - \gamma^T w}{\beta^T u}, \min_{1\leq j\leq n} \frac{a_j^T u + c_j^T w}{b_j^T u} \right] \right\} \\ = \qquad & \sup_{\substack{u=1\\ w\geq 0}} \left\{ \min \left[\frac{u+w}{u}, \frac{-w}{u} \right] \right\} \\ = \qquad & \sup_{\substack{u=1\\ w\geq 0}} \{\min[1+w, -w]\} \\ = \qquad & \sup \left\{ -w \middle| w \geq 0 \right\}. \end{aligned}$$

Alors, la fonction $F_D(\mu)$ peut s'écrire :

$$\begin{aligned} F_D(\mu) &= \inf_{w\geq 0} \left\{ \max\left[-\alpha^T u + \gamma^T w - \mu\beta^T u, \max\left(-a_j^T u - c_j^T w - \mu b_j^T u \right) \right] \right\} \\ &= \inf_{w\geq 0} \{\max[-1-w-\mu, w-\mu]\} \\ &= \inf \left\{ w-\mu \middle| w \geq 0 \right\} \\ &= -\mu \qquad \forall\, \mu \in \mathrm{I\!R}. \end{aligned}$$

D'où, $F_D(\mu)$ est finie et continue. Notre méthode converge donc en un pas pour tout $w^0 > 0$ arbitraire initial.

Chapitre 6

Méthodes de Dinkelbach de régularisation proximale

6.1 Méthode de point proximal inexacte

Considérons le problème de programmation fractionnelle généralisée suivant :

$$(P) \qquad \bar{\theta} = \inf_{x \in S} \left\{ \theta(x) = \max_{1 \leq i \leq p} \left[\frac{f_i(x)}{g_i(x)} \right] \right\}$$

où $S \subseteq \mathbb{R}^n$ est un ensemble convexe fermé non vide, $f_i, g_i : S \to \mathbb{R}$ sont des fonctions continues pour tout $1 \leq i \leq p$, $g_i > 0$ pour tout $x \in S$ et $1 \leq i \leq p$ et où $f_i - \theta g_i$ est convexe pour tout $\theta \geq \bar{\theta}$ de telle sorte que la fonction $F(\cdot, w^k, \theta_k)$ soit convexe sur S.

Soit (x^k, w^k, θ_k). La méthode de régularisation proximale consiste à remplacer le problème de minimisation

$$\min_{x \in S} F(x, w^k, \theta_k)$$

par le problème suivant

$$(P_{w^k, \theta_k, \alpha_k}) \qquad \min_{x \in S} \left\{ F(x, w^k, \theta_k) + \frac{1}{2\alpha_k} \left\| x - x^k \right\|^2 \right\},$$

où $\alpha_k > 0$.

Afin de pouvoir implémenter ce problème, nous ne calculerons qu'une approximation de la solution de ce dernier. Pour ce faire, nous allons approximer la fonction convexe non différentiable $F(\cdot, w^k, \theta_k)$ par une fonction convexe $\varphi(\cdot, w^k, \theta_k)$ de telle sorte qu'on puisse plus facilement résoudre le problème suivant :

$$(AP_{w^k,\theta_k,\alpha_k}) \qquad \min_{x\in S}\left\{\varphi(x, w^k, \theta_k) + \frac{1}{2\alpha_k}\left\|x - x^k\right\|^2\right\}.$$

Commençons par énoncer toutes les propriétés que les fonctions $\varphi(\cdot, w^k, \theta_k)$ doivent satisfaire de telle sorte que la suite $\{\theta_k\}$ converge vers la valeur optimale $\bar{\theta}$ du problème (P) et que la suite $\{x^k\}$ converge vers une solution de ce problème s'il en existe une.

Définition 6.1.1. Soit $c \in (0, 1)$ et soit $w^k > 0$, $\theta_k \geq \bar{\theta}$ et soit $x^k \in S$. Une fonction convexe $\varphi(\cdot, w^k, \theta_k)$ est une c-approximation de $F(\cdot, w^k, \theta_k)$ au point x^k si $\varphi(x, w^k, \theta_k) \leq F(x, w^k, \theta_k)$ pour tout $x \in S$, et si

$$\varphi(x^{k+1}, w^k, \theta_k) \geq \frac{1}{c}F(x^{k+1}, w^k, \theta_k), \tag{6.1}$$

où x^{k+1} est la solution du problème $(AP_{w^k,\theta_k,\alpha_k})$.

Les c-approximations vérifient les propriétés suivantes :

Propriété 6.1.1. *Si $\varphi(\cdot, w^k, \theta_k)$ est une c-approximation de $F(\cdot, w^k, \theta_k)$ au point x^k, alors, au point x^{k+1}, nous avons*

$$\frac{1}{c}F(x^{k+1}, w^k, \theta_k) \leq \varphi(x^{k+1}, w^k, \theta_k) \leq F(x^{k+1}, w^k, \theta_k).$$

En particulier, puisque $c \in (0, 1)$, nous avons

$$\varphi(x^{k+1}, w^k, \theta_k) \leq F(x^{k+1}, w^k, \theta_k) \leq 0. \tag{6.2}$$

L'algorithme de Dinkelbach devient alors :

Algorithme de point proximal inexact

Pas 0 : Choisir $x^0 \in S$, $w^0 > 0$, $\alpha_0 > 0$, $c \in (0,1)$ et prendre $\theta_0 = \theta(x^0)$, et $k = 0$.

Pas 1 : Construire une c-approximation de $F(\cdot, w^k, \theta_k)$ et trouver $x^{k+1} \in S$, l'unique solution du problème

$$(AP_{w^k,\theta_k,\alpha_k}) \qquad \min_{x\in S}\left\{\varphi(x, w^k, \theta_k) + \frac{1}{2\alpha_k}\left\|x - x^k\right\|^2\right\}.$$

Pas 2 : Calculer $\theta_{k+1} = \max\limits_i \dfrac{f_i(x^{k+1})}{g_i(x^{k+1})}$, choisir w^{k+1}, $\alpha_{k+1} > 0$.

Pas 3 : Remplacer k par $k+1$ et retourner au pas 1.

Détaillons brièvement l'interprétation de cet algorithme. Nous pouvons d'abord observer que w^k n'intervient pas dans le calcul de θ_k et que pour tout $w > 0$, nous avons

$$\theta_{k+1} = \max_i \frac{f_i(x^{k+1})}{g_i(x^{k+1})} \Leftrightarrow F(x^{k+1}, w, \theta_k) = 0.$$

Par conséquent, à chaque itération, nous avons que

$$F(x^k, w^k, \theta_k) = 0.$$

De plus, une c-approximation de $F(\cdot, w^k, \theta_k)$ au point x^k satisfait la propriété

$$F(x^k, w^k, \theta_k) - F(x^{k+1}, w^k, \theta_k) \geq c\left[F(x^k, w^k, \theta_k) - \varphi(x^{k+1}, w^k, \theta_k)\right]. \qquad (6.3)$$

Cette inégalité nous permet de conclure que la décroissance de F entre les points x^k et x^{k+1} (c'est-à-dire le membre de gauche de l'inégalité) est plus importante qu'une fraction de la décroissance prédite par le modèle φ (c'est-à-dire le membre de droite de l'inégalité).

6.2 Convergence

Afin d'étudier la convergence de la suite $\{\theta_k\}$, introduisons les notations suivantes. Pour $x \in S$, $w > 0$, et θ, définissons les ensembles :

- $K(x, \theta) = \{j \mid f_j(x) - \theta g_j(x) = F(x, \theta)\}$

– $K(x,w,\theta) = \left\{ j \;\middle|\; \frac{f_j(x)-\theta g_j(x)}{w_j} = F(x,w,\theta) \right\}$

Rappelons également que l'ensemble $J(x)$ est défini par

$$J(x) = \left\{ i \;\middle|\; \frac{f_i(x)}{g_i(x)} = \theta(x) \right\}.$$

Proposition 6.2.1. *Soit $c \in (0,1)$. Alors*

1. *la suite $\{\theta_k\}$ est décroissante et converge vers $\hat{\theta} \geq \bar{\theta}$;*
2. *si $\bar{\theta} > -\infty$ et si $g_i(x^k) \leq \gamma$ et $w_i^k \geq \underline{w} > 0$ pour tout k et $1 \leq i \leq p$, alors $F(x^{k+1}, w^k, \theta_k) \to 0$.*

Démonstration.

1. Puisque, par définition, nous avons

$$F(x^{k+1}, w^k, \theta_k) = \max_{1\leq i \leq p} \left\{ \frac{f_i(x^{k+1}) - \theta_k g_i(x^{k+1})}{w_i^k} \right\},$$

nous obtenons donc, pour tout $1 \leq i \leq p$, que

$$F(x^{k+1}, w^k, \theta_k) \geq \frac{f_i(x^{k+1}) - \theta_k g_i(x^{k+1})}{w_i^k}. \tag{6.4}$$

Soit $i^* \in J(x^{k+1})$. Ainsi,

$$\begin{aligned} & \theta(x^{k+1}) = \tfrac{f_{i^*}(x^{k+1})}{g_{i^*}(x^{k+1})} \\ \Leftrightarrow \quad & f_{i^*}(x^{k+1}) = \theta(x^{k+1}) g_{i^*}(x^{k+1}). \end{aligned} \tag{6.5}$$

Ainsi, par (6.4) et (6.5), nous obtenons

$$\begin{aligned} 0 & \overset{(6.2)}{\geq} F(x^{k+1}, w^k, \theta_k) \geq \tfrac{f_{i^*}(x^{k+1}) - \theta_k g_{i^*}(x^{k+1})}{w_{i^*}^k} && \forall\, i^* \in J(x^{k+1}) \\ & \overset{(6.5)}{=} \tfrac{1}{w_{i^*}^k} \left(\theta_{k+1} g_{i^*}(x^{k+1}) - \theta_k g_{i^*}(x^{k+1}) \right) && \forall\, i^* \in J(x^{k+1}) \\ & = \underbrace{\frac{g_{i^*}(x^{k+1})}{w_{i^*}^k}}_{\geq 0 \text{ par hypothèse}} (\theta_{k+1} - \theta_k) && \forall\, i^* \in J(x^{k+1}). \end{aligned} \tag{6.6}$$

Par conséquent,

$$\begin{aligned} & \theta_{k+1} - \theta_k \leq 0 \\ \Leftrightarrow\ & \theta_{k+1} \leq \theta_k \\ \Leftrightarrow\ & \{\theta_k\} \text{ est décroissante.} \end{aligned}$$

Dès lors,

$$\begin{aligned}
\left\|x^{k+1}-x\right\|^2 &= \left\langle x^{k+1}-x, x^{k+1}-x\right\rangle \\
&= \left\langle x^{k+1}-x^k+x^k-x, x^{k+1}-x^k+x^k-x\right\rangle \\
&= \left\langle x^{k+1}-x^k, x^{k+1}-x^k\right\rangle + \left\langle x^k-x, x^k-x\right\rangle + 2\underbrace{\left\langle x^{k+1}-x^k, x^k-x\right\rangle}_{\stackrel{(6.10)}{\leq}\alpha_k\left[F(x,w^k,\theta_k)+\epsilon_k\right]} \\
&\leq \left\|x^{k+1}-x^k\right\|^2 + \left\|x^k-x, x^k-x\right\|^2 + 2\alpha_k\left[F(x,w^k,\theta_k)+\epsilon_k\right].
\end{aligned}$$

3. Développons d'abord deux inégalités :
   - Par la définition de ϵ_k et par (6.1), nous savons que

$$\epsilon_k + \alpha_k^{-1}\left\|x^{k+1}-x^k\right\|^2 \stackrel{\text{déf de } \epsilon_k}{=} -\varphi(x^{k+1},w^k,\theta_k) \stackrel{(6.1)}{\leq} -\frac{1}{c}F(x^{k+1},w^k,\theta_k).$$

   - En remplaçant k par $k-1$ dans l'inégalité (6.7), nous obtenons que

$$0 \leq F(x^k,w^{k-1},\theta_{k-1}) + \frac{\gamma}{\underline{w}}\left(\theta_{k-1}-\theta_k\right).$$

En combinant ces deux inégalités, nous avons que

$$\epsilon_k + \alpha_k^{-1}\left\|x^{k+1}-x^k\right\|^2 \leq \frac{1}{c}\left[F(x^k,w^{k-1},\theta_{k-1}) - F(x^{k+1},w^k,\theta_k)\right] + \frac{\gamma}{\underline{w}c}\left[\theta_{k-1}-\theta_k\right].$$

Sommons, à présent, cette inégalité pour $k=1$ jusque $k=q$:

$$\sum_{k=1}^{q}\left\{\epsilon_k + \alpha_k^{-1}\left\|x^{k+1}-x^k\right\|^2\right\} \leq \frac{1}{c}\left[F(x^1,w^0,\theta_0) - F(x^{q+1},w^q,\theta_q)\right] + \frac{\gamma}{\underline{w}c}\left[\theta_0-\theta_q\right].$$

Or, par la proposition précédente, nous savons que $F(x^{q+1},w^q,\theta_q)\to 0$ et que $\theta_q\to\hat{\theta}$. Par conséquent, la série $\sum_{k\geq 1}\left[\epsilon_k + \alpha_k^{-1}\left\|x^{k+1}-x^k\right\|^2\right]$ est convergente. □

Etudions, à présent, la convergence de la suite $\{\theta_k\}$ vers $\bar{\theta}$.

Théorème 6.2.1. *Soit $c\in(0,1)$. Supposons que $0<\nu\leq g_i(x^k)\leq\gamma$ et que $0<\underline{w}\leq w_i^k\leq\bar{w}$ pour tout k et $1\leq i\leq p$. Supposons aussi que $\sum_{k\geq 0}\alpha_k=+\infty$ et que soit $\alpha_k\leq\bar{\alpha}$ pour tout k ou soit $\alpha_k\leq\alpha_{k+1}$ pour tout k. Alors, la suite $\{\theta_k\}$ générée par l'algorithme de point proximal inexact converge vers $\bar{\theta}$, la valeur optimale du problème (P).*

Démonstration.

Puisque, par la proposition 6.2.1, nous savons que la suite $\{\theta_k\}$ converge vers $\hat{\theta} \geq \bar{\theta}$, il ne nous reste plus qu'à montrer que $\hat{\theta} = \bar{\theta}$.

Remarquons d'abord que, lorsque $\hat{\theta} = -\infty$, $\bar{\theta} = -\infty$ et donc $\hat{\theta} = \bar{\theta}$. Par conséquent, supposons que $\hat{\theta} > -\infty$ et montrons que $\hat{\theta} = \bar{\theta}$ dans ce cas.

Soit $x \in S$. Pour $j \in K(x, w^k, \theta_k)$, nous avons que

$$F(x, w^k, \theta_k) = \frac{(f_j(x) - \theta_k g_j(x))}{w_j^k},$$

En combinant cette dernière égalité avec $\theta(x) \geq \frac{f_j(x)}{g_j(x)}$, nous obtenons

$$F(x, w^k, \theta_k) \leq (\theta(x) - \theta_k)\, \frac{g_j(x)}{w_j^k}.$$

Or, par hypothèse, nous savons que $g_i(x) \geq \nu$ et que $w_j^k \leq \bar{w}$. D'où

$$F(x, w^k, \theta_k) \leq (\theta(x) - \theta_k)\, \frac{\nu}{\bar{w}} \qquad \text{si} \quad \theta(x) - \theta_k \leq 0. \tag{6.11}$$

Montrons, maintenant, que, pour tout $x \in S$, nous avons

$$\limsup_{k \to \infty} \left(F(x, w^k, \theta_k) \right) \geq 0. \tag{6.12}$$

Pour ce faire, supposons, par l'absurde, que cette inégalité n'est pas vérifiée c'est-à-dire que

$$\exists\, \epsilon > 0,\ \tilde{x} \in S \text{ et } k_\epsilon \text{ t.q. } F(\tilde{x}, w^k, \theta_k) < -\epsilon \quad \forall\, k \geq k_\epsilon. \tag{HA}$$

Par le point 2 du lemme 6.2.1 avec $x = \tilde{x}$, nous obtenons, pour tout $k \geq k_\epsilon$, que

$$\begin{aligned} \left\|x^{k+1} - \tilde{x}\right\|^2 &\leq \left\|x^k - \tilde{x}\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\alpha_k \left[F(\tilde{x}, w^k, \theta_k) + \epsilon_k\right] \\ &\overset{(HA)}{<} \left\|x^k - \tilde{x}\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\alpha_k (-\epsilon + \epsilon_k) \\ &= \left\|x^k - \tilde{x}\right\|^2 + 2\alpha_k \left[\epsilon_k + (2\alpha_k)^{-1} \left\|x^{k+1} - x^k\right\|^2\right] - 2\alpha_k \epsilon. \end{aligned} \tag{6.13}$$

Avant de continuer cette preuve, citons d'abord un lemme qui nous sera utile par la suite :

Lemme 6.2.2. *Soit* $\{u^k\} \subseteq \mathbb{R}^+$. *Si* $u^{k+1} \leq u^k + \epsilon_k$ *et si* $\sum_{k \geq 1} \epsilon_k < \infty$. *Alors* $u^k \to u$.

Considérons, à présent, deux cas distincts :

- Si $\alpha_k \leq \bar{\alpha}$. Puisque $2\alpha_k\epsilon > 0$, nous pouvons déduire de l'inégalité (6.13)

$$\left\|x^{k+1} - \tilde{x}\right\|^2 \leq \left\|x^k - \tilde{x}\right\|^2 + 2\bar{\alpha}\left[\epsilon_k + (2\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2\right].$$

Puisque, par le lemme 6.2.1, la série $\sum_{k\geq 1}\left[\epsilon_k + (2\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2\right]$ est convergente et par le lemme 6.2.2, nous pouvons dire que la suite $\{\left\|x^k - \tilde{x}\right\|^2\}$ converge vers $u \geq 0$.

En sommant l'inégalité (6.13) pour $k = k_\epsilon$ jusque $k = q$ et en utilisant $\alpha_k \leq \bar{\alpha}$, nous obtenons

$$\left\|x^{q+1} - \tilde{x}\right\|^2 - \left\|x^{k_\epsilon} - \tilde{x}\right\|^2 \leq 2\bar{\alpha}\sum_{k=k_\epsilon}^{q}\left[\epsilon_k + (\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2\right] - 2\epsilon\sum_{k=k_\epsilon}^{q}\alpha_k.$$

En faisant tendre q vers l'infini et en utilisant l'hypothèse $\sum_{k\geq 0}\alpha_k = +\infty$, nous obtenons

$$\begin{aligned} u - \left\|x^{k_\epsilon} - \tilde{x}\right\|^2 &\leq 2\bar{\alpha}\sum_{k=k_\epsilon}^{+\infty}\left[\epsilon_k + (\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2\right] - 2\epsilon\underbrace{\sum_{k=k_\epsilon}^{+\infty}\alpha_k}_{=+\infty} \\ &< -\infty. \end{aligned}$$

Cette dernière inégalité n'est pas possible. Par conséquent, ce que nous avions supposé n'est pas correct et nous avons donc

$$\limsup_{k\to\infty}\left(F(x, w^k, \theta_k)\right) \geq 0.$$

- Si $\alpha_k \leq \alpha_{k+1}$ pour tout k. En divisant les deux membres de l'inégalité (6.13) par $2\alpha_k$, nous obtenons

$$\begin{aligned} (2\alpha_k)^{-1}\left\|x^k - \tilde{x}\right\|^2 + \left[\epsilon_k + (2\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|\right] - \epsilon \geq\ & (2\alpha_k)^{-1}\left\|x^{k+1} - \tilde{x}\right\|^2 \\ \geq\ & (2\alpha_{k+1})^{-1}\left\|x^{k+1} - \tilde{x}\right\|^2. \quad (6.14) \end{aligned}$$

Par un raisonnement similaire au raisonnement effectué au point précédent, nous pouvons conclure que la suite $\left\{(2\alpha_k)^{-1}\left\|x^k - \tilde{x}\right\|^2\right\}$ converge vers $u \geq 0$ car $\epsilon > 0$ et car la série $\sum_{k\geq 1}\left[\epsilon_k + (2\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2\right]$ est convergente.

En sommant l'inégalité (6.14) pour $k = k_\epsilon$ jusque $k = q$, nous obtenons

$$(2\alpha_{q+1})^{-1} \left\|x^{q+1} - \tilde{x}\right\|^2 - (2\alpha_{k_\epsilon})^{-1} \left\|x^{k_\epsilon} - \tilde{x}\right\|^2$$
$$\leq \sum_{k=k_\epsilon}^{q} \left[\epsilon_k + (2\alpha_k)^{-1} \left\|x^{k+1} - x^k\right\|^2\right] - \epsilon(q - k_\epsilon).$$

En faisant tendre q vers l'infini, nous obtenons

$$u - (2\alpha_{k_\epsilon})^{-1} \left\|x^{k_\epsilon} - \tilde{x}\right\|^2 < -\infty,$$

ce qui n'est pas possible. Par conséquent, ce que nous avions supposé est faux et nous avons donc

$$\limsup_{k\to\infty} \left(F(x, w^k, \theta_k)\right) \geq 0.$$

En considérant le fait que, pour tout $x \in S$, $\limsup\limits_{k\to\infty} \left(F(x, w^k, \theta_k)\right) \geq 0$, montrons que $\hat{\theta} = \bar{\theta}$:

- D'abord, si $\theta(x) \geq \theta_k$ pour un ensemble infini d'indices k, alors $\theta(x) \geq \hat{\theta}$ car $\theta_k \geq \hat{\theta}$.
- D'un autre côté, $\theta(x) < \theta_k$ pour tout $k \geq k_0$. Dès lors, en reprenant l'inégalité (6.11) et en utilisant (6.12), nous avons pour $k \geq k_0$

$$0 \leq \limsup_{k\to\infty} \left(F(x, w^k, \theta_k)\right) \leq \limsup_{k\to\infty} \left((\theta(x) - \theta_k) \underbrace{\frac{\nu}{\bar{w}}}_{>0} \right).$$

Donc

$$\theta(x) \geq \theta_k.$$

D'où,

$$\theta(x) \geq \theta_k \geq \hat{\theta}.$$

Par conséquent, dans les deux cas, nous avons

$$\theta(x) \geq \hat{\theta}.$$

Finalement, puisque x est arbitraire, nous avons que

$$\bar{\theta} \geq \hat{\theta}$$

et donc

$$\hat{\theta} = \bar{\theta}.$$

$\square$

Démontrons, à présent, la convergence de la suite $\{x^k\}$. Pour ce faire, nous aurons besoin du lemme suivant :

Lemme 6.2.3. *Soit $\bar{z}$ une valeur d'adhérence de la suite $\{z^k\}$ qui satisfait :*

$$\left\|z^{k+1} - \bar{z}\right\|^2 \leq \left\|z^k - \bar{z}\right\|^2 + \delta_k$$

où $\{\delta_k\}$ est une suite de nombres non négatifs telle que $\sum_{k\geq 0} \delta_k < +\infty$.

Alors, toute la suite $\{z^k\}$ converge vers $\bar{z}$.

Notons néanmoins que, pour démontrer la convergence de la suite $\{\theta_k\}$, aucune hypothèse n'avait été nécessaire sur l'existence ou non d'une solution au problème (P). Par contre, pour la convergence de la suite $\{x^k\}$, il est nécessaire de supposer que le problème (P) possède une solution.

Théorème 6.2.2. *Soit $c \in (0,1)$. Supposons que $0 < \nu \leq g_i(x^k) \leq \gamma$ et que $0 < \underline{w} \leq w_i^k \leq \bar{w}$ pour tout k et $1 \leq i \leq p$. Supposons aussi que $\sum_{k\geq 0} \alpha_k = +\infty$ et que soit $\alpha_k \leq \bar{\alpha}$ pour tout k ou soit $\alpha_k \leq \alpha_{k+1}$ pour tout k. Alors*

1. *toute valeur d'adhérence de la suite $\{x^k\}$ est solution du problème (P) ;*
2. *si $\alpha_k \leq \bar{\alpha}$ pour tout k et si l'ensemble des solutions du problème (P) est non vide, alors la suite $\{x^k\}$ converge vers une solution de (P).*

Démonstration.

1. Notons par x^* une valeur d'adhérence de la suite $\{x^k\}$. Alors, il existe une sous-suite $\{x^{n_k}\}$ qui converge vers x^* et puisque $\theta(x)$ est une fonction continue, nous avons que

$$\theta(x^{n_k}) \to \theta(x^*).$$

Or, par le lemme 6.2.1, nous savons que

$$\theta(x^{n_k}) = \theta_{n_k} \to \bar{\theta}.$$

Par unicité de la limite, nous avons

$$\theta(x^*) = \bar{\theta}$$

et x^* est une solution du problème (P).

2. Montrons d'abord que la suite $\{x^k\}$ est bornée. Soit $\bar{x}$ une solution du problème (P). Alors

$$F(\bar{x}, w^k, \theta_k) \leq 0.$$

En effet, comme $\theta_k \geq \bar{\theta} = \theta(\bar{x}) = \max_i \frac{f_i(\bar{x})}{g_i(\bar{x})}$, nous avons

$$\begin{aligned}
\theta_k \geq \frac{f_i(\bar{x})}{g_i(\bar{x})} &\Leftrightarrow f_i(\bar{x}) - \theta_k g_i(\bar{x}) \leq 0 \\
&\Rightarrow \frac{f_i(\bar{x}) - \theta_k g_i(\bar{x})}{w_i} \leq 0 \quad \text{car} \quad w_i > 0 \quad \forall\, i \\
&\Rightarrow \max_i \left\{ \frac{f_i(\bar{x}) - \theta_k g_i(\bar{x})}{w_i} \right\} \leq 0 \\
&\Leftrightarrow F(\bar{x}, w^k, \theta_k) \leq 0.
\end{aligned}$$

Par le second point du lemme 6.2.1, en prenant $x = \bar{x}$, nous avons

$$\begin{aligned}
\left\|x^{k+1} - \bar{x}\right\|^2 &\leq \left\|x^k - \bar{x}\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\alpha_k \left[\underbrace{F(\bar{x}, w^k, \theta_k)}_{\leq 0} + \epsilon_k\right] \\
&\leq \left\|x^k - \bar{x}\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\alpha_k \epsilon_k \\
&= \left\|x^k - \bar{x}\right\|^2 + 2 \underbrace{\alpha_k}_{\leq \bar{\alpha}} \left[\epsilon_k + (2\alpha_k)^{-1} \left\|x^{k+1} - x^k\right\|^2\right] \\
&\leq \left\|x^k - \bar{x}\right\|^2 + 2\bar{\alpha} \left[\epsilon_k + (2\alpha_k)^{-1} \left\|x^{k+1} - x^k\right\|^2\right].
\end{aligned}$$

Or, par le troisième point du lemme 6.2.1, nous savons que la série suivante est convergente

$$\sum_{k \geq 1} \left[\epsilon_k + (2\alpha_k)^{-1} \left\|x^{k+1} - x^k\right\|^2\right].$$

Par conséquent, par le lemme 6.2.2, la suite $\left\{\left\|x^k - \bar{x}\right\|^2\right\}$ converge. La suite $\{x^k\}$ est donc bornée.

De plus, par la définition de c-approximation, nous savons que

$$\varphi(\tilde{x}^k, w^k, \theta_k) \leq F(\tilde{x}^k, w^k, \theta_k) \qquad \forall\, \tilde{x}^k \in S$$

et que

$$\varphi(x^{k+1}, w^k, \theta_k) \geq \frac{1}{c} F(x^{k+1}, w^k, \theta_k) \quad \text{si} \quad x^{k+1} \text{ est solution de } (AP_{w^k,\theta_k,\alpha_k}).$$

D'où, (6.15) s'écrit

$$F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1} \left\|\tilde{x}^k - x^k\right\|^2 \geq \frac{1}{c} F(x^{k+1}, w^k, \theta_k). \tag{6.16}$$

Or, en utilisant (6.7), nous obtenons finalement

$$F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1} \left\|\tilde{x}^k - x^k\right\|^2 \geq \frac{1}{c} F(x^{k+1}, w^k, \theta_k) \geq \frac{\gamma}{\underline{w}c} \left(\theta_{k+1} - \theta_k\right). \tag{6.17}$$

Puisque la suite $\{x^k\}$ converge vers $x^* \in S^*$, nous avons que $x^k \in B(S^*, \delta) \cap S$ pour k suffisamment grand. Dès lors, en utilisant les hypothèses de ce théorème, nous obtenons

$$F(x^k, \bar{\theta}) \geq \kappa \left\|x^k - \tilde{x}^k\right\|^2 \text{ pour } k \text{ suffisamment grand.} \tag{6.18}$$

Soit $j \in K(x^k, \bar{\theta})$. Dès lors,

$$F(x^k, \bar{\theta}) = f_j(x^k) - \bar{\theta} g_j(x^k). \tag{6.19}$$

Or, par la propriété 3.3.2, nous avons, pour $\theta_k \geq \bar{\theta}$, que $f_j(x^k) \leq \theta_k g_j(x^k)$. Par conséquent, (6.19) peut s'écrire

$$\begin{aligned} F(x^k, \bar{\theta}) &= \underbrace{f_j(x^k)}_{\leq \theta_k g_j(x^k)} - \bar{\theta} g_j(x^k) \\ &\leq \underbrace{(\theta_k - \bar{\theta})}_{\geq 0} \underbrace{g_j(x^k)}_{\leq \gamma} \\ &\leq \gamma(\theta_k - \bar{\theta}). \end{aligned} \tag{6.20}$$

En combinant les résultats (6.18) et (6.20), nous obtenons pour k suffisamment grand

$$\kappa \left\|x^k - \tilde{x}^k\right\|^2 \leq \gamma(\theta_k - \bar{\theta}). \tag{6.21}$$

Raisonnons d'une façon similaire pour un indice $j \in K(\tilde{x}^k, w^k, \theta_k)$. Dès lors, pour un tel indice, nous avons

$$F(\tilde{x}^k, w^k, \theta_k) = \frac{1}{w_j^k} \left[f_j(\tilde{x}^k) - \theta_k g_j(\tilde{x}^k) \right] \tag{6.22}$$

et puisque $f_j(\tilde{x}^k) \leq \theta(\tilde{x}^k) g_j(\tilde{x}^k) = \bar{\theta} g_j(\tilde{x}^k)$, nous obtenons

$$\begin{aligned} F(\tilde{x}^k, w^k, \theta_k) &\leq \underbrace{(\bar{\theta} - \theta_k)}_{\leq 0} \frac{\overbrace{g_j(\tilde{x}^k)}^{\geq \tau}}{\underbrace{w_j^k}_{\leq \bar{w}}} \\ &\leq \tfrac{\tau}{\bar{w}} \left(\bar{\theta} - \theta_k\right). \end{aligned} \tag{6.23}$$

En utilisant le résultat (6.23), nous obtenons pour k suffisamment grand

$$\frac{\tau}{\bar{w}}(\bar{\theta} - \theta_k) + (2\alpha_k)^{-1} \left\|\tilde{x}^k - x^k\right\|^2 \geq F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1} \left\|\tilde{x}^k - x^k\right\|^2, \tag{6.24}$$

et en utilisant cette dernière inégalité avec (6.21) et (6.17), nous obtenons

$$\frac{\tau}{\bar{w}}(\bar{\theta} - \theta_k) + (2\alpha_k)^{-1} \frac{\gamma}{\kappa}(\theta_k - \bar{\theta}) \geq \frac{\gamma}{\underline{w}c}(\theta_{k+1} - \theta_k). \tag{6.25}$$

Ecrivons cette dernière inégalité différemment :

$$\begin{aligned} \frac{\tau}{\bar{w}}(\bar{\theta} - \theta_k) + (2\alpha_k)^{-1} \frac{\gamma}{\kappa}(\theta_k - \bar{\theta}) &\geq \frac{\gamma}{\underline{w}c}(\theta_{k+1} - \theta_k) \\ &= \frac{\gamma}{\underline{w}c}(\theta_{k+1} - \bar{\theta} + \bar{\theta} - \theta_k). \end{aligned}$$

En divisant cette dernière expression par $\frac{\gamma}{\underline{w}c}$ et en arrangeant les termes, nous obtenons :

$$\begin{aligned} & \frac{\underline{w}c}{\gamma} \left[\frac{\tau}{\bar{w}}(\bar{\theta} - \theta_k) + (2\alpha_k)^{-1} \frac{\gamma}{\kappa}(\theta_k - \bar{\theta})\right] \geq (\theta_{k+1} - \bar{\theta} + \bar{\theta} - \theta_k) \\ \Leftrightarrow\ & \frac{\underline{w}c}{\gamma} \left[\frac{\tau}{\bar{w}}(\bar{\theta} - \theta_k) + (2\alpha_k)^{-1} \frac{\gamma}{\kappa}(\theta_k - \bar{\theta})\right] + (\theta_k - \bar{\theta}) \geq (\theta_{k+1} - \bar{\theta}) \\ \Leftrightarrow\ & \underbrace{\left[c\left(-\frac{\tau \underline{w}}{\gamma \bar{w}} + \frac{\underline{w}}{2\alpha_k \kappa}\right) + 1\right]}_{(*)} (\theta_k - \bar{\theta}) \geq (\theta_{k+1} - \bar{\theta}). \end{aligned}$$

Afin de prouver la convergence linéaire, il nous reste à montrer que $(*)$ est compris entre 0 et 1 :

- Montrons d'abord que $1 + c\left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}}\right) < 1$.

  Si $\liminf_{k \to \infty} \alpha_k > \dfrac{\gamma \bar{w}}{2\kappa\tau}$, alors

$$\begin{aligned} & \frac{\underline{w}}{2\kappa} \frac{2\kappa\tau}{\gamma \bar{w}} > \frac{\underline{w}}{2\alpha_k \kappa} \\ \Leftrightarrow\ & \frac{\underline{w}\tau}{\gamma \bar{w}} > \frac{\underline{w}}{2\alpha_k \kappa} \qquad (\star) \\ \Leftrightarrow\ & \left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\underline{w}\tau}{\gamma \bar{w}}\right) < 0. \end{aligned}$$

Comme $c \in (0,1)$, nous obtenons

$$c\left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}}\right) < 0.$$

Par conséquent,

$$c\left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}}\right) + 1 < 1.$$

- Il nous reste donc à montrer que $1 + c\left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}}\right) > 0$.

  Développons cette inégalité afin d'arriver à une expression qui est toujours correcte :

$$\begin{aligned} & c\left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}}\right) > -1 \\ \Leftrightarrow\ & \frac{c\underline{w}}{2\alpha_k \kappa} - \frac{c\tau \underline{w}}{\gamma \bar{w}} > -1 \\ \Leftrightarrow\ & \frac{c\underline{w}}{2\alpha_k \kappa} > -1 + \frac{c\tau \underline{w}}{\gamma \bar{w}} \\ \Leftrightarrow\ & \underbrace{\frac{c\underline{w}}{2\alpha_k \kappa}}_{< \frac{c\underline{w}\tau}{\gamma \bar{w}} \text{ par } (\star)} > \frac{-\gamma \bar{w} + c\tau \underline{w}}{\gamma \bar{w}} \\ \Leftrightarrow\ & \frac{c\underline{w}\tau}{\gamma \bar{w}} > \frac{-\gamma \bar{w} + c\tau \underline{w}}{\gamma \bar{w}} \\ \Leftrightarrow\ & c\underline{w}\tau > -\gamma \bar{w} + c\tau \underline{w} \\ \Leftrightarrow\ & 0 > -\underbrace{\gamma \bar{w}}_{>0} \end{aligned}$$

  Cette dernière inégalité étant toujours vérifiée, nous avons donc que

$$c\left(\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}}\right) > -1.$$

Par conséquent, la suite $\{\theta_k\}$ converge linéairement vers $\bar{\theta}$ lorsque α_k est suffisamment grand. □

Théoriquement, puisque $c \in (0,1)$ et puisque $\frac{\underline{w}}{2\alpha_k \kappa} - \frac{\tau \underline{w}}{\gamma \bar{w}} < 0$, le meilleur taux de convergence est atteint lorsque c est proche de 1.

Introduisons, à présent, la notion de c-approximation forte afin d'obtenir la convergence superlinéaire :

Définition 6.2.1. Soit $c \in (0,1)$ et soient $w^k > 0$, $\theta_k \geq \bar{\theta}$ et $x^k \in S$. Une fonction convexe $\varphi(\cdot, w^k, \theta_k)$ est une c-approximation forte de $F(\cdot, w^k, \theta_k)$ au point x^k si $\varphi(x, w^k, \theta_k) \leq F(x, w^k, \theta_k)$ pour tout $x \in S$ et si

$$F(x^{k+1}, w^k, \theta_k) - \varphi(x^{k+1}, w^k, \theta_k) \leq \frac{1-c}{\alpha_k} \left\| x^{k+1} - x^k \right\|^2, \tag{6.26}$$

où x^{k+1} est la solution du problème $(AP_{w^k, \theta_k, \alpha_k})$.

Démontrons, avant tout, une proposition qui nous sera utile afin de démontrer la convergence superlinéaire de la suite $\{\theta_k\}$.

Proposition 6.2.2. *Soit $c \in (0,1)$. Une c-approximation forte de $F(\cdot, w^k, \theta_k)$ au point x^k est aussi une c-approximation de $F(\cdot, w^k, \theta_k)$ au point x^k.*

Démonstration.
Par (6.8), nous avons que

$$\alpha_k^{-1}(x^k - x^{k+1}) \in \partial \left[\varphi(\cdot, w^k, \theta_k) + \psi_S \right] (x^{k+1}).$$

Par la définition de sous-différentiel, nous obtenons

$$\varphi(x^k, w^k, \theta_k) - \varphi(x^{k+1}, w^k, \theta_k) \geq \alpha_k^{-1} \left\| x^{k+1} - x^k \right\|^2. \tag{6.27}$$

Or, par hypothèse, $\varphi(x^k, w^k, \theta_k) \leq F(x^k, w^k, \theta_k)$. De plus, comme $\theta_k = \max_i \frac{f_i(x^k)}{g_i(x^k)}$, nous avons que le maximum est atteint pour un certain i^* c'est-à-dire

$$\theta_k = \frac{f_{i^*}(x^k)}{g_{i^*}(x^k)}.$$

Donc

$$f_i(x^k) - \theta_k g_i(x^k) \begin{cases} \leq 0, & \text{si} \quad i \neq i^*, \\ = 0, & \text{si} \quad i = i^*. \end{cases}$$

Par conséquent,

$$F(x^k, w^k, \theta_k) = \max_{1 \leq i \leq p} \left\{ \frac{f_i(x^k) - \theta_k g_i(x^k)}{w_i^k} \right\} = 0. \tag{6.28}$$

En utilisant (6.26) dans (6.27), nous avons

$$\begin{aligned}
&\underbrace{\varphi(x^k, w^k, \theta_k)}_{\leq 0} - \varphi(x^{k+1}, w^k, \theta_k) \geq \alpha_k^{-1} \left\| x^{k+1} - x^k \right\|^2 \\
\Leftrightarrow\ & -\varphi(x^{k+1}, w^k, \theta_k) \geq \underbrace{\alpha_k^{-1} \left\| x^{k+1} - x^k \right\|^2}_{(\star\,)}.
\end{aligned}$$

Or, $(\star\,) \geq \frac{1}{1-c}\left(F(x^{k+1}, w^k, \theta_k) - \varphi(x^{k+1}, w^k, \theta_k)\right)$.
D'où,

$$\begin{aligned}
&-\varphi(x^{k+1}, w^k, \theta_k) \geq \frac{1}{1-c}\left(F(x^{k+1}, w^k, \theta_k) - \varphi(x^{k+1}, w^k, \theta_k)\right) \\
\Leftrightarrow\ & \frac{c}{1-c}\varphi(x^{k+1}, w^k, \theta_k) \geq \frac{1}{1-c}F(x^{k+1}, w^k, \theta_k) \\
\Leftrightarrow\ & \varphi(x^{k+1}, w^k, \theta_k) \geq \frac{1}{c}F(x^{k+1}, w^k, \theta_k) \\
\Leftrightarrow\ & (6.1).
\end{aligned}$$

□

Théorème 6.2.4. *Supposons que l'ensemble S^* des solutions du problème (P) est non vide, que la fonction $F(\cdot, \bar{\theta})$ satisfait la condition suivante*

$$\exists\, \delta > 0,\ \kappa > 0 \text{ t.q. } F(x, \bar{\theta}) \geq \kappa \ dist(x, S^*)^2 \qquad \forall\, x \in B(S^*, \delta) \cap S,$$

et que $\tau = \inf_{x \in S^*} \min_i g_i(x) > 0$. Supposons également que, à chaque itération, $\varphi(\cdot, w^k, \theta_k)$ est une c-approximation forte de $F(\cdot, w^k, \theta_k)$ avec $c > \frac{1}{2}$ et que la suite $\{x^k\}$ converge vers une solution de (P). Alors la suite $\{\theta_k\}$ converge superlinéairement vers $\bar{\theta}$ si α_k tend vers $+\infty$ lorsque $k \to \infty$ si, à chaque itération, w_i^k est choisi égal à $\beta g_i(x^k)$ avec $\beta > 0$.

Démonstration.
Puisque $F(x^{k+1}, w^k, \theta_k) = \max_{1 \leq i \leq p} \left\{ \dfrac{f_i(x^{k+1}) - \theta_k g_i(x^{k+1})}{w_i^k} \right\}$ et puisque, pour $i \in J(x^{k+1})$,

$\theta_{k+1} = \frac{f_i(x^{k+1})}{g_i(x^{k+1})}$, nous avons

$$\begin{aligned} F(x^{k+1}, w^k, \theta_k) &\geq \frac{f_i(x^{k+1}) - \theta_k g_i(x^{k+1})}{w_i^k} \\ &= (\theta_{k+1} - \theta_k)\frac{g_i(x^{k+1})}{w_i^k} \\ &= (\theta_{k+1} - \bar{\theta} + \bar{\theta} - \theta_k)\frac{g_i(x^{k+1})}{w_i^k} \\ &= \underbrace{(\theta_{k+1} - \bar{\theta})}_{\geq 0}\frac{g_i(x^{k+1})}{w_i^k} - \underbrace{(\theta_k - \bar{\theta})}_{\geq 0}\frac{g_i(x^{k+1})}{w_i^k} \quad \text{car} \quad \theta_k \geq \bar{\theta} \quad \forall\, k \\ &\geq (\theta_{k+1} - \bar{\theta})\min_i \frac{g_i(x^{k+1})}{w_i^k} - (\theta_k - \bar{\theta})\max_i \frac{g_i(x^{k+1})}{w_i^k}. \end{aligned} \tag{6.29}$$

Soit $\tilde{x}^k \in S^*$ tel que $\left\|x^k - \tilde{x}^k\right\|^2 = d(x^k, S^*)$. Puisque, par construction de l'algorithme, x^{k+1} est solution du problème $(AP_{w^k,\theta_k,\alpha_k})$, nous avons que

$$\varphi(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|\tilde{x}^k - x^k\right\|^2 \geq \varphi(x^{k+1}, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2. \tag{6.30}$$

Puisque $\varphi(\cdot, w^k, \theta_k)$ est une c-approximation forte c'est-à-dire puisque $\varphi(\tilde{x}^k, w^k, \theta_k) \leq F(\tilde{x}{\cdot}w^k, \theta_k)$, (6.30) s'écrit

$$F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|\tilde{x}^k - x^k\right\|^2 \geq \underbrace{\varphi(x^{k+1}, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|x^{k+1} - x^k\right\|^2}_{(\star)}.$$

En utilisant (6.26), cette dernière inégalité peut s'écrire

$$F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|\tilde{x}^k - x^k\right\|^2 \geq F(x^{k+1}, w^k, \theta_k) + \frac{2c-1}{2\alpha_k}\left\|x^{k+1} - x^k\right\|^2.$$

Or, comme $c > \frac{1}{2}$, nous avons

$$\frac{2c-1}{2\alpha_k} > 0.$$

Par conséquent,

$$F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|\tilde{x}^k - x^k\right\|^2 \geq F(x^{k+1}, w^k, \theta_k). \tag{6.31}$$

En combinant (6.29) et (6.31), nous obtenons

$$F(\tilde{x}^k, w^k, \theta_k) + (2\alpha_k)^{-1}\left\|\tilde{x}^k - x^k\right\|^2 \geq (\theta_{k+1} - \bar{\theta})\min_i \frac{g_i(x^{k+1})}{w_i^k} - (\theta_k - \bar{\theta})\max_i \frac{g_i(x^{k+1})}{w_i^k}.$$

En utilisant les expressions (6.21) et (6.23), nous avons

$$\underbrace{F(\tilde{x}^k, w^k, \theta_k)}_{\leq (\bar{\theta} - \theta_k)\frac{g_i(\tilde{x}^k)}{w_i^k}} + (2\alpha_k)^{-1}\underbrace{\left\|\tilde{x}^k - x^k\right\|^2}_{\leq \frac{\gamma}{\kappa}(\theta_k - \bar{\theta})} \geq (\theta_{k+1} - \bar{\theta})\min_i \frac{g_i(x^{k+1})}{w_i^k} - (\theta_k - \bar{\theta})\max_i \frac{g_i(x^{k+1})}{w_i^k}$$

$$\Leftrightarrow \left[\max_i \frac{g_i(x^{k+1})}{w_i^k} - \min_i \frac{g_i(\tilde{x}^k)}{w_i^k} + \frac{\gamma}{2\alpha_k \kappa}\right](\theta_k - \bar{\theta}) \geq (\theta_{k+1} - \bar{\theta})\min_i \frac{g_i(x^{k+1})}{w_i^k}.$$

Par (6.21), nous savons que les suites $\{x^k\}$ et $\{\tilde{x}^k\}$ convergent vers la même limite. En combinant ceci avec le choix de w : $w_i^k = \beta g_i(x^k)$ pour tout k et $1 \leq i \leq p$, nous avons

$$\max_i \frac{g_i(x^{k+1})}{w_i^k} \to \frac{1}{\beta}, \quad \min_i \frac{g_i(\tilde{x}^k)}{w_i^k} \to \frac{1}{\beta} \quad \text{et} \quad \min_i \frac{g_i(x^{k+1})}{w_i^k} \to \frac{1}{\beta}$$

lorsque $k \to \infty$ et donc

$$\left[\underbrace{\max_i \frac{g_i(x^{k+1})}{w_i^k}}_{\to\frac{1}{\beta}} - \underbrace{\min_i \frac{g_i(\tilde{x}^k)}{w_i^k}}_{\to\frac{1}{\beta}} + \frac{\gamma}{2\alpha_k\kappa}\right](\theta_k - \bar{\theta}) \geq (\theta_{k+1} - \bar{\theta}) \underbrace{\min_i \frac{g_i(x^{k+1})}{w_i^k}}_{\to\frac{1}{\beta}}$$

$$\Leftrightarrow \frac{\gamma}{2\alpha_k\kappa}(\theta_k - \bar{\theta}) \geq (\theta_{k+1} - \bar{\theta})\frac{1}{\beta}.$$

Par conséquent, comme $\alpha_k \to +\infty$ lorsque $k \to \infty$, nous obtenons

$$\frac{(\theta_{k+1} - \bar{\theta})}{(\theta_k - \bar{\theta})} \to 0.$$

□

Notons qu'un résultat semblable a été démontré dans le théorème 4.2.2. En effet, dans ce dernier, nous avions démontré la convergence superlinéaire de la suite $\{\theta_k\}$ mais aucun terme de régularisation n'y apparaissait c'est-à-dire que $\alpha_k = +\infty$ et toutes les composantes de w_i^k de w étaient égales à $g_i(x^k)$.

6.3 Construction des c-approximations

Afin de pouvoir implémenter la méthode décrite dans la section précédente, nous devons, à présent, préciser comment construire une c-approximation de la fonction $F(\cdot, w^k, \theta_k)$ au point x^k afin que le problème $(AP_{w^k,\theta_k,\alpha_k})$ soit plus facile à résoudre que le problème $(P_{w^k,\theta_k,\alpha_k})$.

Afin de faciliter l'écriture, notons la fonction $F(\cdot, w^k, \theta_k)$ par F^k et la c-approximation (forte) de F^k au point x^k par φ^k.

Lorsque l'ensemble S est décrit par des égalités et/ou des inégalités linéaires et lorsque φ^k est une fonction linéaire convexe par morceaux, le problème $(AP_{w^k,\theta_k,\alpha_k})$

est équivalent à un problème quadratique convexe et peut alors être résolu par des méthodes numériques efficaces connues. En effet, soit $\varphi^k(x) = \max\limits_{1\leq q\leq m} \{\langle a_q, x\rangle + b_q\}$ où $a_q \in \mathbb{R}^n$ et $b_q \in \mathbb{R}$ pour tout $1 \leq q \leq m$. Alors, le problème $(AP_{w^k,\theta_k,\alpha_k})$ est équivalent au problème suivant :

$$\begin{cases} \min \; v + \frac{1}{2\alpha_k}\left\|x - x^k\right\|^2 \\ \text{s.c} \;\; v \geq \langle a_q, x\rangle + b_q, \quad q = 1, \cdots, m \\ \qquad x \in S \end{cases}$$

La construction de la c-approximation φ^k est basée sur l'hypothèse classique de la programmation convexe non différentielle suivante :

En tout point $y \in S$, un sous-gradient de F^k en y est disponible. Ce sous-gradient est noté $s(y)$.

Afin d'obtenir une fonction convexe linéaire par morceaux pour φ^k, nous allons la construire morceau par morceau à l'aide de modèles successifs

$$\varphi_j^k, \qquad j = 1, 2, \cdots$$

jusqu'à ce que $\varphi_{j_k}^k$ soit, si possible, une c-approximation (forte) de la fonction F^k au point x^k pour $j_k \geq 1$.

Pour $j = 1, 2, \cdots$, notons par y_j^k l'unique solution du problème

$$(P_j^k) \qquad \min_{y\in S}\left\{\varphi_j^k(y) + \frac{1}{2\alpha_k}\left\|y - x^k\right\|^2\right\}$$

et prenons $\varphi^k = \varphi_{j_k}^k$ et $x^{k+1} = y_{j_k}^k$.

Définissons également les fonctions affines l_j^k, $j = 1, 2, \cdots$ par

$$l_j^k(y) = \varphi_j^k(y_j^k) + \langle \gamma_j^k, y - y_j^k\rangle \qquad \forall\, y \in \mathbb{R}^n,$$

où $\gamma_j^k = \alpha_k^{-1}(x^k - y_j^k)$.

Remarquons que

$$\min_{y\in S}\left\{\varphi_j^k(y) + \frac{1}{2\alpha_k}\left\|y - x^k\right\|^2\right\} \Leftrightarrow \min\left\{\varphi_j^k(y) + \frac{1}{2\alpha_k}\left\|y - x^k\right\|^2 + \psi_S(y)\right\}.$$

D'où, par l'optimalité de y_j^k, nous avons

$$\begin{aligned} & 0 \in \partial \left[\varphi_j^k + \frac{1}{2\alpha_k} \left\| y_j^k - x^k \right\|^2 + \psi_S \right] (y_j^k) \\ \Leftrightarrow\ & 0 \in \frac{1}{\alpha_k}(y_j^k - x^k) + \partial \left[\varphi_j^k + \psi_S \right] (y_j^k) \\ \Leftrightarrow\ & \frac{1}{\alpha_k}(x^k - y_j^k) \in \partial \left[\varphi_j^k + \psi_S \right] (y_j^k) \\ \Leftrightarrow\ & \gamma_j^k \in \partial \left[\varphi_j^k + \psi_S \right] (y_j^k). \end{aligned}$$

Alors, nous pouvons observer que

$$l_j^k(y_j^k) = \varphi_j^k(y_j^k) + \underbrace{\left\langle \gamma_j^k, y_j^k - y_j^k \right\rangle}_{=0} = \varphi_j^k(y_j^k)$$

et, par la définition de sous-différentiel,

$$\gamma_j^k \in \partial \left[\varphi_j^k + \psi_S \right] (y_j^k) \Leftrightarrow \varphi_j^k(y) \geq \varphi_j^k(y_j^k) + \left\langle \gamma_j^k, y - y_j^k \right\rangle = l_j^k(y) \qquad \forall\, y \in S.$$

Afin d'obtenir une c-approximation (forte) $\varphi_{j_k}^k$ de F^k au point x^k, imposons trois conditions sur les modèles successifs φ_j^k, $j = 1, 2, \dots$. Supposons donc que les fonctions approximantes convexes φ_j^k satisfont :

(C_1) $\varphi_j^k \leq F^k$ sur S pour $j = 1, 2, \cdots$,

(C_2) $\varphi_{j+1}^k \geq F^k(y_j^k) + \left\langle s(y_j^k), \cdot - y_j^k \right\rangle$ sur S pour $j = 1, 2, \cdots$,

(C_3) $\varphi_{j+1}^k \geq l_j^k$ sur S pour $j = 1, 2, \cdots$

où $s(y_j^k)$ désigne le sous-gradient de F^k disponible au point y_j^k.

Plusieurs modèles remplissent ces conditions. Par exemple, pour le premier modèle φ_1^k, nous pouvons prendre la fonction linéaire

$$\varphi_1^k(y) = F^k(x^k) + \left\langle s(x^k), y - x^k \right\rangle \qquad \forall\, y \in \mathbb{R}^n.$$

Vérifions la condition (C_1) : puisque $s(x^k) \in \partial F^k(x^k)$, par la définition de sous-différentiel, nous avons

$$F^k(y) \geq F^k(x^k) + \left\langle s(x^k), y - x^k \right\rangle = \varphi_1^k(y) \qquad \Rightarrow \qquad (C_1) \text{ est satisfaite.}$$

Les modèles suivants φ_j^k, $j = 2, \cdots$ peuvent être construits de différentes façons. Citons deux exemples :

1. Pour $j = 1, 2, \cdots$

$$\varphi_{j+1}^k(y) = \max\left\{l_j^k(y), F^k(y_j^k) + \langle s(y_j^k), y - y_j^k\rangle\right\} \qquad \forall\, y \in \mathbb{R}^n. \tag{6.32}$$

Nous pouvons facilement constater que les conditions (C_2) et (C_3) sont satisfaites ainsi que (C_1) car chaque morceau linéaire de ces fonctions est inférieur à F^k.

2. Pour $j = 1, 2, \cdots$

$$\varphi_{j+1}^k(y) = \max_{0 \le q \le j}\left\{F^k(y_q^k) + \langle s(y_q^k), y - y_q^k\rangle\right\} \qquad \forall\, y \in \mathbb{R}^n \tag{6.33}$$

où $y_0^k = x^k$.

Puisque $s(y_q^k) \in \partial F^k(y_q^k)$ pour $q = 0, \ldots, j$, la condition (C_1) est satisfaite. Par la définition de φ_{j+1}^k, (C_2) est vérifiée et puisque $\varphi_{j+1}^k \ge \varphi_j^k \ge l_j^k$, (C_3) est satisfaite.

Si nous comparons (6.32) et (6.33), nous constatons que l_j^k joue le même rôle que les j fonctions linéaires $F^k(y_q^k) + \langle s(y_q^k), y - y_q^k\rangle$, $q = 0, \cdots, j-1$. D'où, la fonction l_j^k est appelée *fonction affine aggrégée.*

Citons, à présent, l'algorithme permettant de construire une c-approximation de F^k au point x^k.

Algorithme de la c-approximation forte

Soit $x^k \in S$ et $c \in (\frac{1}{2}, 1)$. Soit $j = 1$.

Pas 1 : Choisir φ_j^k une fonction convexe linéaire par morceaux qui satisfait les conditions (C_1), (C_2) et (C_3) et résoudre le problème (P_j^k) pour obtenir y_j^k.

Pas 2 : Si $F^k(y_j^k) - \varphi_j^k(y_j^k) \le (1-c)\alpha_k^{-1}\left\|y_j^k - x^k\right\|^2$, alors prendre $x^{k+1} = y_j^k$, $j_k = j$ et STOP ; la fonction $\varphi_{j_k}^k$ est une c-approximation forte de F^k au point x^k et x^{k+1} le prochain itéré.

Pas 3 : Prendre $j := j + 1$ et retourner au pas 1.

Nous pouvons également obtenir une c-approximation en remplaçant, dans le pas 2, l'inégalité par

$$\varphi_j^k(y_j^k) \ge c^{-1}F^k(y_j^k).$$

Puisque une c-approximation forte est également une c-approximation, il est évident que si nous obtenons une c-approximation forte après un nombre fini d'itérations, nous obtiendrons également une c-approximation après un nombre fini d'itérations. Par conséquent, nous allons uniquement considérer une c-approximation forte dans le théorème qui suit. De plus, le fait que φ_j^k satisfait les conditions (C_1) à (C_3) signifie que φ_j^k satisfait (C_1) et, si $j \geq 2$, φ_j^k satisfait (C_2) et (C_3) avec $j+1$ remplacé par j. Le théorème suivant permet de prouver que si x^k n'est pas un minimum de F^k et si les modèles φ_j^k, $j = 1, \cdots$ satisfont les hypothèses (C_1) à (C_3), alors il existe un indice $j_k \in \mathbb{N}_0$ tel que $\varphi_{j_k}^k$ est une c-approximation forte de F^k au point x^k, i.e. , la procédure s'arrête au second pas après un nombre fini d'itérations.

Théorème 6.3.1. *Supposons que les modèles φ_j^k, $j = 1, 2, \cdots$ satisfont les conditions (C_1) à (C_3), et soit, pour tout j, y_j^k l'unique solution du problème (P_j^k). Soit, également, $\bar{x}^k$ l'unique solution du problème $(P_{w^k, \theta_k, \alpha_k})$. Alors*

1. *$F^k(y_j^k) - \varphi_j^k(y_j^k) \to 0$ et $y_j^k \to \bar{x}^k$ quand $j \to +\infty$.*
2. *Si $x^k \neq \bar{x}^k$, alors l'algorithme de la c-approximation forte s'arrête après un nombre fini d'itérations j_k avec $\varphi_{j_k}^k$ une c-approximation forte de F^k au point x^k et avec $x^{k+1} = y_{j_k}^k$.*

En insérant l'algorithme de la c-approximation forte dans le pas 1 de l'algorithme de point proximal inexact, nous obtenons l'algorithme suivant :

Algorithme de faisceaux

Choisir $x^0 \in X$, $w^0 > 0$, $\alpha_0 > 0$, $c \in (\frac{1}{2}, 1)$ et soit $\theta_0 = \theta(x^0)$, $y_0^0 = x^0$ et $k = 0$, $j = 1$.

Pas 1 : Choisir une fonction convexe linéaire par morceaux φ_j^k satisfaisant les hypothèses (C_1) à (C_3) et résoudre

$$(P_j^k) \qquad \min_{y \in S} \left\{ \varphi_j^k(y) + \frac{1}{2\alpha_k} \left\| y - x^k \right\|^2 \right\},$$

pour obtenir l'unique solution optimale y_j^k.

Pas 2 : Si $F^k(y_j^k) - \varphi_j^k(y_j^k) \leq (1-c)\alpha_k^{-1} \left\| y_j^k - x^k \right\|^2$, alors prendre $x^{k+1} = y_j^k$, $y_0^{k+1} = x^{k+1}$, $\theta_{k+1} = \theta(x^{k+1})$, choisir $w^{k+1} > 0$, $\alpha_{k+1} > 0$, prendre $k := k+1$ et $j = 0$.

Pas 3 : Prendre $j := j + 1$ et retourner au pas 2.

Un autre algorithme faisceaux peut être obtenu en remplaçant, dans le pas 2, l'inégalité correspondant à une c-approximation forte c'est-à-dire $F(y_j^k) - \varphi_j^k(y_j^k) \leq (1-c)\alpha_k^{-1} \left\| y_j^k - x^k \right\|^2$ par l'inégalité correspondant à une c-approximation suivante :

$$\varphi_j^k(y_j^k) \geq c^{-1} F^k(y_j^k).$$

Chapitre 7

Dualité en programmation fractionnelle convexe

7.1 Introduction

Considérons le problème de programmation fractionnelle généralisée suivant :

$$(P) \qquad \inf_{x \in S} \left(\max_{i \in I} \frac{f_i(x)}{g_i(x)} \right)$$

où l'ensemble $S \subseteq \mathbb{R}^n$ est compact, les fonctions $f_i, g_i : \mathcal{T} \to \mathbb{R}$, $i \in I = \{1, \cdots, p\}$, $p \geq 1$, sont continues, $\mathcal{T}$ est un ouvert contenant S, $g_i(x) > 0$ pour tout $x \in S$ et $i \in I$.

Remarquons que le problème d'optimisation (P) possède une solution optimale puisque la fonction $x \to \max_{i \in I} f_i(x)/g_i(x)$ est finie et continue sur l'ensemble compact S.

Par la suite, nous serons amenés à introduire le problème dual associé au problème (P). Pour ce faire, supposons, à présent, que l'ensemble S est convexe, que les fonctions f_i, $i \in I$, sont convexes sur S et que les fonctions g_i, $i \in I$, sont concaves sur S.

Introduisons également les fonctions suivantes :

$$f(x)^T = (f_1(x), \cdots, f_p(x)), \qquad g(x)^T = (g_1(x), \cdots, g_p(x))$$

2. La fonction $y \to h(x,y)$ étant un rapport de fonctions linéaires en y, par la propriété 1.2.3.4, $h(x,y)$ est strictement quasiconvexe et strictement quasiconcave sur Σ.

3. Puisque $y \in \Sigma$, par un raisonnement semblable à celui du point 1, nous pouvons conclure que la fonction $x \to h(x,y)$ est quasiconvexe sur S.

Il nous reste à prouver l'égalité (7.1). La fonction $y \to h(x,y)$ étant quasiconvexe sur Σ, elle atteint son maximum en un des sommets du polyhèdre convexe c'est-à-dire $(1, 0, \cdots, 0), \ldots, (0, \cdots, 0, 1)$. Par conséquent,

$$\max_{i \in I} \frac{f_i(x)}{g_i(x)} = \max_{y \in \Sigma} \frac{y^T f(x)}{y^T g(x)}.$$

$\square$

7.2 Algorithme dual

Supposons que les hypothèses de la proposition 7.1.1 sont satisfaites. Puisque les ensembles S et Σ sont compacts et que, comme démontré dans la proposition précédente, la fonction $h(x,y)$ est quasiconvexe en x et quasiconcave en y, il suit, par le théorème du minimax, que

$$\min_{x \in S} \left\{ \max_{y \in \Sigma} h(x,y) \right\} = \max_{y \in \Sigma} \left\{ \min_{x \in S} h(x,y) \right\}.$$

Or, par définition de $h(x,y)$, nous avons

$$\max_{y \in \Sigma} \left\{ \min_{x \in S} h(x,y) \right\} \stackrel{\text{def}}{=} \max_{y \in \Sigma} \left\{ \min_{x \in S} \frac{y^T f(x)}{y^T g(x)} \right\} \tag{7.2}$$

et par (7.1), nous avons

$$\min_{x \in S} \left\{ \max_{y \in \Sigma} h(x,y) \right\} \stackrel{(7.1)}{=} \min_{x \in S} \left\{ \max_{i \in I} \frac{f_i(x)}{g_i(x)} \right\}. \tag{7.3}$$

Dès lors, en combinant (7.2) et (7.3), nous obtenons

$$\min_{x \in S} \left\{ \max_{i \in I} \frac{f_i(x)}{g_i(x)} \right\} = \max_{y \in \Sigma} \left\{ \min_{x \in S} \frac{y^T f(x)}{y^T g(x)} \right\} \tag{7.4}$$

Définissons, à présent, la fonction $c : \Sigma \to \mathbb{R}$ par

$$c(y) = \min_{x \in S} \frac{y^T f(x)}{y^T g(x)} \stackrel{\text{def}}{=} \min_{x \in S} h(x,y). \tag{7.5}$$

Notons que cette fonction c est continue sur Σ puisque la fonction h est continue sur l'ensemble compact S et est strictement quasiconcave puisqu'elle n'est autre qu'un infimum de fonctions strictement quasiconcaves. Par conséquent, par (7.4), le problème à résoudre est équivalent au problème d'optimisation strictement quasiconcave suivant :

$$(Q) \qquad \max_{y\in\Sigma} c(y).$$

Rappelons que, dans ce type de problème, un minimum local est également un minimum global.

Le problème (Q) correspond à un nouveau problème dual du problème (P). En effet, lorsque, dans un problème dual standard d'un problème de programmation fractionnelle généralisée, une partie de l'ensemble des contraintes est dualisée, le problème (Q) peut être considéré comme un problème dual partiel d'un problème (P) de programmation fractionnelle généralisée, puisque seuls les rapports sont dualisés.

La fonction c étant continue sur l'ensemble compact Σ, le problème dual (Q) possède une solution y^*. Ceci ne signifie pas que la solution x^* du problème d'optimisation associé à $c(y^*)$ est une solution du problème primal (P). Néanmoins, nous avons le résultat suivant :

Proposition 7.2.1. *Si x^* est une solution optimale du problème (P) et y^* une solution optimale du problème (Q), alors x^* est une solution optimale du problème d'optimisation associé à $c(y^*)$.*

Démonstration.
D'une part, puisque x^* est une solution optimale du problème (P) et par (7.1), nous avons

$$\begin{aligned}\theta(x^*) &= \max_{i\in I}\frac{f_i(x^*)}{g_i(x^*)}\\ &\overset{(7.1)}{=} \max_{y\in\Sigma}\frac{y^Tf(x^*)}{y^Tg(x^*)}.\end{aligned}$$

similaire.

1. Montrons d'abord que $c(y) > c(\bar{y}) \Rightarrow F_D(y, c(\bar{y})) > 0$.
Puisque $g(x) > 0$ pour tout $x \in S$, nous avons

$$c(\bar{y})g(x) < c(y)g(x).$$

Par conséquent,

$$\begin{aligned} F_D(y, c(\bar{y})) &= \min_{x\in S} \left\{ y^T \left[f(x) - c(\bar{y})g(x) \right] \right\} \\ &> \min_{x\in S} \left\{ y^T \left[f(x) - c(y)g(x) \right] \right\} \\ &= F_D(y, c(y)). \end{aligned} \tag{7.9}$$

Or, comme pour tout $x \in S$, nous avons que $c(y) \leq \frac{y^T f(x)}{y^T g(x)}$ et pour un certain $\hat{x} \in S$, $c(y) = \frac{y^T f(\hat{x})}{y^T g(\hat{x})}$ car S est compact, nous obtenons que

$$F_D(y, c(y)) = 0. \tag{7.10}$$

Finalement, par (7.9) et (7.10), nous obtenons

$$F_D(y, c(\bar{y})) > 0.$$

2. Montrons ensuite que $F_D(y, c(\bar{y})) > 0 \Rightarrow c(y) > (\bar{y})$.
D'une part, puisque $F_D(y, c(\bar{y})) > 0$, nous avons

$$y^T \left[f(x) - c(\bar{y})g(x) \right] > 0 \qquad \forall\, x \in S$$

c'est-à-dire

$$c(\bar{y}) < \frac{y^T f(x)}{y^T g(x)} \qquad \forall\, x \in S.$$

D'autre part, nous avons

$$c(y) = \frac{y^T f(\hat{x})}{y^T g(\hat{x})} \qquad \text{pour un certain } \hat{x} \in S.$$

Par conséquent,

$$c(y) > c(\bar{y}).$$

$\square$

Supposons que nous sommes à l'itération k où l'on connaît le point y^k. Selon le lemme précédent, il suffit d'augmenter la valeur de F_D afin de faire croître la valeur de c. Une stratégie possible est donnée par

$$F_D(y^{k+1}, c(y^k)) = \max_{y \in \Sigma} F_D(y, c(y^k)). \tag{7.11}$$

Notons que ce problème peut être considéré comme le problème paramétrique de paramètre $c(y^k)$ associé au problème (Q) :

$$(Q) \qquad \max_{y \in \Sigma} c(y) = \max_{y \in \Sigma} \left\{ \min_{x \in S} \left[\frac{y^T f(x)}{y^T g(x)} \right] \right\}$$

$$\downarrow$$

$$(Q_{c(y^k)}) \qquad \max_{y \in \Sigma} F_D(y, c(y^k)) = \max_{y \in \Sigma} \left\{ \min_{x \in S} \left[y^T (f(x) - c(y^k) g(x)) \right] \right\}.$$

Introduisons, à présent, l'algorithme dual :

Algorithme dual (AD)

Pas 0 **:** Soit $y^0 \in \Sigma$, calculer $c(y^0) = \min_{x \in S} \dfrac{y^{0T} f(x)}{y^{0T} g(x)}$, et soit $k = 1$.

Pas 1 **:** Déterminer $y^k = \arg\max_{y \in \Sigma} F_D(y, c(y^{k-1}))$.

Pas 2 **:** Si $F_D(y^k, c(y^{k-1})) = 0$, alors y^{k-1} est une solution optimale de valeur $c(y^{k-1})$ et STOP. Sinon aller au pas 3.

Pas 3 **:** Calculer $c(y^k)$. Prendre $k := k + 1$ et retourner au pas 1.

Examinons maintenant la procédure à suivre afin de pouvoir retrouver la solution optimale du problème (P). Pour ce faire, remarquons d'abord que

$$F_D(y^{k+1}, c(y^k)) = F(c(y^k)). \tag{7.12}$$

En effet,

$$\begin{aligned}
F_D(y^{k+1}, c(y^k)) &\overset{(7.11)}{=} \max_{y\in\Sigma} F_D(y, c(y^k)) \\
&\overset{\text{def}}{=} \max_{y\in\Sigma}\left\{\min_{x\in S}\left[y^T(f(x) - c(y^k)g(x))\right]\right\} \\
&= \min_{x\in S}\left\{\max_{y\in\Sigma}\left[y^T(f(x) - c(y^k)g(x))\right]\right\} \\
&\qquad \text{par le théorème min-max de Von Neumann} \\
&= \min_{x\in S}\left\{\max_{i\in I}\left[f_i(x) - c(y^k)g_i(x)\right]\right\} \\
&\qquad \text{car le maximum est atteint en un des sommets du polyhèdre } \Sigma \\
&\overset{\text{def}}{=} F(c(y^k)),
\end{aligned}$$

Regardons ce qu'il se passe lors du pas 2 de l'algorithme dual. A cette étape, nous avons que

$$F_D(y^k, c(y^{k-1})) = 0$$

et donc, par (7.12),

$$F(c(y^{k-1})) = 0.$$

D'où, par la propriété 3.3.4, les problèmes

$$(P) \quad \bar{\theta} = \inf_{x\in S}\left\{\max_{i\in I}\frac{f_i(x)}{g_i(x)}\right\}$$

et

$$(P_\theta) \quad F(\theta) = \inf_{x\in S}\left\{\max_{i\in I}\left[f_i(x) - \theta g_i(x)\right]\right\}$$

avec $\theta = c(y^{k-1})$ possèdent le même ensemble de solutions.
Par conséquent, toute solution du problème (P_θ) avec $\theta = c(y^{k-1})$ est également solution du problème primal (P).

Examinons, à présent, le pas 1 de l'algorithme dual. Lors de ce pas, à l'itération k, nous devons résoudre le problème d'optimisation, où $\mu = c(y^k)$, suivant :

$$(Q_\mu) \qquad \max_{y\in\Sigma}\left\{\min_{x\in S}\left[y^T(f(x) - \mu g(x))\right]\right\}.$$

La résolution de ce problème nécessite beaucoup de temps. D'un autre côté, quand nous appliquons l'algorithme de type Dinkelbach au problème primal (P), le problème (P_μ), qui semble plus simple, doit être résolu à chaque pas. Néanmoins, sous certaines hypothèses, il est possible de transformer une solution optimale x^{k+1} du problème $(P_{c(y^k)})$ en une solution optimale y^{k+1} de $(Q_{c(y^k)})$. Montrons cela dans le cas particulier où

$$S = \left\{x \in \mathbb{R}^n \middle| p_j(x) \leq 0,\ j = 1, \cdots, s\right\},$$

où $p_j : \mathbb{R}^n \rightarrow \mathbb{R}$, $j = 1, \cdots, s$ sont des fonctions convexes et différentiables. Notons que le problème $(P_{c(y^k)})$ est équivalent au problème suivant

$$(P_k) \qquad \begin{cases} \min\ t \\ \text{s.c. } f_i(x) - c(y^k)g_i(x) - t \leq 0, \quad i = 1, \cdots, p, \\ \qquad p_j(x) \leq 0, \quad j = 1, \cdots, s. \end{cases}$$

En plus des hypothèses traditionnelles sur les fonctions f_i et g_i, supposons également que ces fonctions sont différentiables.
Soient x^{k+1} et t_{k+1} une solution optimale du problème (P_k) et définissons l'ensemble des contraintes actives correspondant

$$I' = \left\{i \middle| f_i(x^{k+1}) - c(y^k)g_i(x^{k+1}) = t_{k+1}\right\} \quad \text{et} \quad J' = \left\{j \middle| p_j(x^{k+1}) = 0\right\}.$$

Avant de continuer cette analyse, écrivons la condition de Slater pour notre problème (P_k) :

$$\boxed{\exists \tilde{x} \in \mathbb{R}^n \text{ t.q. } p_j(\tilde{x}) < 0 \qquad \forall\, j = 1, \cdots, s.}$$

Si cette condition est satisfaite, les conditions de Karush-Kuhn-Tucker sont alors satisfaites. Par conséquent, il existe des scalaires non négatifs u_i, $i \in I'$, et v_j, $j \in J'$ tels que

$$(KKT) \qquad \begin{cases} \displaystyle\sum_{i \in I'} u_i \nabla q_i(x^{k+1}) + \sum_{j \in J'} v_j \nabla p_j(x^{k+1}) = 0, \\ \displaystyle\sum_{i \in I'} u_i = 1, \\ (u_{I'}, v_{J'}) \geq 0, \end{cases}$$

où $q_i(x) = f_i(x) - c(y^k)g_i(x)$.
Notons que, le couple (x^{k+1}, t_{k+1}) étant optimal, l'ensemble des contraintes actives

I' est non vide.

Le résultat suivant indique comment associer une solution optimale du problème $(Q_{c(y^k)})$ aux multiplicateurs de Lagrange $u_i,\ i \in I'$.

Proposition 7.2.2. *Si la qualification de contraintes de Slater est satisfaite sur S, alors une solution optimale $\bar{y}$ du problème $(Q_{c(y^k)})$ est donnée par*

$$\bar{y}_i = \begin{cases} 0, & si \quad i \notin I', \\ u_i, & si \quad i \in I'. \end{cases}$$

Démonstration.

Nous devons prouver que $\bar{y}$ est une solution du problème

$$(Q_{c(y^k)}) \qquad \max_{y \in \Sigma} F_D(y, c(y^k))$$

où

$$F_D(y, c(y^k)) = \min_{x \in S} \left\{ y^T \left[f(x) - c(y^k)(x) \right] \right\}.$$

Il faut donc montrer que

$$F_D(\bar{y}, c(y^k)) = \max_{y \in \Sigma} F_D(y, c(y^k)).$$

La condition de Slater étant satisfaite par hypothèse, les conditions de Karush-Kuhn-Tucker sont vérifiées. Par conséquent, $u_{I'} \in \Sigma$ et donc, par la définition de $\bar{y}_i$, $\bar{y} \in \Sigma$. De plus, nous obtenons

$$\begin{aligned}
\sum_{1 \le i \le p} \bar{y}_i q_i(x^{k+1}) &= \sum_{i \in I'} \bar{y}_i q_i(x^{k+1}) + \sum_{i \notin I'} \bar{y}_i q_i(x^{k+1}) \\
&= \sum_{i \in I'} u_i q_i(x^{k+1}) + 0 \qquad \text{par définition de } \bar{y}_i \\
&= \sum_{i \in I'} u_i t_{k+1} \qquad \text{par définition de } I' \\
&= t_{k+1} \underbrace{\sum_{i \in I'} u_i}_{\overset{(KKT)}{=} 1} \\
&= t_{k+1}. \qquad\qquad (*)
\end{aligned}$$

D'un autre côté, nous avons

$$
\begin{aligned}
t_{k+1} &= \min_{x\in S}\left\{\max_{1\le i\le p} q_i(x)\right\} \qquad t_{k+1} \text{ est une solution optimale du problème } (P_{c(y^k)})\\
&= \min_{x\in S}\left\{\max_{1\le i\le p}\left[f_i(x)-c(y^k)g_i(x)\right]\right\}\\
&= F(c(y^k))\\
&= \max_{y\in\Sigma} F_D(y,c(y^k)) \qquad \text{car par (7.12), } F(c(y^k)) = F_D(y^{k+1},c(y^k)). \qquad (**)
\end{aligned}
$$

En combinant $(*)$ et $(**)$, nous obtenons

$$\sum_{1\le i\le p} \bar{y}_i q_i(x^{k+1}) = \bar{y}\; q(x^{k+1}) = \max_{y\in\Sigma} F_D(y,c(y^k))$$

où $q(x) = (q_i(x))_{1\le i\le p} = f(x) - c(y^k)g(x)$. Puisque, par définition, $F_D(\bar{y},c(y^k)) = \min_{x\in S} \bar{y}^T q(x)$, il nous reste à montrer que

$$x^{k+1} \text{ est la solution optimale du problème } \min_{x\in S} \bar{y}^T q(x).$$

Ce problème étant convexe, il nous suffit de montrer que x^{k+1} satisfait les conditions de KKT associées c'est-à-dire

$$\sum_{1\le i\le p} \bar{y}_i \nabla q_i(x) + \sum_{j\in J'} v_j \nabla p_j(x^{k+1}) = 0 \tag{7.13}$$

avec $v_{J'} \ge 0$.

Or, par les conditions de KKT associées au problème (P_k), nous avons

$$
\begin{aligned}
&\sum_{i\in I'} u_i \nabla q_i(x^{k+1}) + \sum_{j\in J'} v_j \nabla p_j(x^{k+1}) = 0\\
\Leftrightarrow\ &\sum_{1\le i\le p} \bar{y}_i \nabla q_i(x^{k+1}) + \sum_{j\in J'} v_j \nabla p_j(x^{k+1}) = 0 \qquad \text{par définition de } \bar{y}\\
\Leftrightarrow\ &(7.13).
\end{aligned}
$$

Par conséquent, $\bar{y}$ est une solution du problème $(Q_{c(y^k)})$. □

7.2.1 Interprétation géométrique de l'algorithme dual

Définissons, d'abord, pour tout $y \in \Sigma$, la fonction $F_y : \mathbb{R} \to \mathbb{R}$ par

$$F_y(\mu) = \min_{x\in S}\left\{y^T\left[f(x) - \mu g(x)\right]\right\}.$$

Pour $\mu \in \mathbb{R}$ et $y \in \Sigma$, nous avons

$$\begin{aligned}
F_y(\mu) &= \min_{x \in S} \left\{ y^T \left[f(x) - \mu g(x) \right] \right\} \\
&\leq \min_{x \in S} \left\{ \max_{y \in \Sigma} \left\{ y^T \left[f(x) - \mu g(x) \right] \right\} \right\} \\
&= \min_{x \in S} \left\{ \max_{i \in I} \left[f_i(x) - \mu g_i(x) \right] \right\} \\
&\qquad \text{car le maximum est atteint en un des sommets du polyhèdre } \Sigma \\
&\overset{\text{def}}{=} F(\mu).
\end{aligned}$$

Ainsi, pour $\mu = c(y^k)$, nous avons $F_{y^{k+1}}(\mu) = F(\mu)$. De plus, calculer $c(y^{k+1})$ revient à trouver la racine de $F_{y^{k+1}}(\mu) = 0$. En effet, puisque par construction de $c(y^{k+1})$, nous avons

$$c(y^{k+1}) \leq \frac{y^{k+1T} f(x)}{y^{K+1T} g(x)}$$

pour tout $x \in S$ et pour un certain $\hat{x} \in S$, nous avons même l'égalité

$$c(y^{k+1}) = \frac{y^{k+1T} f(x)}{y^{K+1T} g(x)}.$$

Donc, nous avons, pour tout $x \in S$,

$$y^{k+1T} \left\{ f(x) - c(y^{k+1}) g(x) \right\} \geq 0$$

avec l'égalité pour $x = \hat{x}$. Par conséquent,

$$F_{y^{k+1}}(c(y^{k+1})) = \min_{x \in S} \left\{ y^{k+1T} \left[f(x) - c(y^{k+1}) g(x) \right] \right\} = 0.$$

En résumé, au point $c(y^k)$, nous calculons y^{k+1} et nous obtenons $F_{y^{k+1}}(c(y^k)) = F(c(y^k)) > 0$. Alors, $c(y^{k+1})$ est la racine de $F_{y^{k+1}}$.

Par conséquent,

$$\begin{aligned}
-td &= -ty^T g(\bar{x}) \\
&= -ty^T g(\bar{x}) + y^T f(\bar{x}) - y^T f(\bar{x}) + y^T \mu g(\bar{x}) - y^T \mu g(\bar{x}) \\
&= y^T \left(f(\bar{x}) - (\mu + t) g(\bar{x}) \right) - y^T \left(f(\bar{x}) - \mu g(\bar{x}) \right) \\
&\geq \min_{x \in S} \left\{ y^T \left(f(x) - (\mu + t) g(x) \right) \right\} - \min_{x \in S} \left\{ y^T \left(f(x) - \mu g(x) \right) \right\} \\
&= F_y(\mu + t) - F_y(\mu).
\end{aligned}$$

Par cette dernière inégalité, nous pouvons conclure que d est un sous-gradient de la fonction $-F_y$ puisque $d \in \partial(-F_y(\mu))$ si

$$\begin{aligned}
-F_y(\mu + t) &\geq -F_y(\mu) + \langle d, \mu + t - \mu \rangle \\
F_y(\mu + t) &\leq F_y(\mu) - td.
\end{aligned}$$

En procédant de la même manière pour la borne inférieure de l'intervalle (7.15), nous avons que $\inf\limits_{x \in S_y(\mu)} \left\{ y^T g(x) \right\}$ est également un sous-gradient. Par conséquent, la première inclusion est vérifiée.

2. $\boxed{\subseteq}$ Considérons une suite x_n, $n \geq 1$ avec $x_n \in S_y(\mu + \frac{1}{n})$. L'ensemble S étant compact, la suite $\{x_n, n \geq 1\}$ possède une valeur d'adhérence $x_\infty \in S$ et il existe une sous-suite $\{x_{n_l} : l \geq 1\}$ telle que

$$\lim_{l \uparrow \infty} x_{n_l} = x_\infty.$$

Les fonctions $\mu \mapsto F_y(\mu), x \mapsto y^T f(x)$ et $x \mapsto y^T g(x)$ étant continues, nous obtenons que

$$\begin{aligned}
F_y(\mu) &= \lim_{l \uparrow \infty} F_y \left(\mu + \frac{1}{n_l} \right) \\
&\stackrel{\text{def}}{=} \lim_{l \uparrow \infty} \left(y^T f(x_{n_l}) - \left(\mu + \frac{1}{n_l} \right) y^T g(x_{n_l}) \right) \\
&= y^T f(x_\infty) - \mu y^T g(x_\infty).
\end{aligned}$$

Dès lors, $x_\infty \in S_y(\mu)$.

La fonction $x \mapsto y^T g(x)$ étant continue

$$\forall\, \delta > 0, \quad \exists\, l_\delta \geq 1 \text{ t.q. } \forall\, l \geq l_\delta \quad y^T g(x_{n_l}) \leq y^T g(x_\infty) + \delta \leq \sup_{x \in S_y(\mu)} \left\{ y^T g(x) \right\} + \delta. \tag{7.16}$$

D'où, pour tout $d \in \partial(-F_y)(\mu)$ et $l \geq l_\delta$, nous avons

$$
\begin{aligned}
-\frac{1}{n_l}d &\geq F_y\left(\mu + \frac{1}{n_l}\right) - F_y(\mu) \quad \text{par définition de sous-gradient.} \\
&\geq y^T\left(f(x_{n_l}) - \left(\mu + \frac{1}{n_l}\right)g(x_{n_l})\right) - y^T\left[f(x_{n_l}) - \mu g(x_{n_l})\right] \quad \text{par définition de } F_y(\cdot). \\
&= -\frac{1}{n_l}y^T g(x_{n_l}) \\
&\overset{(7.16)}{\geq} -\frac{1}{n_l}\left(\sup_{x \in S_y(\mu)}\left\{y^T g(x)\right\} + \delta\right).
\end{aligned}
$$

En multipliant les deux membres de cette dernière inégalité par $-n_l$, nous obtenons

$$d \leq \sup_{x \in S_y(\mu)}\left\{y^T g(x)\right\} + \delta.$$

Puisque $\delta > 0$ est arbitraire, nous avons

$$d \leq \sup_{x \in S_y(\mu)}\left\{y^T g(x)\right\}. \tag{7.17}$$

En considérant une suite $x_n \in S_y\left(\mu - \frac{1}{n}\right)$ et en utilisant un raisonnement semblable à celui-ci, nous pouvons montrer que

$$d \geq \inf_{x \in S_y(\mu)}\left\{y^T g(x)\right\}. \tag{7.18}$$

Par conséquent, en combinant (7.17) et (7.18), nous obtenons, pour $d \in \partial(-F_y)(\mu)$, que

$$d \in \left[\inf_{x \in S_y(\mu)}\left\{y^T g(x)\right\}, \sup_{x \in S_y(\mu)}\left\{y^T g(x)\right\}\right].$$

□

Soient

$$
\begin{aligned}
\delta_{k+1} &:= \min\left\{y^{k+1T}g(x) : x \text{ est solution de } \min_{x \in S}\frac{y^{k+1T}f(x)}{y^{k+1T}g(x)}\right\} \\
&= \min\left\{y^{k+1T}g(x) : x \in S_{y^{k+1}}(c(y^{k+1}))\right\}
\end{aligned}
$$

et

$$\Delta_k(y) := \max\left\{y^T g(x) : x \in S_y(c(y^k))\right\}.$$

Il suit du lemme 7.2.2 que

$$\Delta_k(y) \in \partial(-F_y)(c(y^k)) \quad \text{et} \quad \delta_{k+1} \in \partial(-F_{y^{k+1}})(c(y^{k+1})).$$

Afin de démontrer la convergence de l'algorithme dual, nous aurons également besoin du résultat suivant :

Théorème 7.2.1. *La suite y_k, $0 \leq k < k^*$, ne contient pas de solution optimale du problème (Q) et les valeurs $c(y^k)$, $0 \leq k < k^*$, sont croissantes. De plus, si k^* est fini, alors $c(y^{k^*}) = \mu^*$ pendant que, pour $k^* = +\infty$, toute valeur d'adhérence de la suite y^k, $k \geq 0$, est une solution optimale de (Q). Finalement, si $k^* = +\infty$ et y^* est une solution optimale de (Q), alors l'inégalité*

$$0 \leq \mu^* - c(y^{k+1}) \leq \left(1 - \frac{\Delta_k(y^*)}{\delta_{k+1}}\right)(\mu^* - c(y^k)) \tag{7.19}$$

est vérifiée pour tout $k \geq 0$.

Démonstration.
- Soit y_k un point non optimal. Par conséquent, par le lemme 7.2.1, nous avons

$$F_D(y^{k+1}, c(y^k)) > 0.$$

Par le lemme 7.2.1, nous obtenons également

$$c(y^{k+1}) > c(y^k).$$

Alors, nous pouvons déduire que la suite $c(y^k)$ est croissante.
- Observons que, pour k^* fini,

$$F_D(y^{k^*}, c(y^{k^*-1})) = 0.$$

Dès lors, en appliquant le lemme 7.2.1, y^k est une solution de (Q). Par conséquent, par (7.4), nous avons

$$c(y^{k^*}) = \mu^*.$$

- Notons que $c(y^k)$, $k \geq 0$, est croissante pour $k^* = +\infty$, et puisque $c(y^k) \leq \mu^* < +\infty$ pour tout $k \geq 0$, nous aurons donc que $\lim_{k \uparrow \infty} c(y^k)$ existe et est finie. De plus, par le lemme 7.2.2 et la définition de sous-gradient, nous obtenons, pour toute solution optimale y^* de (Q), que

$$F_{y^*}(c(y^*)) - F_{y^*}(c(y^k)) \leq -(c(y^*) - c(y^k))\Delta_k(y^*). \tag{7.20}$$

Théorème 7.2.2. *Soit $\{y^k\}$ la suite générée par l'algorithme dual. Alors toute valeur d'adhérence de la suite $\{y^k\}$ est une solution optimale du problème (Q). De plus, la suite $\{c(y^k)\}$ est croissante et converge linéairement vers $\bar{\theta}$. Finalement, si, pour toute solution optimale y^* de (Q), le problème $\min_{x \in S} \dfrac{y^{*T} f(x)}{y^{*T} g(x)}$ possède une seule solution, alors la convergence de la suite $\{c(y^k)\}$ est superlinéaire.*

Démonstration.
Les deux premiers points de ce théorème, à savoir que toute valeur d'adhérence de la suite $\{y^k\}$ est une solution optimale du problème (Q) et que la suite $\{c(y^k)\}$ est croissante, ont été démontrés auparavant. Il nous reste donc à montrer la convergence linéaire et superlinéaire de la suite $\{c(y^k)\}$:
- La convergence linéaire provient de (7.19). Afin d'avoir un taux linéaire de convergence, il suffit d'observer le comportement de $1 - \frac{\Delta_k(y^*)}{\delta_{k+1}}$ lorsque $k \to \infty$ pour une solution optimale arbitraire de (Q). Soit $\delta_\infty := \limsup_{k \uparrow \infty} \delta_{k+1}$. Par la définition de limite supérieure, il existe une sous-suite $\mathcal{K} \in \mathbb{N}$ telle que

$$\delta_\infty = \lim_{k \in \mathcal{K},\, k \uparrow \infty} \delta_{k+1}. \tag{7.24}$$

De plus, si nous considérons la suite $\{y^{k+1} : k \in \mathcal{K}\} \subseteq \Sigma$, par la compacité de Σ, nous pouvons trouver une sous-suite $\mathcal{K}_l \subseteq \mathcal{K}$ qui satisfait

$$\lim_{k \in \mathcal{K}_l,\, k \uparrow \infty} y^{k+1} = y^* \tag{7.25}$$

où y^* est une valeur d'adhérence. Par le théorème 7.2.1, y^* est une solution optimale de (Q).
Considérons, à présent, la suite $1 - \frac{\Delta_k(y^*)}{\delta_{k+1}}$ pour ce point y^*. Notons que la fonction $y \mapsto \partial(-F_y)(c(y))$ est semi-continue supérieurement. Par conséquent, comme $\delta_{k+1} \in \partial(-F_{y^{k+1}})(c(y^{k+1}))$, nous obtenons, par (7.24) et par (7.25), que

$$\delta_\infty \in \partial(-F_{y^*})(c(y^*)). \tag{7.26}$$

D'un autre côté, par le lemme 7.2.2, nous avons

$$\Delta_k(y^*) \in \partial(-F_{y^*})(c(y^k)).$$

De plus, puisque la suite $c(y^k)$ converge de façon monotone vers $c(y^*)$, puisque la fonction $\mu \mapsto -F_{y^*}(\mu)$ est convexe et puisque $\Delta_k(y^*) \in \partial(-F_{y^*})(c(y^k))$, nous obtenons que

$$\Delta_k(y^*) \leq \Delta_{k+1}(y^*) \leq \cdots \leq a_*$$

où $a_* \in \partial(-F_{y^*})(c(y^*))$. Dès lors, $\lim_{k\uparrow\infty} \Delta_k(y^*) := \Delta_\infty(y^*)$ existe et par la semi-continuité supérieure de la fonction $\mu \mapsto \partial(-F_{y^*})(\mu)$, nous obtenons que

$$\Delta_\infty(y^*) \in \partial(-F_{y^*})(c(y^*)).$$

Comme $\Delta_\infty(y^*) \leq a_*$ pour tout $a_* \in \partial(-F_{y^*})(c(y^*))$, nous obtenons, par le lemme 7.2.2, que

$$\Delta_\infty(y^*) = \min_{x \in S_{y^*}(c(y^*))} \left\{ y^{*T} g(x) \right\}. \tag{7.27}$$

Pour conclure notre analyse du comportement de la suite $1 - \frac{\Delta_k(y^*)}{\delta_{k+1}}$ lorsque $k \to \infty$, observons, par (7.26) et (7.27), que

$$0 \leq \limsup_{k\uparrow\infty} \left(1 - \frac{\Delta_k(y^*)}{\delta_{k+1}}\right) = 1 - \liminf_{k\uparrow\infty} \frac{\Delta_k(y^*)}{\delta_{k+1}} = 1 - \frac{\Delta_\infty(y^*)}{\delta_\infty} < 1.$$

Par conséquent, la suite $\{c(y^k)\}$ converge bien linéairement.

- En ce qui concerne la convergence superlinéaire, il nous suffit de montrer que

$$1 - \frac{\Delta_k(y^*)}{\delta_{k+1}} \to 0$$

ou encore que

$$\frac{\Delta_\infty(y^*)}{\delta_\infty} = 1.$$

Or, nous avons

$$\begin{aligned} \Delta_\infty(y^*) &\overset{(7.27)}{=} \min_{x \in S_{y^*}(c(y^*))} \left\{ y^{*T} g(x) \right\} \\ &= y^{*T} g(x^*). \end{aligned}$$

Nous avons également, par (7.26) et par le lemme 7.2.2, que

$$\begin{aligned} \delta_\infty &\in \partial(-F_{y^*})(c(y^*)) \\ &\in \left[\inf_{x \in S_{y^*}(c(y^*))} \left\{ y^{*T} g(x) \right\}, \sup_{x \in S_{y^*}(c(y^*))} \left\{ y^{*T} g(x) \right\} \right] \end{aligned}$$

Afin de montrer que $\dfrac{\Delta_\infty(y^*)}{\delta_\infty} = 1$, c'est-à-dire que $\Delta_\infty(y^*) = \delta_\infty$, montrons que l'intervalle $\left[\inf_{x \in S_{y^*}(c(y^*))} \left\{y^{*T} g(x)\right\}, \sup_{x \in S_{y^*}(c(y^*))} \left\{y^{*T} g(x)\right\}\right]$ est réduit à un singleton. Pour ce faire, supposons que le problème

$$\min_{x \in S} y^{*T} \left[f(x) - c(y^*) g(x)\right]$$

possède une seule solution que nous noterons x^*. Montrons, alors, en deux étapes, que cette solution x^* est aussi l'unique solution du problème

$$\min_{x \in S} \frac{y^{*T} f(x)}{y^{*T} g(x)} = c(y_*).$$

1. Nous avons

$$y^{*T} f(x^*) = c(y^*) g(x^*).$$

D'où,

$$y^{*T} \left[f(x^*) - c(y_*) g(x^*)\right] = 0.$$

2. Pour tout $x \neq x^*$, nous avons

$$\begin{array}{rrcl} & \frac{y^{*T} f(x)}{y^{*T} g(x)} & > & c(y^*) \\ \Leftrightarrow & y^{*T} f(x) & > & y^{*T} c(y^*) g(x) \\ \Leftrightarrow & y^{*T} \left[f(x) - c(y^*) g(x)\right] & > & 0 \overset{1.}{=} y^{*T} \left[f(x^*) - c(y^*) g(x^*)\right]. \end{array}$$

Par conséquent, x^* est l'unique solution du problème

$$\min_{x \in S} \frac{y^{*T} f(x)}{y^{*T} g(x)} = c(y^*).$$

Dès lors, nous aurons que

$$\delta_\infty = y^{*T} g(x^*) = \Delta_\infty(y^*).$$

$\square$

Conclusion

Dans ce mémoire, nous nous sommes intéressés aux méthodes primales et duales de type Dinkelbach permettant de résoudre des problèmes de programmation fractionnelle généralisée. Nous avons développé l'algorithme de type Dinkelbach dans le cas général mais aussi dans les cas compact et linéaire. Cet algorithme consiste à remplacer le problème de programmation fractionnelle généralisée par un problème de recherche de zéros d'une fonction paramétrique et possède un taux de convergence linéaire. Ensuite, dans le cas compact, nous avons modifié l'algorithme décrit précédemment pour obtenir, dans ce cas, un taux de convergence superlinéaire. Nous avons également montré que les problèmes de programmation fractionnelle généralisée avec un ensemble convexe pouvaient être résolus plus efficacement en utilisant une méthode de point proximal inexacte plutôt qu'une méthode exacte principalement lorsque le problème initial possède plusieurs solutions. La stratégie consistait à ajouter un terme de régularisation proximale à la fonction paramétrique associée au problème initial et à construire l'algorithme sur base des ε-approximations. Sous certaines conditions, cette méthode converge superlinéairement. Finalement, pour le problème de programmation fractionnelle généralisée dans le cas convexe, nous avons construit un dual sans saut de dualité et montré que ce dernier pouvait également être résolu via une approche paramétrique. Cette méthode possède un taux de convergence linéaire voire superlinéaire sous certaines conditions.

D'un point de vue numérique, selon les différents articles utilisés dans ce mémoire, il semblerait que la méthode utilisant les ε-approximations et la méthode duale soient plus efficaces que les méthodes primales de type Dinkelbach.

Bibliographie

[1] J.P. Crouzeix, J.A. Ferland & S. Schaible, *An Algorithm for Generalized Fractional Programs*, Journal of Optimization Theory and Applications, Vol. 47, pp 35-49, Septembre 1985.

[2] J.P. Crouzeix, J.A. Ferland & S. Schaible, *A Note on an Algorithm for Generalized Fractional Programs*, Journal of Optimization Theory and Applications, Vol. 50, No 1, pp 183-187, Juillet 1986.

[3] J.J. Strodiot, J.P. Crouzeix, J.A. Ferland & V.H.Nguyen, *An Inexact Proximal Point Method for Solving Generalized Fractional Programs*, Département de Mathématiques, FUNDP, Juin 2005.

[4] A.I. Barros, J.B.G. Frenk, S. Schaible & S. Zhang, *A New Algorithm for Generalized Fractional Programs*, Mathematical Programming, Vol. 72, pp 147-175, 1996.

[5] R. Jagannathan & S. Schaible, *Duality in Generalized Fractional Programming via Farkas' Lemma*, Journal of Optimization Theory ans Applications, Vol. 41, No 3, pp 417-423, Novembre 1983.

[6] O.L. Mangasarian, *Nonlinear Programming*, Classics in Applied Mathematics 10, SIAM, 1994.

[7] J.J. Strodiot, *An introduction to Nonsmooth Optimization*, Summer school on Optimization and Applied Mathematics, Can Tho University, Août 2003.