

UNIVERSITÉ MOULOUD MAMMERI DE TIZI-OUZOU
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES

MÉMOIRE DE MAGISTER

En Mathématiques

Option : Recherche Opérationnelle et Optimisation

THÈME

Application de la Programmation bi-niveaux au Problème de Contrôle Optimal

PRÉSENTÉ PAR :

Mme HARRACHE Fazia épouse HERMIME

DEVANT LE JURY COMPOSÉ DE :

M. HAMADOUCHE	Djamal	Professeur	UMMTO	Président
M. AIDENE	Mohamed	Professeur	UMMTO	Rapporteur
M. OUANES	Mohand	M. C. A.	UMMTO	Examineur
M. SADI	Bachir	M. C. A.	UMMTO	Examineur
M. OUKACHA	Brahim	M. C. A.	UMMTO	Examineur

2013

REMERCIEMENTS

Je remercie d'abord et avant tout le bon Dieu tout puissant de m'avoir donné le courage et la force pour réaliser ce travail.

Je désire remercier Monsieur Aidene Mohamed, mon directeur de mémoire, de m'avoir proposé ce sujet, et de sa disponibilité, son aide précieuse et ses conseils qui m'ont accompagné tout au long de ce travail. Je lui suis très reconnaissante d'avoir cru en mes capacités et de m'avoir accordé sa confiance, et aussi pour ses qualités scientifiques, pédagogiques et humaines.

Il m'est très agréable de remercier Monsieur Hamadouche Djamel qui m'a fait l'honneur de présider le jury de soutenance. Je tiens également à remercier Messieurs Sadi Bachir, Oukacha Brahim et Ouanes Mohand qui ont accepté d'être examinateurs de mon mémoire. Je tiens très vivement à exprimer ma reconnaissance à ces honorables enseignants, et à tous mes enseignants de l'université Mouloud Mammeri de Tizi-Ouzou pour les connaissances appréciables qu'ils m'ont données tout au long de mes études.

Je remercie chacun et chacune de mes professeurs qui se sont succédés, dès mes premiers pas à la maternelle jusqu'au DES, pour m'éduquer et m'apprendre la science et les bonnes manières. C'est grâce à eux et à leurs encouragements que j'ai pu poursuivre avec réussite mes études.

Il m'est impossible de passer sous silence l'immense influence qu'ont eue mes amis sur moi, tout particulièrement M. Hamid Karim, M. Belhadj, Daya, Houria, Nacera, Lynda, Souad, Ahcene pour leurs précieux soutiens. Je remercie tout les enseignants avec qui je travaille à l'établissement privé "Nedjma" sans oublier notre aimable directrice. Je tiens à vous remercier tous.

J'aimerais remercier mes parents, toute ma famille et ma belle famille pour leur constant soutien. Merci à mon oncle Ahmed et sa femme pour m'avoir encourager sans cesse pour aller toujours plus loin.

J'exprime enfin un geste modeste de remerciement et de reconnaissance à mon mari pour son infinie patience et sa compréhension et j'adresse toute ma tendresse maternelle à ma fille Lila.

Table des matières

Remerciements	i
Table des matières	iv
Introduction	1
1 Concepts fondamentaux	6
1.1 Convexité	6
1.1.1 Ensembles convexes	6
1.1.2 Fonctions convexes	9
1.2 Programmation mathématique	10
1.2.1 Optimisation sous contraintes	10
1.2.2 Notion de solution	11
1.2.3 Lagrangien et condition KKT	13
1.3 Fonctions multivoques	16
1.3.1 Définitions	16
1.3.2 Continuité des fonctions multivoques	17
1.4 Optimisation paramétrique	17
1.4.1 Définitions et notations	18
1.4.2 Conditions de régularité	19
2 Programmation mathématique bi-niveaux	21
2.1 Programmation bi-niveaux	21
2.1.1 Formulation mathématique d'un problème de programmation bi-niveaux(PBN)	21
2.1.2 Formulation mathématique d'un problème de programmation multi-niveaux	22
2.1.3 Multiplicité des solutions optimales de second niveau .	24
2.1.4 Autres formulations du problème bi-niveaux	26
2.2 Programmation linéaire bi-niveaux (PBL)	27
2.2.1 Complexité de la programmation bi-niveaux	31

2.2.2	Effet du déplacement des contraintes	31
2.2.3	Cas de l'unicité des solutions du niveau inférieur	33
2.2.4	Condition d'optimalité	35
2.2.5	Méthodes de résolution du problème bi-niveaux linéaire	38
2.2.6	Relation PBL- Problèmes classiques	42
3	Contrôle optimal bi-niveaux	47
3.1	Contrôle Optimal	47
3.1.1	Position du problème	47
3.2	Contrôlabilité	49
3.2.1	Contrôlabilité des systèmes linéaires	49
3.2.2	Contrôlabilité des systèmes linéaires autonomes	50
3.3	Stabilisation	52
3.3.1	Bouclage statique	52
3.3.2	concepts de stabilité	53
3.4	Principe du maximum de Pontryaguin	54
3.5	Éléments d'analyse	56
3.5.1	La dérivée directionnelle généralisée	57
3.5.2	Cône normal de Clarke	58
3.6	Problème du contrôle optimal bi-niveaux	61
3.6.1	Etude du problème	62
3.6.2	Différentiabilité de la fonction valeur	64
3.6.3	Conditions nécessaire d'optimalité	65
3.6.4	Cas particulier	66
	Conclusion	83
	Bibliographie	85

INTRODUCTION

La programmation mathématique se propose pour objet l'étude théorique des problèmes d'optimisation, ainsi que la conception et la mise en œuvre des algorithmes de résolution.

Depuis le début de l'histoire de l'humanité, l'homme a toujours été confronté à des situations où il doit décider quelle action va-t-il entreprendre afin qu'il puisse tirer le meilleur résultat dans différentes circonstances. Euclide, prolifique mathématicien de l'Antiquité, et qui a rédigé l'un des plus célèbres textes de l'histoire des mathématiques, *les Éléments*, formulait déjà des problèmes d'optimisation au III^e siècle avant J. C. Sir Isaac Newton (1642-1727), auteur du célèbre *Mathematical Principles of Natural Philosophy*, ainsi que Gottfried Wilhelm Leibniz (1646-1716) offrirent, à la fin du XVII^e siècle, les premiers outils de résolution de certains problèmes d'optimisation relatifs à la géométrie et à la physique. Ils avaient en effet inventé, simultanément mais néanmoins indépendamment, le calcul différentiel. Mais ce n'est que récemment qu'on a pu vraiment élaborer des méthodes scientifiques pour résoudre une classe importante de ce genre de problèmes. Dans les années 30, une approche systématique des problèmes de décision commence à se développer notamment aux Etats-Unis d'Amérique qui étaient en proie à une dépression économique sans précédent, et c'était principalement pour accompagner le programme *New-Deal* qui avait pour objectif de contrecarrer les effets de la grande crise économique de 1929. Parallèlement, des tentatives similaires sont apparues dans d'autres régions du monde touchées aussi par la crise.

Durant la seconde guerre mondiale, le Royaume-Uni fait appel à un groupe de scientifiques dans différentes disciplines pour mettre en œuvre leurs compétences scientifiques dans le but d'apporter assistance aux militaires dans l'élaboration de nouvelles stratégies et techniques de guerre. Par la suite, les Etats-Unis, encouragés par les bons résultats obtenus sur le

terrain par le Royaume-Uni, entament la même démarche. Les méthodes appliquées par ces scientifiques sont appelées *Recherches opérationnelles* (R.O), vu qu'elles étaient destinées à des recherches dans le domaine des opérations militaires ; c'est aussi, à cette période ,et pour les mêmes raisons, qu'apparaît la *Programmation Mathématique*. On associe généralement la naissance de cette discipline à la découverte de la *méthode du simplexe* en 1947 par G. B. Dantzing, qui était conseiller pour l'armée de l'air américaine. Beaucoup de théories en R.O ont été développées durant cette période comme réponses aux problèmes en lien direct avec des besoins militaires.

Après la fin de la seconde guerre mondiale, les techniques de la Recherche Opérationnelle ont été introduites dans le domaine des affaires, l'industrie, le commerce, la gestion et dans beaucoup d'autres domaines de l'activité quotidienne de l'homme. L'apparition des ordinateurs, et la puissance sans cesse croissante de ces machines, va permettre de propulser la programmation mathématique pour devenir aujourd'hui une branche très active des mathématiques appliquées.

Les problèmes d'optimisation à deux niveaux (ou bi-niveaux) ont été formulés pour la première fois en 1934 par H.V. Stackelberg dans un écrit sur l'économie de marché. D'ailleurs, un problème particulier de programmation mathématique à deux niveaux, traité durant des années en théorie des jeux, porte le nom de *jeu de Stackelberg*.

Dans les années mille neuf cent soixante-dix du XX^e siècle, les problèmes de programmation bi-niveaux ont été introduits dans le domaine de l'optimisation par J. Bracken et J. Mc Gill (*Mathematical programs with optimization problems in the constraints*. Operation Research, 1973. *Defense applications of mathematical programs with optimization problems in the constraints*. Operation Research, 1974.). Toutefois Candler et Norton furent les premiers à utiliser la terminologie de programmation à deux niveaux ou à plusieurs niveaux dans un rapport de la Banque Mondiale en 1977. Ainsi ces problèmes ont connu un développement en théorie, comme dans les applications pratiques puisqu'ils étaient, et le sont toujours, un thème de recherche auquel des mathématiciens, des économistes et des ingénieurs ont consacré beaucoup de travaux et de contributions.

Ces problèmes ont plusieurs applications dans différents domaines scientifiques : En économie, ils peuvent servir comme modèle qui permet aux décideurs à différents niveaux de responsabilités de prendre les meilleures

décisions pour un meilleur rendement [58, 69]. En sciences expérimentales et l'ingénierie [19, 48], en transport urbain, politique et militaire [57], en contrôle de pollution et même en chimie [33]. les problèmes de programmation à deux niveaux constituent des modèles adéquats dans plusieurs situations.

La théorie du contrôle analyse les propriétés des systèmes commandés, c'est-à-dire des systèmes sur lesquels on peut agir au moyen d'une *commande* (ou contrôle). Le but est alors d'amener le système d'un état initial donné à un certain état final, en respectant éventuellement certains critères.

Du point de vue mathématique, un système de contrôle est un système dynamique dépendant d'un paramètre appelé *le contrôle*. Pour le modéliser, on peut avoir recours à des équations différentielles, intégrales, aux dérivées partielles, stochastiques, etc. Pour cette raison la théorie du contrôle est à l'interconnexion de nombreux domaines mathématiques.

La théorie du contrôle optimal s'inscrit dans la continuité du calcul des variations. Elle est apparue après la seconde guerre mondiale, répondant à des besoins pratiques de guidage, notamment dans le domaine de l'aéronautique et de la dynamique du vol. Historiquement, la théorie du contrôle optimal est très liée à la mécanique classique, en particulier au principe variationnel de la mécanique (équations d'Euler-Lagrange, par exemple). Le point clé de cette théorie est *le pincipe du maximum de Pontryagin*, formulé par L. S. Pontryagin en 1956, qui donne une condition nécessaire d'optimalité et permet ainsi de calculer la *trajectoire optimale*. Les points forts de la théorie ont été la découverte de la méthode de programmation dynamique, l'introduction de l'analyse fonctionnelle dans la théorie des systèmes optimaux, la découverte des liens entre la solution d'un problème de contrôle optimal et des résultats de la théorie de stabilité de Lyapunov.

L'objet de ce travail porte sur le problème d'optimisation bi-niveaux et son application au problème de contrôle optimal. Le problème de programmation bi-niveaux est contrôlé par deux décideurs en interdépendance hiérarchique. Le premier est appelé *Leader* ou décideur du niveau supérieur, le second est appelé *Suiveur* ou décideur du niveau inférieur. Le contrôle des variables de décision est partagé entre les deux décideurs. En rajoutant l'hypothèse de la linéarité des fonctions objectifs et des contraintes du *Leader* et du *Suiveur*, on obtient sa version appelée *problème bi-niveaux linéaire*. Dans ce cas, il a été prouvé que la solution est atteinte en un certain point extrême

de l'ensemble des contraintes qui est un polyèdre. Malgré sa simplicité apparente, le problème de programmation bi-niveaux linéaire est un problème compliqué et difficile à résoudre. Beaucoup d'approches ont été proposées dans la littérature pour sa résolution. Les plus populaires sont basées sur sa reformulation en un programme mathématique à un seul niveau.

Ce présent travail est divisé en trois chapitres, organisé de la manière suivante :

Le premier chapitre présente les concepts fondamentaux qui vont servir de références toute au long de ce document. Les éléments de l'analyse convexe jouent un rôle important dans les problèmes de programmation mathématique. L'optimisation sous contraintes et les fonctions multivoques constituent des éléments de base pour les problèmes de programmation bi-niveaux.

Le deuxième chapitre est partagé en deux parties : La première est consacrée aux problèmes de programmation bi-niveaux. Deux approches sont dégagées pour le traitement de ce genre de problèmes selon l'absence ou l'existence de la coopération entre le *Leader* et le *Suiveur*. On parle de la *position optimiste* si le *leader* se trouve en situation de coopération avec le *suiveur*, et de la *position pessimiste* dans le cas contraire. Quant à la seconde partie, elle est vouée au problème de programmation bi-niveaux linéaire. Pour ce modèle les fonctions objectives et les contraintes sont linéaires. Malgré la linéarité, ces problèmes sont NP-difficiles ; Jeroslow est le premier à montrer la complexité de ces problèmes [63, 28, 39].

Le troisième et le dernier chapitre de ce mémoire est consacré à l'étude du problème de contrôle optimal bi-niveaux. Notre démarche est inspirée des outils de l'analyse non régulière (non lisse) [24, 25, 26, 60], c'est-à-dire les fonctions intervenant dans la formulation du problème ne sont pas nécessairement différentiables au sens usuel. Ce type de problèmes a été étudié pour la première fois par Chen et Cruz [23]. Nous commençons par une présentation des notions essentielles d'un problème de contrôle optimal, la contrôlabilité et la stabilité. Puis nous exposons quelques notions de l'analyse non régulière avec lesquelles sera traité le problème de contrôle optimal bi-niveaux. Nous entamons la partie traitant les problèmes de contrôle bi-niveaux par sa formu-

lation mathématique [70, 71] et poser certaines hypothèses qui vont définir le cadre de notre étude. Pour formuler une condition nécessaire d'optimalité, la démarche consiste à ramener le problème à deux niveaux en un problème à un seul niveau en faisant intervenir la fonction valeur du problème du niveau inférieur. Pour finir ce chapitre, nous présentons un exemple modélisant une situation d'intérêt économique et environnementale.

Ce travail se termine par une conclusion et quelques perspectives.

Chapitre 1

Concepts fondamentaux

1.1 Convexité

Dans cette partie, on se limite à quelques faits fondamentaux de l'analyse convexe qui nous sont indispensables pour la suite de ce travail. Le lecteur peut consulter [17, 16, 12, 29, 49, 56] au sujet de la convexité.

1.1.1 Ensembles convexes

Définition 1.1. Un sous-ensemble C de \mathbb{R}^n est dit *convexe* si :

$$\alpha x + (1 - \alpha)y \in C, \quad \forall x, y \in C, \forall \alpha \in [0, 1].$$

Autrement dit, un sous ensemble convexe contient toujours le segment $[x, y]$ joignant deux de ces points x et y .

Notons que l'ensemble vide et un ensemble ne contenant qu'un seul point sont, par convention, considérés comme des ensembles convexes.

Exemple 1.1. Les ensembles suivants sont des exemples remarquables d'ensembles convexes dans \mathbb{R}^n .

- i. *Hyperplan.* $H = \{x : a^T x = \alpha\}$, où a est un vecteur non nul de \mathbb{R}^n , et α un scalaire.
- ii. *Demi-espace.* $S = \{x : a^T x \leq \alpha\}$, où a est un vecteur non nul de \mathbb{R}^n et α un scalaire.

- iii. *Polyèdre.* $P = \{x : Ax \leq b\}$, où A est une $m \times n$ -matrice et b un m -vecteur.
- iv. *Voisinage.* $U_\delta(x^0) = \{x \in \mathbb{R}^n : \|x - x^0\| < \delta\}$, où x^0 est un vecteur fixé de \mathbb{R}^n et $\delta > 0$.

Définition 1.2. Un vecteur x de \mathbb{R}^n est appelé *combinaison convexe* des p vecteurs x_1, x_2, \dots, x_p s'il existe des constantes non négatives $\lambda_1, \lambda_2, \dots, \lambda_p$ avec $\lambda_1 + \lambda_2 + \dots + \lambda_p = 1$ telles que $x = \sum_{i=1}^p \lambda_i x_i$.

On appelle *enveloppe convexe* d'un sous-ensemble quelconque C de \mathbb{R}^n qu'on note $\text{conv}(C)$ l'ensemble de toutes les combinaisons convexes des points de C .

Définition 1.3. Soit $M \subset \mathbb{R}^n$. M est dit *ensemble affine* si :

$$(1 - \lambda)x + \lambda y \in M \quad \text{pour tout } x \in M, y \in M \text{ et } \lambda \in \mathbb{R}.$$

Un ensemble affine est aussi appelé par certains auteurs *variété affine* ou *variété linéaire*. L'intersection d'une famille d'ensembles affines est aussi un ensemble affine. Ainsi, pour $S \subset \mathbb{R}^n$, il existe un plus petit ensemble affine qui contient S (à savoir l'intersection de la famille des ensembles affines M tels que $S \subset M$). Cet ensemble est appelé *enveloppe affine* de S , il est noté par $\text{aff}S$.

Définition 1.4. Soit C une partie convexe de \mathbb{R}^n . Un point $x \in C$ est un *point extrême* de C si et seulement si x ne peut pas s'écrire comme combinaison convexe $\lambda y + (1 - \lambda)z$ telle que $y \in C, z \in C, 0 < \lambda < 1$ et $y \neq z$.

Définissons, maintenant, la notion d'*hyperplan d'appui* (*hyperplan de support*) d'un ensemble.

Définition 1.5. Soit $S \subset \mathbb{R}^n$, et x_0 un point qui appartient à la frontière de S . Si $a \neq 0$ satisfait $a^T x \leq a^T x_0$ pour tout $x \in S$, alors l'hyperplan $\{x : a^T x = a^T x_0\}$ est appelé *hyperplan de support* de S au point x_0 .

L'interprétation géométrique est que l'hyperplan $\{x : a^T x = a^T x_0\}$ est tangent à S au point x_0 , et l'ensemble S est contenu dans le demi-espace $\{x : a^T x \leq a^T x_0\}$.

Un résultat important, appelé *le théorème de l'hyperplan de support*, énonce que pour tout ensemble non vide et convexe C , et pour tout point x_0 de la frontière de C , il existe un hyperplan de support de C au point x_0 .

Définition 1.6. On appelle *ensemble polyédral* un ensemble qui est une intersection finie de demi-espaces fermés de \mathbb{R}^n , autrement dit l'ensemble de solutions d'un système d'inéquations $Ax \leq b$, avec A une $m \times n$ -matrice, $x \in \mathbb{R}^n$ et $b \in \mathbb{R}^m$.

Définition 1.7. Un polyèdre convexe borné est appelé *polytope convexe*.

Définition 1.8. Un *polygone convexe* est un polytope convexe à deux dimensions.

Une caractéristique topologique importante des ensembles convexes est due à la notion de l'intérieur relatif qu'on définit en regardant l'ensemble convexe comme sous-ensemble de son enveloppe affine.

Définition 1.9. Soit C une partie convexe de \mathbb{R}^n . On dit que x est un point *intérieur relatif* de C si $x \in C$ et s'il existe une sphère ouverte S de centre x telle que $(S \cap \text{aff}C) \subset C$. L'ensemble de tous les points intérieurs relatifs de C est appelé *intérieur relatif* de C , il est noté $ri(C)$.

Définition 1.10. Une *face* d'un ensemble convexe C est une partie convexe C' de C telle que chaque segment de C avec un point intérieur relatif de C' a ses deux points extrémités dans C' .

Définition 1.11. Un sous-ensemble M est un *cône* si :

$$\forall x \in M, \quad \forall \alpha \geq 0, \quad \alpha x \in M$$

Un *cône* est donc un union de demi-droites fermées issues de l'origine.

Définition 1.12. Un *cône* M est dit *convexe* si et seulement s'il est stable par addition, c'est-à-dire :

$$\forall x \in M, \quad \forall y \in M, \quad x + y \in M$$

1.1.2 Fonctions convexes

Définition 1.13. Soit C un sous-ensemble convexe de \mathbb{R}^n . Une fonction $f : C \rightarrow \mathbb{R}$ est dite *convexe* si :

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad \forall x, y \in C, \forall \alpha \in [0, 1].$$

f est dite *strictement convexe* si l'inégalité stricte est toujours vérifiée pour tous $x, y \in C$ avec $x \neq y$ et pour tout $\alpha \in]0, 1[$.

Remarque 1.1. Une fonction $f : C \rightarrow \mathbb{R}$, où C est un ensemble convexe de \mathbb{R}^n est dite *concave* (respectivement *strictement concave*) si $(-f)$ est convexe (respectivement strictement convexe).

On est parfois amené à considérer des fonctions convexes prenant des valeurs $+\infty$ ou $-\infty$, et on a la définition suivante :

Définition 1.14. On dit que f est une fonction convexe de \mathbb{R}^n dans $\overline{\mathbb{R}} = [-\infty, +\infty]$ (bornes incluses) si et seulement si :

$$\left. \begin{array}{l} \forall x \text{ tel que } f(x) \neq +\infty \\ \forall y \text{ tel que } f(y) \neq +\infty \end{array} \right\} \Rightarrow f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Définition 1.15. On définit l'*épigraphe* d'une fonction $f : X \rightarrow [-\infty, +\infty]$ où $X \subset \mathbb{R}^n$, comme un sous-ensemble de \mathbb{R}^{n+1} donné par :

$$\text{epi}(f) = \{(x, \alpha) | x \in X, \alpha \in \mathbb{R}, f(x) \leq \alpha\}.$$

Le *domaine effectif* de f est l'ensemble

$$\text{dom}(f) = \{x \in X | f(x) < \infty\}.$$

La propriété fondamentale due à la notion de convexité apparaît dans le théorème suivant :

Théorème 1.1. [52] *Si X est un sous-ensemble convexe de \mathbb{R}^n et $f : X \rightarrow]-\infty, +\infty[$ est une fonction convexe, alors un minimum local de f sur X est aussi un minimum global de f sur X .
Si de plus f est strictement convexe, alors il existe au plus un minimum global de f sur X .*

Dans le cas des fonctions différentiables, le théorème ci-après, donne une caractérisation de la convexité.

Théorème 1.2. [13] *Soit $C \in \mathbb{R}^n$ un sous-ensemble convexe et soit $f : C \rightarrow \mathbb{R}$ une fonction différentiable sur C . Alors*

i. La fonction f est convexe si et seulement si

$$f(z) \geq f(x) + (z - x)^T \nabla f(x), \quad \forall x, z \in C.$$

ii. Si l'inégalité précédente est stricte quand $x \neq z$, alors f est strictement convexe.

1.2 Programmation mathématique

Nous nous plaçons dans \mathbb{R}^n qui, muni de la norme euclidienne

$$\|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$$

est un espace vectoriel normé.

1.2.1 Optimisation sous contraintes

Un problème d'optimisation sous contraintes est représenté comme suit :

$$\begin{cases} \min_{x \in X} f(x) \\ g(x) \leq 0, \\ h(x) = 0, \end{cases} \quad (1.1)$$

où $x \in \mathbb{R}^n$, $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ appelée *variable de décision*,
 $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est la *fonction objectif* (on dit aussi : fonction économique),

$g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et $h : \mathbb{R}^n \longrightarrow \mathbb{R}^q$ sont des fonctions vectorielles appelées *fonctions contraintes*, avec :

$$g(x) = (g_1(x), g_2(x), \dots, g_p(x)) \text{ et } h(x) = (h_1(x), h_2(x), \dots, h_q(x)).$$

Remarque 1.2. On dit généralement que le problème (1.1) est un problème de *minimisation*. La recherche du maximum d'une fonction f se ramène immédiatement au problème de minimisation de $F = -f$. Dans ce qui suit, nous n'envisageons que le cas de la minimisation.

1.2.2 Notion de solution

Définition 1.16. L'ensemble $\Omega = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1 \cdots p, h_j(x) = 0, j = 1 \cdots q\}$ est appelé *ensemble des solutions réalisables* du problème (1.1)

Définition 1.17. On appelle *solution optimale* (ou encore : *optimum global*) de (1.1) une solution x^* qui minimise $f(x)$ sur l'ensemble de toutes les solutions réalisables, c'est-à-dire telle que : $f(x^*) \leq f(x)$ pour toute solution $x \in \Omega$.

Définition 1.18. soit $x^* \in \Omega$. On dit que x^* est un *minimum local* s'il existe $\epsilon > 0$ tel que :

$$f(x^*) \leq f(x), \forall x \in B_\epsilon(x^*)$$

Où $B_\epsilon(x^*) = \{x \in \mathbb{R}^n \mid \|x - x^*\| < \epsilon\}$.

Cet ensemble est communément appelé boule de rayon ϵ centrée en x^* .

Définition 1.19. Soit $x^* \in \Omega$. On dit que $y \in \mathbb{R}^n$ est une *direction admissible* s'il existe $\alpha > 0$ tel que :

$$x^* + ty \in \Omega, \forall t \in [0, \alpha].$$

Remarque 1.3. On notera qu'il est souvent possible de caractériser les optimums locaux d'un problème, c'est-à-dire de donner les conditions nécessaires et/ou suffisantes pour qu'une solution x soit un optimum local. Par contre, il est généralement impossible de caractériser les optimums globaux sauf dans le cas particulier de la *programmation convexe*.

Illustrons les notions d'optimum local et d'optimum global par la figure (1.1) sur une fonction à une seule variable.

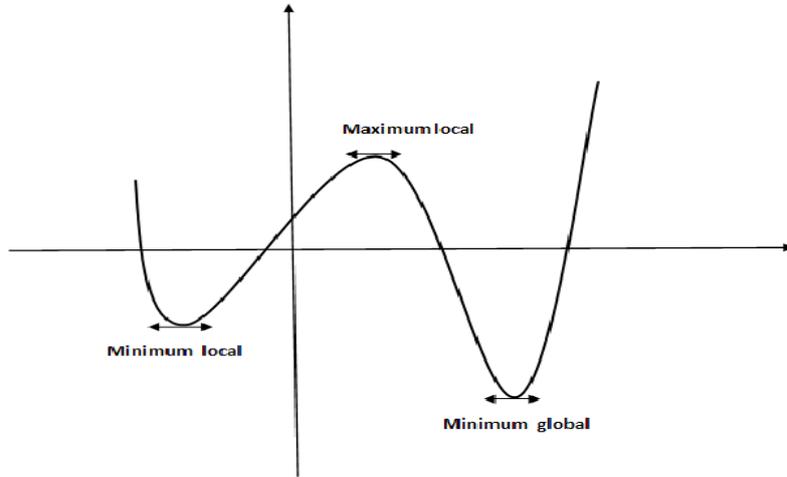


FIG. 1.1 – Optimum local et optimum global

Remarque 1.4. On note, généralement, l'ensemble des solutions optimales du problème (1.1) par

$$\arg \min_{x \in X} f(x).$$

Mais il arrive parfois qu'on rencontre, par abus d'écriture, ce qui suit :

$$x^* = \arg \min_{x \in X} f(x).$$

Comme nous en ferons souvent usage le long de ce document, rappelons le célèbre théorème de Weierstrass qui concerne l'existence d'une solution optimale pour un problème d'optimisation.

Théorème 1.3 (Weierstrass). *Soit $K \subset \mathbb{R}^n$ une partie compacte de \mathbb{R}^n et soit $f : K \rightarrow \mathbb{R}$ une fonction réelle continue, alors le problème d'optimisation*

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

admet une solution optimale $x^ \in K$.*

1.2.3 Lagrangien et condition KKT

Considérons le problème de programmation avec des contraintes d'égalité

$$\begin{cases} \min_x f(x), \\ h_j(x) = 0, j = 1, \dots, r \quad (r < n). \end{cases} \quad (1.2)$$

Lagrange transforma ce problème avec contraintes en un problème sans contraintes en introduisant ce qu'on appelle les *multiplieurs de lagrange* $\lambda_i = 1, \dots, r$ dans la formulation de la *fonction de Lagrange* (ou le *lagrangien*) suivante :

$$L(x, \lambda) = f(x) + \sum_{j=1}^r \lambda_j h_j(x). \quad (1.3)$$

Supposons que les fonctions f et h_j ($j = 1, \dots, r$) sont continûment différentiables et faisons l'hypothèse que la matrice jacobienne

$$\frac{\partial h(x^*)}{\partial x} = [\nabla h_1(x^*), \dots, \nabla h_r(x^*)]$$

est de rang r . Alors une condition nécessaire pour que x^* soit solution optimale de (1.2) est qu'il existe un vecteur λ^* tel que :

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x^*, \lambda^*) &= 0, \quad i = 1, \dots, n, \\ \frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) &= 0, \quad j = 1, \dots, r. \end{aligned}$$

Pour un problème avec contraintes d'inégalité

$$\begin{cases} \min_x f(x), \\ g_j(x) \leq 0, j = 1, \dots, m, \end{cases} \quad (1.4)$$

on définit encore le lagrangien par

$$L(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x). \quad (1.5)$$

Karush et par la suite Kuhn et Tucker énoncèrent des conditions nécessaires pour que x^* soit solution de (1.4). Ces conditions sont connues sous le nom de *conditions KKT*, et s'expriment en utilisant le lagrangien $L(x, \lambda)$.

Considérons le problème (1.4), et supposons que les fonctions f, g_j ($j = 1, \dots, m$) sont continûment différentiables, et supposons qu'ils existent les multiplicateurs de Lagrange λ^* au point x^* .

Théorème 1.4. [66](conditions KKT) *Si x^* est une solution optimale du problème (1.4) alors les conditions suivantes sont satisfaites*

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^m \lambda_j^* \frac{\partial g_j}{\partial x_i}(x^*) &= 0, i = 1 \dots n, \\ g_j(x^*) &\leq 0, j = 1 \dots m, \\ \lambda_j^* g_j(x^*) &= 0, j = 1 \dots m, \\ \lambda_j^* &\geq 0, j = 1 \dots m. \end{aligned}$$

Considérons maintenant le problème de programme mathématique non linéaire avec contraintes d'égalité et d'inégalité suivant :

$$\begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ g_i(x) \leq 0 \quad i = 1 \dots p, \\ h_j(x) = 0 \quad j = 1 \dots q. \end{cases} \quad (1.6)$$

Où les fonctions f, g_i et h_j sont au moins deux fois continûment différentiables.

Définition 1.20. Le *Lagrangien* du programme mathématique (1.6) est défini par :

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \mu_j h_j(x)$$

Définition 1.21. L'ensemble

$$\Lambda(x) = \{(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q : \lambda^T g(x) = 0, \nabla_x L(x, \lambda, \mu) = 0\}$$

est appelé l'ensemble des *multiplicateurs de Lagrange* correspondant au point $x \in \mathbb{R}^n$.

Théorème 1.5. [59](Conditions nécessaires du 1^{er} ordre)
Soit x^ un minimum local régulier de (1.6). Alors, il existe des multiplicateurs*

$\lambda^* \in \mathbb{R}^p$ et $\mu^* \in \mathbb{R}^q$ tels que :

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^q \mu_j^* \nabla h_j(x^*) &= 0. \\ \nabla_x L(x^*, \lambda^*, \mu^*) &= 0 \\ \lambda_i^* &\geq 0 \quad i = 1 \dots p \\ \lambda_i^* g_i(x^*) &= 0 \quad i = 1 \dots p \\ g_i(x^*) &\leq 0 \quad i = 1 \dots p \\ h_j(x^*) &= 0 \quad i = 1 \dots q \end{aligned}$$

Les conditions nécessaires du premier ordre sont souvent appelées les conditions de Karush-Kuhn-Tucker (KKT). Dans le cas convexe (c'est-à-dire lorsque $f(x)$ et $g_i(x)$ sont des fonctions convexes et $h_j(x)$ des fonctions affines), un point x^* régulier est un minimum global pour le programme (1.6) si et seulement s'il satisfait les conditions de KKT.

Les conditions du deuxième ordre utilisent les notations suivantes :

$$\begin{aligned} I(x) &= \{i = 1 \dots p : g_i(x) = 0\}, \\ I'(x) &= \{i = 1 \dots p : g_i(x) = 0, \lambda_i > 0\}, \\ T(x) &= \{y \in \mathbb{R}^n : \nabla g_i(x)y \leq 0, i \in I(x), \nabla h_j(x)y = 0, j = 1 \dots q\}, \\ T'(x) &= \{y \in \mathbb{R}^n : \nabla g_i(x)y \leq 0, i \in I'(x), \nabla h_j(x)y = 0, j = 1 \dots q\}. \end{aligned}$$

L'ensemble $I(x)$ correspond à l'ensemble des indices des contraintes d'inégalités actives au point x .

L'ensemble $I'(x)$ ne contient que les indices des contraintes d'inégalités actives qui vérifient la complémentarité stricte. $T(x)$ est le cône des directions admissibles au point x qui est inclus dans le cône élargi $T'(x)$. Ces ensembles permettent la caractérisation de la matrice hessienne réduite du Lagrangien.

Théorème 1.6. [59](Conditions nécessaires du 2ème ordre) Soit x^* un minimum local régulier de (1.6) alors, il existe des multiplicateurs $\lambda^* \in \mathbb{R}^p$ et $\mu^* \in \mathbb{R}^q$ tels que :

- Les conditions de KKT sont satisfaites,
- $y^T \nabla^2 L(x^*, \lambda^*, \mu^*) y \geq 0, \quad \forall y \in T(x^*)$.

Différentes formulations du lagrangien et des conditions KKT dans différents contextes de programmes mathématiques existent dans la littérature. Nous renvoyons le lecteur à [13, 16, 65, 52].

1.3 Fonctions multivoques

1.3.1 Définitions

Soient X et Y deux ensembles non vides. Une *fonction multivoque* (on dit aussi *multi-fonction* ou *correspondance*) Γ de X dans Y est une correspondance qui associe à chaque élément x de X un sous-ensemble $\Gamma(x)$ de Y . Le sous-ensemble $\Gamma(x)$ est appelé *image* de x par la fonction multivoque Γ .

L'ensemble $D_\Gamma = \{x \in X, \Gamma(x) \neq \emptyset\}$ est appelé *domaine* (ou *ensemble de définition*) de Γ , et l'ensemble

$$\bar{Y} = \bigcup_{x \in X} \Gamma(x)$$

est appelé *ensemble des valeurs* de Γ .

Si B est un sous-ensemble de Y , l'ensemble

$$\Gamma^{-1}(B) = \{x \in X : \Gamma(x) \subset B\} \subset X$$

est appelé *image réciproque* de B par Γ .

L'ensemble

$$\text{graph}\Gamma = \{(x, y) \in X \times Y : y \in \Gamma(x)\}$$

est appelé *graphe* de la fonction multivoque Γ .

Définition 1.22. [63] Soient A et Y deux ensembles arbitraires et soit la fonction multivoque $\Gamma : A \longrightarrow 2^Y$. Considérons la famille de parties \mathcal{D} de Y avec $\mathcal{D} = \{\Gamma(a) : a \in A\}$.

- i. Une *fonction choisie* de Γ est la fonction $\gamma : A \longrightarrow Y$ telle que $\gamma(a) \in \Gamma(a), \forall a \in A$.
- ii. Le *produit cartésien* de \mathcal{D} est l'ensemble des fonctions choisies de Γ , noté $\prod_{a \in A} \Gamma(a)$, tel que :

$$\prod_{a \in A} \Gamma(a) = \{\gamma : A \longrightarrow Y \mid \gamma(a) \in \Gamma(a) \quad \forall a \in A\}.$$

1.3.2 Continuité des fonctions multivoques

Nous introduisons ci-après la notion de continuité pour les fonctions multivoques.

Définition 1.23. Soient $X \subset \mathbb{R}^k$, $Y \subset \mathbb{R}^l$, et considérons la fonction multivoque

$$\Gamma : X \longrightarrow 2^Y.$$

- i. On dit que Γ est semi-continue inférieurement au point $z \in X$ si pour chaque partie ouverte Z avec $\Gamma(z) \cap Z \neq \emptyset$, il existe un voisinage ouvert $U_\delta(z)$ de z tel que $\Gamma(t) \cap Z \neq \emptyset$ pour tout $t \in U_\delta(z)$.
- ii. On dit que Γ est semi-continue supérieurement au point $z \in Y$ si pour chaque partie ouverte Z avec $\Gamma(z) \subset Z$, il existe un voisinage ouvert $U_\delta(z)$ de z tel que $\Gamma(t) \subset Z$ pour tout $t \in U_\delta(z)$.
- iii. On dit que la fonction multivoque Γ est continue au point $z \in X$ si elle est à la fois semi-continue inférieurement et semi-continue supérieurement au point z .

Définition 1.24. Soit X un ensemble ouvert de \mathbb{R}^n et Y une partie de \mathbb{R}^m . Une fonction multivoque $\Gamma : X \longrightarrow 2^Y$ est dite *uniformément compacte* au point $\tilde{x} \in X$ s'il existe un voisinage $U_\delta(\tilde{x})$ de \tilde{x} tel que l'ensemble $\Gamma(U_\delta(\tilde{x}))$ soit borné.

1.4 Optimisation paramétrique

Dans cette section, on s'intéressera à quelques résultats de l'optimisation paramétrique qui seront d'une grande utilité pour aborder les problèmes de la programmation mathématique bi-niveaux. L'optimisation paramétrique constitue un domaine de recherche très actif, toutefois, on abordera que les points dont on aura besoin pour la suite, c'est-à-dire dans le traitement des problèmes bi-niveaux. Voir, par exemple, les références suivantes pour plus de détails sur l'optimisation paramétrique [8, 42].

1.4.1 Définitions et notations

Considérons le problème d'optimisation suivant :

$$\min_x \{f(x, y) : g(x, y) \leq 0, h(x, y) = 0\}, \quad (1.7)$$

où $f : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^p$, $h : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^q$ sont des fonctions différentiables.

Notons :

$$\begin{aligned} g(x, y) &= (g_1(x, y), g_2(x, y), \dots, g_p(x, y)), \\ h(x, y) &= (h_1(x, y), h_2(x, y), \dots, h_q(x, y)). \end{aligned}$$

On notera par $\nabla_x g(x, y)$ le gradient de g au point x pour y fixé.

Définition 1.25. Le problème (1.7) est appelé *problème d'optimisation paramétrique convexe* si toutes les fonctions $f(\cdot, y)$, $g_i(\cdot, y)$, $i = 1, \dots, p$ sont convexes et les fonctions $h_j(\cdot, y)$, $j = 1, \dots, q$ sont affines sur \mathbb{R}^n pour tout $y \in \mathbb{R}^m$ fixé. De même, il sera appelé *problème d'optimisation paramétrique linéaire* quand toutes les fonctions $f(\cdot, y)$, $g_i(\cdot, y)$, $i = 1, \dots, p$, $h_j(\cdot, y)$, $j = 1, \dots, q$ sont linéaires sur \mathbb{R}^n pour tout $y \in \mathbb{R}^m$ fixé.

On donnera ci-après quelques définitions inhérentes au problème paramétrique (1.7) et auxquelles on fera souvent référence.

Définition 1.26. Considérons le problème (1.7).

i. Définissons la fonction suivante :

$$V : \mathbb{R}^m \longrightarrow \mathbb{R}$$

$$V(y) = \min_x \{f(x, y) : g(x, y) \leq 0, h(x, y) = 0\}.$$

V est appelée *fonction de la valeur optimale* du problème (1.7).

ii. Définissons la fonction multivoque suivante :

$$M : \mathbb{R}^m \longrightarrow 2^{\mathbb{R}^n}$$

$$M(y) = \arg \min_x \{f(x, y) : g(x, y) \leq 0, h(x, y) = 0\}.$$

M est appelée *correspondance des solutions optimales* du problème (1.7). Par définition, la fonction multivoque M associe à chaque $y \in \mathbb{R}^m$ l'ensemble des solutions optimales globales du problème (1.7).

iii. Définissons la fonction multivoque suivante :

$$\Psi : \mathbb{R}^m \longrightarrow 2^{\mathbb{R}^n}$$

$$\Psi(y) = \{x \in \mathbb{R}^n : g(x, y) \leq 0, h(x, y) = 0\}.$$

Ψ est appelée *correspondance des solutions admissibles* du problème (1.7).

iv. L'ensemble des solutions optimales locales du problème (1.7) est noté par :

$$M_{loc}(y) = \{x \in M(y) : \exists \varepsilon > 0, f(x, y) \leq f(z, y), \forall z \in \Psi(y) \cap U_\varepsilon(x)\},$$

où : $U_\varepsilon(x) = \{z \in \mathbb{R}^n : \|x - z\| \leq \varepsilon\}$ est un voisinage ouvert de x .

1.4.2 Conditions de régularité

La formulation de l'existence de solutions pour un problème de programmation mathématique résulte d'une façon générale de ses conditions de régularité. On donnera quelques unes, pour cela, considérons le problème (1.7).

Définition 1.27. Soit $y^0 \in \mathbb{R}^m$ et $x^0 \in \Psi(y^0)$. On dira que la *qualification des contraintes de Mangasarian-Fromowitz* (MFCQ)¹ est satisfaite en (x^0, y^0) s'il existe une direction $d \in \mathbb{R}^n$ telle que :

$$\nabla_x g_i(x^0, y^0)d < 0, \text{ pour tout } i \in I(x^0, y^0) = \{j : g_j(x^0, y^0) = 0\},$$

$$\nabla_x h_j(x^0, y^0)d = 0, \text{ pour tout } j = 1, \dots, q,$$

et les gradients $\{\nabla_x h_j(x^0, y^0), j = 1, \dots, q\}$ sont linéairement indépendants.

On désigne par $I(x, y) = \{j : g_j(x, y) = 0\}$, l'ensemble des *contraintes actives* en $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$.

On notera par :

$$L(x, y, \lambda, \mu) = f(x, y) + \lambda^T g(x, y) + \mu^T h(x, y),$$

la *fonction lagrangienne* du problème (1.7).

Et par :

$$\Lambda(x, y) = \{(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q : \lambda \geq 0, \lambda^T g(x, y) = 0, \nabla_x L(x, y, \lambda, \mu) = 0\},$$

¹En anglais : Mangasarian-Fromowitz Constraint Qualification

Les *multiplicateurs de Lagrange* correspondant au point $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$.

La *qualification des contraintes de Mangasarian-Fromowitz stricte*, en abrégé (SMFCQ) en (x^0, y^0) pour le problème (1.7) quant à elle s'écrit : il existe un vecteur de multiplicateurs de Lagrange (λ, μ) et une direction d tels que :

$$\begin{aligned} \lambda &\geq 0, \lambda^T g(x^0, y^0) &= 0, \\ L(x^0, y^0, \lambda, \mu) &= \nabla f(x^0, y^0) + \lambda^T \nabla g(x^0, y^0) + \mu^T \nabla h(x^0, y^0) &= 0, \\ \nabla g_i(x^0, y^0)^T d &< 0, \\ \forall i : g_i(x^0, y^0) &= \lambda_i &= 0 \\ \nabla g_i(x^0, y^0)^T d &= 0, \forall i : \lambda_i = 0, \\ \nabla h_j(x^0, y^0)^T d &\leq 0, \forall j = 1, \dots, q. \end{aligned}$$

Définition 1.28. L'ensemble

$$SP(y) = \{x \in \Psi(y) : \Lambda(x, y) \neq \emptyset\}$$

est appelé *ensemble des solutions stationnaires* du problème (1.7).

Définition 1.29. On dit que la *qualification des contraintes de l'indépendance linéaire* (LICQ)² est satisfaite au point (x^0, y^0) si les gradients

$$\{\nabla_x g_i(x^0, y^0), i \in I(x^0, y^0)\} \cup \{\nabla_x h_j(x^0, y^0), j = 1, \dots, q\}$$

sont linéairement indépendants.

Théorème 1.7. [33] *Considérons le problème (1.7) au point (x^0, y^0) dans $\mathbb{R}^n \times \mathbb{R}^m$ avec $x^0 \in M_{loc}(y^0)$. Alors (MFCQ) est satisfaite en (x^0, y^0) si et seulement si $\Lambda(x^0, y^0)$ est un polyèdre non vide, convexe et compact. Et on a*

$$M(y^0) \subseteq M_{loc}(y^0) \subseteq SP(y^0)$$

Théorème 1.8. [33] *Considérons le problème (1.7). Si on a :*

i. L'ensemble $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : g(x, y) \leq 0, h(x, y) = 0\}$ est non vide et compact.

ii. (MFCQ) est satisfaite en tout point (x, y) avec $x \in \Psi(y)$.

Alors la correspondance des solutions $M(y)$ est semi-continue supérieurement et la fonction de la valeur optimale $V(\cdot)$ du problème (1.7) est continue en y .

²En anglais : Linear Independence Constraint Qualification

Chapitre 2

Programmation mathématique bi-niveaux

2.1 Programmation bi-niveaux

Un problème de programmation à deux niveaux sont des problèmes d'optimisation dont les contraintes sont déterminées, en partie, par un autre problème d'optimisation. En d'autres termes, ce sont des programmes mathématiques hiérarchiques où l'ensemble de toutes les variables est partitionné entre deux vecteurs x et y , une valeur du vecteur x étant donnée, le vecteur y est choisit comme étant solution optimale d'un problème d'optimisation paramétré en x .

2.1.1 Formulation mathématique d'un problème de programmation bi-niveaux(PBN)

Dans le cas général un problème de programmation bi-niveaux (PBN) est un problème qui se présente comme suit [33] :

$$\left\{ \begin{array}{l} \min_x F(x, y) \\ \\ s.c \left\{ \begin{array}{l} G(x, y) \leq 0 \\ \min_y f(x, y) \\ s.c \ g(x, y) \leq 0 \end{array} \right. \end{array} \right. \quad (2.1)$$

où l'ensemble de toutes les variables est partitionnée entre un vecteur $x \in \mathbb{R}^{n_1}$ représentant le premier niveau de décision (niveau supérieur), et un vecteur $y \in \mathbb{R}^{n_2}$ pour le second niveau de décision (niveau inférieur). De même, les fonctions $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ et $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ sont les fonctions objectives du premier niveau de décision et de second niveau de décision respectivement, et les fonctions vectorielles $G : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_1}$ et $g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_2}$ sont respectivement les contraintes du niveau supérieur et niveau inférieur.

2.1.2 Formulation mathématique d'un problème de programmation multi-niveaux

La formulation générale d'un programme multi-niveaux, qui est en fait une généralisation du problème bi-niveaux est la suivante :

$$\left\{ \begin{array}{l} f_1(x_1, x_2, x_3, \dots, x_p) \longrightarrow \min_{x_1} \\ g_1(x_1, x_2, x_3, \dots, x_p) \leq 0 \\ \text{où } x_2, x_3, \dots, x_p \text{ sont solution de} \\ \\ \left\{ \begin{array}{l} f_2(x_1, x_2, x_3, \dots, x_p) \longrightarrow \min_{x_2} \\ g_2(x_1, x_2, x_3, \dots, x_p) \leq 0 \\ \text{où } x_3, \dots, x_p \text{ sont solution de} \\ \\ \left\{ \begin{array}{l} f_3(x_1, x_2, x_3, \dots, x_p) \longrightarrow \min_{x_3} \\ g_3(x_1, x_2, x_3, \dots, x_p) \leq 0 \\ \vdots \\ \text{où } x_p \text{ est solution de} \\ \\ \left\{ \begin{array}{l} f_p(x_1, x_2, x_3, \dots, x_p) \longrightarrow \min_{x_p} \\ g_p(x_1, x_2, x_3, \dots, x_p) \leq 0 \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right. \quad (2.2)$$

Exemple 2.1. [33] Considérons le problème du niveau inférieur défini comme suit.

$$\left\{ \begin{array}{l} \min_y f(x, y) = -xy, \\ 0 \leq y \leq 1, \end{array} \right.$$

où $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$,

pour tout $x \in \mathbb{R}$,

$$M(x) = \arg \min_y \{-xy : 0 \leq y \leq 1\}.$$

Et considérons le problème de programmation bi-niveaux suivant :

$$\begin{cases} \min_x F(x, y) = x^2 + y^2, \\ y \in M(x), \\ -1 \leq x \leq 1, \end{cases}$$

avec $F : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$.

Après évaluation du problème du niveau inférieur, et en portant sa solution optimale dans la fonction objectif du niveau supérieur, on obtient les résultats suivants :

$$M(x) = \begin{cases} \{0\} & \text{si } x < 0, \\ \{1\} & \text{si } x > 0, \\ [0, 1] & \text{si } x = 0. \end{cases}$$

$$F(x, y(x)) = \begin{cases} x^2 & \text{si } x < 0, \\ 1 + x^2 & \text{si } x > 0, \\ \in [0, 1] & \text{si } x = 0. \end{cases}$$

Remarquons qu'au point $x = 0$, la valeur minimum de la fonction $x \mapsto F(x, y(x))$ (voir figure 2.2) est égale à zéro, mais cette valeur est atteinte uniquement si $F(0; y(0)) = 0$, i.e. dans le cas où le suiveur choisit $y = 0$ quand $x = 0$. Si ce n'est pas le cas, alors le problème bi-niveaux considéré n'a pas de solution.

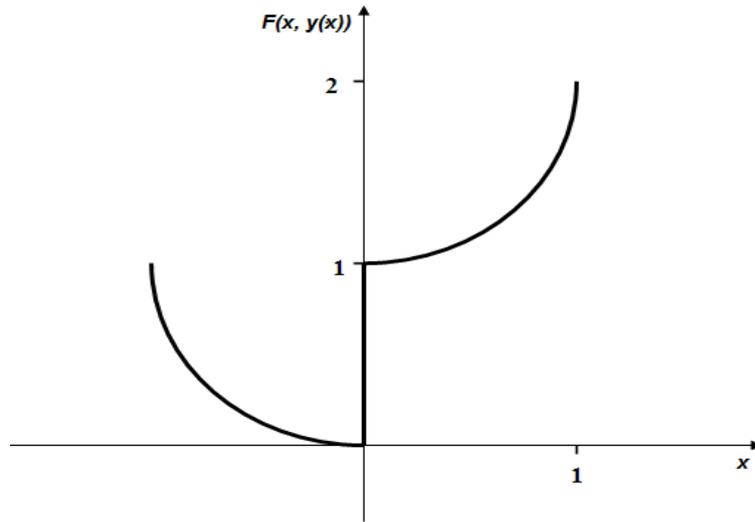


FIG. 2.1 – Le graphe de $x \rightarrow F(x, y(x))$

2.1.3 Multiplicité des solutions optimales de second niveau

Pour un $x \in \mathbf{R}^{n_1}$ fixé, notons par $M(x)$ l'ensemble des solutions optimales du problème de second niveau. On définit ainsi la fonction multivoque

$$M : \mathbf{R}^{n_1} \rightarrow 2^{\mathbf{R}^{n_2}},$$

où $2^{\mathbf{R}^{n_2}}$ désigne l'ensemble de toutes les parties de \mathbf{R}^{n_2} .

On a donc pour tout $x \in \mathbf{R}^{n_1}$,

$$M(x) = \arg \min_y \{f(x, y) : g(x, y) \leq 0\}. \quad (2.3)$$

Soit $y(x) \in M(x)$ la *réaction rationnelle* du suiveur lorsque le leader choisit x . connaissant cette réaction, le problème PBN peut se formuler comme suit :

$$\begin{cases} \min_x F(x, y(x)), \\ G(x, y(x)) \leq 0 \\ y(x) \in M(x) \end{cases} \quad (2.4)$$

Ce problème est dit problème de leader ou problème bi-niveaux.

Théorème 2.1. [33] Si l'ensemble $\{(x, y) : g(x, y) \leq 0, G(x, y) \leq 0\}$ est non vide, compact et pour tout x avec $G(x, y) \leq 0$, la qualification des contraintes de Mangasarian-Fromowitz est satisfaite, alors le problème (2.4) a une solution optimale.

L'approche optimiste

Le *leader* pourrait utiliser cette approche s'il suppose que le *suiveur* est disposé à coopérer, c'est-à-dire que le *suiveur* choisit une solution $y(x) \in M(x)$ la meilleure du point de vue du *leader*. On désigne par :

$$m^0(x) = \min_y \{F(x, y) : y \in M(x)\} \quad (2.5)$$

la valeur optimiste de la fonction économique du premier niveau. Alors, l'approche optimiste du problème de programmation bi-niveaux est réduite à :

$$\min \{m^0(x) : G(x, y) \leq 0\} \quad (2.6)$$

Un couple $(\bar{x}, y(\bar{x}))$ tel que \bar{x} est solution de (2.6), et $y(\bar{x})$ solution de (2.5) pour $x = \bar{x}$, est également une solution optimale pour le problème :

$$\min_{x,y} \{F(x, y) : G(x, y) \leq 0\} \quad (2.7)$$

et vice versa, à condition que le dernier ait une solution optimale. La plus part des contributions à la programmation bi-niveaux sont consacrées à ce problème. Une des raisons pourrait être que ce problème a une solution optimale sous des hypothèses raisonnables. La propriété ci-après énonce les conditions d'existence d'une solution optimale pour le problème (2.7).

Propriété. [34]

Si l'ensemble des solutions réalisables $\Omega = \{(x, y) : G(x, y) \leq 0, g(x, y) \leq 0\}$ est non vide et compact, et l'hypothèse de qualification des contraintes de Mangasarian-Fromowitz est satisfaite pour tout x tel que $G(x, y) \leq 0$, alors le problème (2.7) a une solution.

L'approche pessimiste

L'approche optimiste semble impossible lorsque la coopération n'est pas autorisée, ou alors lorsque l'aptitude du *suiveur* à respecter des ententes n'est

pas assurée. Alors dans cette situation déplaisante, une voie à suivre pour le leader consiste à se prémunir contre les choix indésirables du suiveur. Ce qui conduit au problème :

$$\min\{m^p(x) : G(x, y) \leq 0\} \quad (2.8)$$

Où

$$m^p(x) = \max_y \{F(x, y) : y \in M(x)\}$$

représente la pire valeur de la fonction économique du premier niveau qui peut être atteinte sur l'ensemble des solutions du problème de second niveau.

Propriété. [34]

Supposons que l'application multivoque $M(x)$ est semi-continue inférieurement en tout point x tel que $G(x, y) \leq 0$ et que l'hypothèse de qualification des contraintes de Mangasarian-Fromovitz est satisfaite en chacun de ces points. Alors le problème (2.8) a une solution optimale.

2.1.4 Autres formulations du problème bi-niveaux

Sous l'hypothèse d'univocité de l'application $M(x)$ et celle de la compacité du domaine réalisable, le problème bi-niveaux se reformule comme un problème à un seul niveau en x :

$$\begin{cases} \min_x F(x, y(x)) \\ s.c \begin{cases} x \in \Omega_x^2, \\ G(x, y(x)) \leq 0 \end{cases} \end{cases} \quad (2.9)$$

où

$$\Omega_x^2 = \{x \in \mathbb{R}^{n_1} / \exists y \in \mathbb{R}^{n_2} \text{ avec } g(x, y) \leq 0\}$$

désigne la trace de l'ensemble des solutions réalisables

$$\Omega = \{(x, y) : G(x, y) \leq 0, g(x, y) \leq 0\}.$$

Lorsque l'application $M(x)$ n'est pas univoque, on obtient la formulation :

$$\begin{cases} \min_x F(x, y) \\ s.c \begin{cases} (x, y) \in \Omega, \\ y \in M(x). \end{cases} \end{cases} \quad (2.10)$$

La résolution de ces problèmes est compliquée en raison de la présence des fonctions et des ensembles définis implicitement.

Moyennant des hypothèses de régularité sur le problème de second niveau d'un problème de programmation bi-niveaux (BPN) sous forme particulière, on peut remplacer ce dernier par ses conditions d'optimalité de Karush-Kuhn-Tucker (KKT). Il en résulte un programme mathématique non linéaire à un seul niveau dont la formulation est :

$$PBN_{KKT} : \left\{ \begin{array}{l} \min_{x,y} F(x, y) \\ s.c \left\{ \begin{array}{l} G(x) \leq 0, \\ g(x, y) \leq 0, \\ \nabla_y f(x, y) + \lambda^t \nabla_y g(x, y) = 0, \\ \lambda^t g(x, y) = 0, \\ \lambda \geq 0, \lambda \in \mathbb{R}_+^{m_2}. \end{array} \right. \end{array} \right. \quad (2.11)$$

L'approche de la fonction de la valeur optimale suivante donne une autre formulation du PBN. Définissons la *fonction valeur* du problème de second niveau comme suit : $V : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ telle que :

$$V(x) = \inf_y \{f(x, y) : g(x, y) \leq 0\},$$

où $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ est la droite linéaire achevée. Alors il est évident que le PBN peut se reformuler comme suit :

$$\left\{ \begin{array}{l} \min_{x,y} F(x, y) \\ s.c \left\{ \begin{array}{l} f(x, y) - V(x) \leq 0, \\ g(x, y) \leq 0, \\ G(x, y) \leq 0. \end{array} \right. \end{array} \right. \quad (2.12)$$

2.2 Programmation linéaire bi-niveaux (PBL)

Le problème de programmation linéaire bi-niveaux (PBL) est un cas particulier des problèmes de programmation bi-niveaux, où les fonctions objectifs

et les contraintes du *Leader* et du *Suiveur* sont linéaires.

$$\left\{ \begin{array}{l} \min_x F(x, y) = c_1^t x + d_1^t y, \\ s.c \ A_1 x + B_1 y \leq b_1, \\ x \geq 0, \\ \\ \min_y f(x, y) = c_2^t x + d_2^t y, \\ s.c \ \begin{cases} A_2 x + B_2 y \leq b_2, \\ y \geq 0, \end{cases} \end{array} \right. \quad (2.13)$$

où $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}$, $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ sont les fonctions objectifs du *leader* et du *suiveur* respectivement; $c_i \in \mathbb{R}^{n_1}$, $d_i \in \mathbb{R}^{n_2}$, $b_i \in \mathbb{R}^{m_i}$, $A_i \in \mathbb{R}^{m_i} \times \mathbb{R}^{n_1}$, $B_i \in \mathbb{R}^{m_i} \times \mathbb{R}^{n_2}$, $i = 1, 2$.

Les contraintes $A_1 x + B_1 y \leq d_1$ (respectivement $A_2 x + B_2 y \leq d_2$), sont les contraintes du premier (respectivement du second) niveau.

La fonction linéaire $F(x, y)$ (respectivement $f(x, y)$) représente la fonction économique du *leader* (respectivement du *suiveur*), alors que x (respectivement y) désigne le vecteur des variables du *leader* (respectivement du *suiveur*).

Afin de donner une caractérisation d'une solution du problème PBL (2.13), on a besoin des définitions suivantes :

Définition 2.1.

- i. L'ensemble $\Omega = \{(x, y) : x \geq 0, y \geq 0, A_i x + B_i y \leq b_i, i = 1, 2\}$. définit *l'ensemble des solutions réalisables* du problème PBL (2.13).
- ii. Pour chaque valeur de la variable du premier niveau x , les contraintes de second niveau permettent de définir le *domaine réalisable du suiveur* :

$$\Omega_y(x) = \{y \geq 0 : B_2 y \leq b_2 - A_2 x\}$$

- iii. La *projection* de Ω sur l'ensemble des décisions du leader est :

$$P(x) = \{x \in \mathbb{R}^{n_1} : \exists y \in \mathbb{R}^{n_2}, A_1 x + B_1 y \leq b_1, A_2 x + B_2 y \leq b_2, x \geq 0, y \geq 0\}$$

- iv. Pour $x \in P(x)$, l'ensemble $M(x)$ des *solutions optimales* (ou *réactions rationnelles*) du suiveur est :

$$M(x) = \{y : y \in \arg \min\{f(x, y) : y \in \Omega_y(x)\}\}$$

v. La *fonction valeur* optimale du problème de second niveau pour $x \in P(x)$, fixé est

$$V(x) = \{f(x, y), y \in M(x)\}.$$

vi. L'ensemble réalisable $DI = \{(x, y) \in \Omega : y \in M(x)\}$ est appelé *domaine induit* (ou *région induite*).

Définition 2.2.

- i. (x, y) est un *point réalisable* si $(x, y) \in \Omega$.
- ii. (x, y) est un *point admissible* si (x, y) est un point réalisable et $y \in M(x)$.
- iii. Un point $(x^*, y^*) \in X \times Y$ est une *solution optimale* du problème (2.13) si (x^*, y^*) est admissible et pour tout point admissible (x, y) on a :

$$c_1^t x^* + d_1^t y^* \leq c_1^t x + d_1^t y$$

Propriété. [34] Lorsqu'un PBL admet une solution optimale, cette solution est atteinte en un point extrême du domaine induit .

Exemple 2.2. [63]

considérons l'exemple suivant :

$$\left\{ \begin{array}{l} \min_{x \geq 0} F(x, y) = x - 4y, \\ \left\{ \begin{array}{l} \min_{y \geq 0} f(x, y) = y, \\ -x - y \leq -3, \\ -2x + y \leq 0, \\ 2x + y \leq 12, \\ -3x + 2y \leq -4. \end{array} \right. \end{array} \right.$$

Le polyèdre de la figure (2.2) représente le domaine des contraintes Ω . Conformément à la définition (2.1), nous avons :

$$P(x) = \{x : 1 \leq x \leq 4\}$$

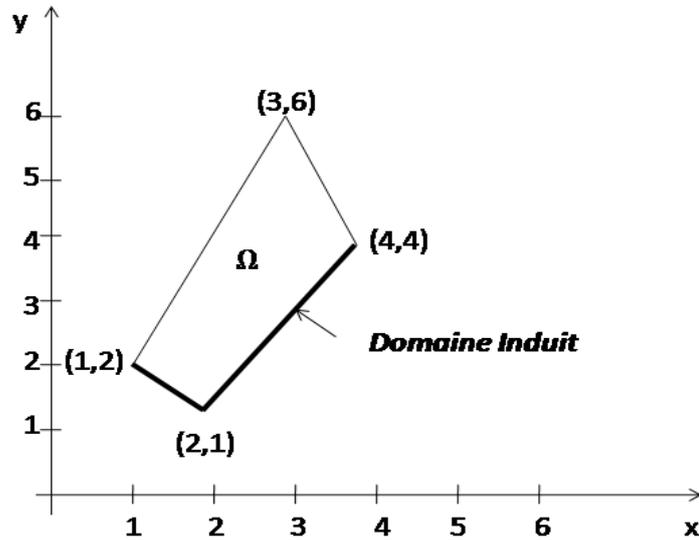


FIG. 2.2 – Géométrie du (PBL)

Pour un point fixé $x \in P(x)$, l'ensemble des solutions admissibles du suiveur $\Omega_y(x)$ est représenté par une ligne horizontale contenue dans le polyèdre Ω au point x .

Puisque l'objectif du suiveur est de minimiser y , l'ensemble de ses réactions rationnelles est l'ensemble des points admissibles se trouvant sur la ligne verticale, et il est donné par :

$$\begin{aligned}
 M(x) &= \max\{-x + 3, (3x - 4)/2\} \\
 &= \begin{cases} -x + 3 & \text{si } 1 \leq x \leq 2 \\ (3x - 4)/2 & \text{si } 2 \leq x \leq 4. \end{cases}
 \end{aligned}$$

Le domaine induit DI est représenté par la portion en gras du périmètre de Ω sur la figure (2.2).

La solution optimale est atteinte au point $(x^*, y^*) = (4, 4)$ avec $F^* = -12$ et $f^* = 4$. Le point $(3, 6)$ procure une meilleure valeur pour le leader mais ce point n'appartient pas au domaine induit.

Si le leader choisit la valeur $x = 3$, la réponse du suiveur sera $y = 2.5$, ce qui donnera $F(3, 2.5) = -7$ et $f(3, 2.5) = 2.5$. Cette situation est clairement meilleure pour le suiveur mais pas pour le leader.

2.2.1 Complexité de la programmation bi-niveaux

On dit qu'une classe de problème est NP-difficile dans le cas où il ne peut exister un algorithme de résolution pour un problème de cette classe en temps borné par une fonction polynomiale de la taille du problème, que si tous les problèmes de cette classe peuvent être résolus en temps polynomial. Le problème de programmation bi-niveaux linéaire est un problème difficile à résoudre malgré que les fonctions objectifs et les fonctions contraintes sont toutes linéaires, le PBL est ni continu partout, ni convexe. Jeroslow (cité dans [28]) est le premier à montrer que le problème PBL est NP-difficile.

2.2.2 Effet du déplacement des contraintes

Plusieurs auteurs se sont intéressés à l'effet que pourrait provoquer le déplacement des contraintes d'un niveau à un autre pour le problème PBL [9], [33], [62]. Nous allons voir à travers un simple exemple que le modèle change complètement.

Exemple 2.3. [51] considérons le problème suivant :

$$\left\{ \begin{array}{l} \min_x F(x, y) = -x - 2y, \\ \left\{ \begin{array}{l} \min_y f(x, y) = -y, \\ -3x + y \leq -3 \\ 3x + y \leq 30 \\ 2x - 3y \geq -12 \\ x + y \leq 14. \end{array} \right. \end{array} \right.$$

La région des contraintes de ce problème est représentée dans la figure (2.3). On remarque que cette ensemble est convexe. La solution de cet exemple est le point $B = (6, 8)$.

La région induite est représentée par le trait en gras. On voit bien qu'elle est non convexe mais continue (connectée).

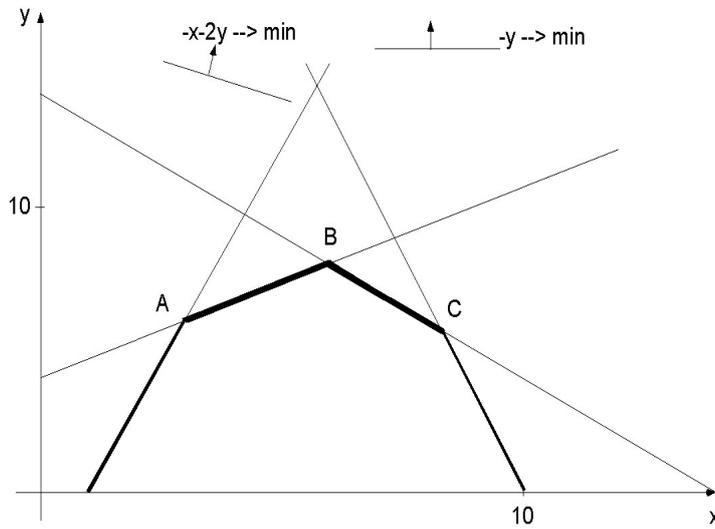


FIG. 2.3 – Représentation graphique de l'exemple (2.3)

Dans le cas de présence des contraintes du niveau supérieur (leader), cette région est souvent discontinue (voir l'exemple suivant).

Exemple 2.4. [51] Considérons le PBL de l'exemple (2.3) où les deux dernières contraintes du suiveur sont déplacées vers le niveau supérieur :

$$\left\{ \begin{array}{l} \min_x F(x, y) = -x - 2y, \\ 2x - 3y \geq -12 \\ x + y \leq 14 \end{array} \right\} \left\{ \begin{array}{l} \min_y f(x, y) = -y, \\ -3x + y \leq -3 \\ 3x + y \leq 30 \end{array} \right.$$

La région des contraintes ainsi que le domaine induit (en gras) sont représentées dans la figure (2.4)

On voit bien d'après le graphe de l'exemple (précédent) que le domaine induit est discontinu. Ceci est dû à la présence des contraintes du niveau supérieur. La solution globale de cet exemple est le point $C = (8, 6)$ et A est une solution locale.

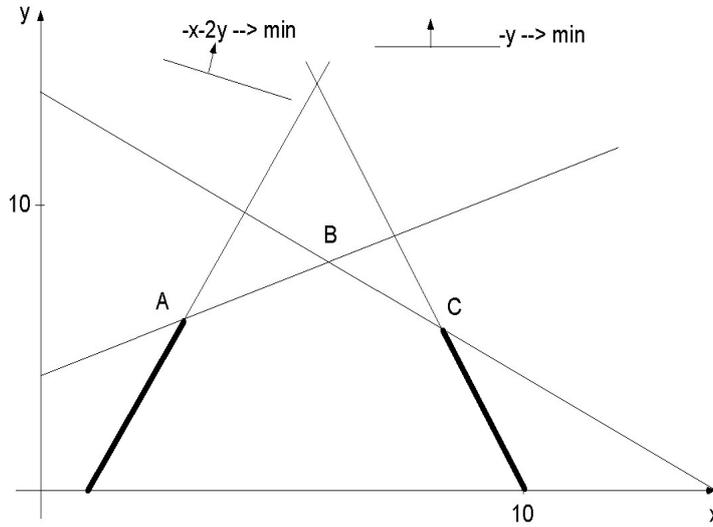


FIG. 2.4 – Représentation graphique de l'exemple (2.4)

2.2.3 Cas de l'unicité des solutions du niveau inférieur

Considérons le PBL (2.13), le domaine d'induction joue un rôle important dans le cas où le problème du niveau inférieur possède une solution optimale unique pour toute valeur x .

Définition 2.3. [51] Considérons le problème (2.13) et supposons que l'ensemble des réactions rationnelles du suiveur contient, au plus, un point pour toute valeur $x \in M(x)$. Alors un point $(\bar{x}, \bar{y}) \in DI$ est une solution optimale du problème (2.13) si

$$c_1\bar{x} + d_1\bar{y} \leq c_1x + d_1y, \quad \forall (x, y) \in DI.$$

Nous allons formuler le problème (2.13) suivant l'approche optimiste. Autrement dit, nous allons nous placer dans le cas où le leader peut influencer le choix du suiveur, ainsi il va le pousser à choisir un point y tel que :

$$y \in \arg \min_y \{c_1x + d_1y : y \in M(x)\},$$

i.e. une meilleure solution dans $M(x)$ du point de vue du leader. Ainsi donc le problème (2.13) est équivalent au problème de programme mathématique suivant :

$$\min_{x,y} \{c_1x + d_1y : (x, y) \in DI\}. \tag{2.14}$$

Définition 2.4. [51] Considérons le PBL (2.13) sous sa formulation optimiste (i.e. considérons le problème (2.14)). Un point $(\bar{x}, \bar{y}) \in DI$ est une solution optimale globale du problème (2.13) si :

$$c_1\bar{x} + d_1\bar{y} \leq c_1x + d_1y, \quad \forall (x, y) \in DI$$

Le point $(\bar{x}, \bar{y}) \in DI$ est une solution optimale locale du problème (2.14) s'il existe un voisinage $U_\delta(\bar{x}, \bar{y})$, $\delta > 0$, de (\bar{x}, \bar{y}) tel que :

$$c_1\bar{x} + d_1\bar{y} \leq c_1x + d_1y, \quad \forall (x, y) \in U_\delta((\bar{x}, \bar{y})) \cap DI$$

Notons qu'un problème de programmation mathématique bi-niveaux avec unicité des solutions du problème du niveau inférieur est un cas particulier des problèmes sous formulation optimiste, et nous allons considérer ce dernier cas pour chercher les conditions suffisantes d'existence de la solution optimale.

Lemme 2.2. [51] *Si le domaine des contraintes Ω du PBL (2.13) est non vide et compact alors le domaine induit DI est fermé.*

Preuve. Puisque Ω est non vide, alors il existe au moins un paramètre $\bar{x} \in P(x)$ tel que $\Omega_y(x) \neq \emptyset$. De l'hypothèse de compacité on a $M(\bar{x}) \neq \emptyset$. Considérons la suite $\{(x^k, y^k)\}_{k \geq 1}$ d'éléments de DI qui converge vers (\bar{x}, \bar{y}) , alors $\bar{y} \in M(\bar{x})$. Par suite $(\bar{x}, \bar{y}) \in DI$, ce qui achève la preuve. \square

Ce lemme donne une caractérisation importante du domaine induit DI sous l'hypothèse que Ω non vide et compact, toutefois il ne garantit pas que DI soit non vide. Dans le cas où il n'existe pas de contraintes couplantes (i.e. les contraintes qui lient les deux variables des niveaux inférieur et supérieur) pour le problème du leader, on a le résultat suivant :

Corollaire 2.3. *Si Ω est compact et non vide et s'il n'existe pas de contraintes couplantes pour le problème du niveau supérieur, alors le problème (2.13) admet une solution optimale.*

Preuve. En effet, dans ce cas le domaine induit DI est non vide, puisqu'il est fermé, il est donc compact ($DI \subset \Omega$). Il s'agit de minimiser un fonction continue sur un compact, le théorème de Weierstrass nous assure l'existence d'une solution optimale. \square

Lemme 2.4. [51] *Supposons que Ω est non vide et compact et que de plus DI est non vide alors DI est une réunion finie d'ensembles polyédraux.*

Corollaire 2.5. *Supposons que :*

- i. Ω est non vide et compact.*
- ii. DI est non vide.*
- iii. $M(x)$ contient au plus un élément pour tout $y \in \Omega_y(x)$.*

Alors DI décrit une fonction linéaire par morceaux.

Corollaire 2.6. *Avec les mêmes hypothèses du corollaire précédent, DI est égal à la réunion de faces de l'ensemble polyédral Ω .*

Théorème 2.7. [51] *Avec les hypothèses :*

- i. Ω est non vide et compact,*
- ii. DI est non vide,*

le problème standard (2.14) admet une solution optimale.

Remarque 2.1. La position pessimiste d'un PBL est un cas plus compliqué et même avec les hypothèses du théorème précédent, l'existence d'une solution optimale n'est pas garantie.

Théorème 2.8. [63] *La solution (x^*, y^*) du PBL est atteinte à un sommet de Ω .*

2.2.4 Condition d'optimalité

Reprenons le problème PBL formulé selon l'approche de la position optimiste, c'est-à-dire le PBL équivalent au problème standard

$$\min_{x,y} \{c_1x + d_1y : (x, y) \in DI\}. \quad (2.15)$$

Proposition 2.9. [51] *Une condition nécessaire et suffisante pour que le couple (x^*, y^*) soit solution locale du problème (2.15) est qu'il existe un vecteur λ^* tel que (x^*, y^*, λ^*) résout localement le problème suivant :*

$$\left\{ \begin{array}{l} \min \quad c_1x + d_1y, \\ A_1x + B_1y \leq b_1, \\ A_2x + B_2y \leq b_2, \\ B_2^T \lambda + d_2 = 0, \\ \lambda^T (b_2 - A_2x - B_2y) = 0, \\ \lambda \geq 0. \end{array} \right. \quad (2.16)$$

Preuve. Le problème (2.16) est le problème dual du problème (2.15). Le résultat du théorème est une conséquence des conditions de Karush-Kuhn-Tucker pour la programmation linéaire. \square

Considérons, maintenant, le problème (2.16), i.e. la formulation de Karush-Kuhn-Tucker d'un problème linéaire bi-niveaux dans la position optimiste. Dans ce qui suit, ce problème sera étudié en le décomposant suivant un système d'indices pour enfin aboutir à une condition nécessaire et suffisante d'optimalité.

Mais tout d'abord, on aura besoin de la définition suivante.

Définition 2.5 (Matrice à rang plein). Soit A une matrice à m lignes et n colonnes. Le *rang* de la matrice A , qu'on note par $rg(A)$, est le nombre maximum de ses vecteurs lignes (ou de ses vecteurs colonnes) linéairement indépendants.

En particulier, nous avons : $rg(A) \leq \min(m, n)$.

Si $rg(A) = \min(m, n)$, on dira que la matrice A est de *rang plein*.

Prenons un ensemble d'indices \tilde{I} vérifiant

$$\{i : \lambda_i^* > 0\} \subset \tilde{I} \subset \{i : (A_2x^* + B_2y^* - b_2)_i = 0\},$$

et considérons le problème suivant :

$$\left\{ \begin{array}{l} \min c_1x + d_1y, \\ A_1x + B_1y \leq b_1, \\ A_2x + B_2y \leq b_2, \\ B_2^T \lambda + d_2 = 0, \\ (A_2x + B_2y - b_2)_i = 0, i \in \tilde{I}, \\ \lambda_i = 0, i \notin \tilde{I}, \\ \lambda \geq 0. \end{array} \right. \quad (2.17)$$

Ici, la matrice $(A_2)_{i \in \tilde{I}}$ est à rang plein. Si (x^*, y^*) est une solution optimale de (2.15), (x^*, y^*, λ^*) est une solution globale de (2.17). A présent, considérons

$$\left\{ \begin{array}{l} \min c_1x + d_1y, \\ A_1x + B_1y \leq b_1, \\ A_2x + B_2y \leq b_2, \\ (A_2x + B_2y - b_2)_i = 0, i \in \tilde{I}, \end{array} \right. \quad (2.18)$$

et la variable duale λ est une solution de

$$\begin{cases} B_1^T \lambda + d_2 = 0, \\ \lambda_i = 0, i \notin \tilde{I}, \\ \lambda > 0, \end{cases} \quad (2.19)$$

avec $\tilde{I} \in D(x, y)$, où

$$D(x, y) = \{I : \exists \lambda \geq 0, B_1^T \lambda + d_2 = 0, \lambda_i = 0, \forall i \notin I, \\ (A_2 x + B_2 y - b_2)_i = 0, \forall i \in I, (B_2)_{i \in I} \text{ est à rang plein}\}.$$

Alors, si (x^*, y^*) est une solution optimale de (2.15), (x^*, y^*) est aussi une solution optimale de (2.18) pour tout $\tilde{I} \in D(x^*, y^*)$. De plus, si (x^*, y^*) est une solution optimale de (2.18) pour tout $\tilde{I} \in D(x^*, y^*)$, alors (x^*, y^*) est solution optimale locale de (2.15).

En gardant les mêmes notations, le théorème qui suit donne une condition nécessaire et suffisante d'optimalité utilisant les conditions KKT dans le cas optimiste.

Théorème 2.10. [51] *Le point (x^*, y^*) est une solution optimale du problème (2.15) si et seulement si c'est une solution optimale du problème (2.18) pour tout $\tilde{I} \in D(x^*, y^*)$.*

Preuve.

Condition nécessaire : Soit (x^*, y^*) solution optimale de (2.15). Puisque les conditions de Karush-Kuhn-Tucker sont nécessaires pour l'optimalité de (x^*, y^*) , ceci implique que (x^*, y^*, λ^*) est une solution optimale de (2.18) pour toute solution λ^* de (2.19).

Condition suffisante : En utilisant la contraposée de l'implication, on va montrer que si (x^*, y^*) n'est pas une solution de (2.15) alors il n'est pas solution de (2.18).

Supposons que $(x^*, y^*) \in DI$ n'est pas une solution optimale de (2.15). Anisi, il existe une suite $\{(x_n, y_n)\}$ qui converge vers (x^*, y^*) avec

$$(x_n, y_n) \in DI, c_1 x_n + d_1 y_n < c_1 x^* + d_1 y^* \quad \text{pour tout } n.$$

Comme on a $(x_n, y_n) \in DI : A_1 x_n + B_1 y_n \leq b_1, x_n \in P(y_n)$ i.e.

$$\exists \lambda_n \geq 0, A_2^T \lambda_n + c_2 = 0, \lambda_n^T (b_2 - A_2 x_n - B - 2y_n) = 0,$$

on peut prendre les λ_n comme sommets de

$$\lambda \geq 0, A_2^T \lambda + c_2 = 0, \lambda^T (b_2 - A_2 x_n - B_2 y_n) = 0,$$

de là les λ_n sont aussi sommets de

$$\{\lambda \geq 0, A_2^T \lambda + c_2 = 0, \lambda_i = 0, \text{ pour tout } i \notin I_n\}$$

pour un certain ensemble d'indices I_n satisfaisant

$$(b_2 - A_2 x_n - B - 2y_n)_i = 0, \quad i \in I_n.$$

Mais le nombre des différents ensembles $I \in \{1, 2, \dots, q\}$ est fini, il existe une suite $\{(x_n, y_n, \lambda_n)\}$ avec les λ_n sont sommets de

$$\{\lambda \geq 0 : A_2^T \lambda + c_2 = 0, \lambda_i = 0, \text{ pour } i \in \widehat{I}\}$$

pour un ensemble fixé \widehat{I} satisfaisant $(b_2 - A_2 x^* - B - 2y^*)_i = 0, i \in \widehat{I}$ pour tout n . Ceci signifie que $\widehat{I} \in D(x^*, y^*)$. Donc (x^*, y^*) ne peut pas être une solution optimale de (2.18) pour cet ensemble \widehat{I} .

□

2.2.5 Méthodes de résolution du problème bi-niveaux linéaire

Pour la résolution du problème de programmation bi-niveaux linéaire, de nombreux algorithmes ont été mis au point et particulièrement pour le cas linéaire [28],[63]. Dans cette section nous allons présenter parmi ces algorithmes ceux qui sont relativement exploitables.

D'une manière générale, il existe trois approches pour la résolution d'un PBL.

Approche basées sur la reformulation KKT

C'est l'approche qui a donné naissance au plus grand nombre de méthodes pour la résolution des problèmes PBL. Le principe est de transformer le PBL en un problème à un seul niveau en utilisant les conditions KKT suivant la proposition suivante :

Proposition 2.11. [63] *Une condition nécessaire pour que (x^*, y^*) soit solution optimale du PBL (2.13) est qu'il existe deux vecteurs u^* et v^* tels que (x^*, y^*, u^*, v^*) soit solution de :*

$$\begin{cases} \min c_1x + d_1y, \\ A_1x + B_1y \leq b_1, \\ uA_2 - v = -c_2 \\ u(b_2 - A_2x - B_2y) + vx = 0, \\ A_2x + B_2y \leq b_2, \\ x \geq 0, y \geq 0, u \geq 0, v \geq 0. \end{cases} \quad (2.20)$$

Pour résoudre le problème PBL, le principe consiste donc à résoudre directement le programme mathématique équivalent (2.20). En pratique, la résolution n'est pas vraiment aussi simple, car la contrainte dite de *complémentarité* dans (2.20) n'est pas linéaire. Il s'agit de la contrainte

$$u(b_2 - A_2x - B_2y) + vx = 0.$$

Plusieurs méthodes sont développées dans cette approche, en cherchant à chaque fois un *compromis* avec la contrainte de complémentarité. Nous allons donner un aperçu sur quelques méthodes. Pour une lecture détaillée, le lecteur peut consulter [63, 28, 9, 33, 35].

- **La méthode de *Branch and Bound*** : C'est la méthode qui connaît le plus de succès et on l'applique pour les problèmes bi-niveaux convexes. Des algorithmes basés sur cette idée sont développés par Falk en 1982 [10], leur technique consiste à entourer l'ensemble des solutions admissibles du problème non convexe (2.20) par un polyèdre et qu'on subdivise ensuite en sous-ensembles disjoints. On minimise la fonction objectif de (2.20) sur chaque sous-ensemble, puis on sélectionne la plus petite valeur parmi toutes les valeurs obtenues. Après un test de cette valeur, on la prend comme valeur optimale globale si le test est positif, sinon on subdivise le sous-ensemble correspondant à cette valeur en nouveaux polyèdres et on recommence l'opération.

Fortuny-Amat et Mc Carl (1984) proposent une méthode similaire à la précédente, plus tard, Bard et Moore présentent un algorithme plus efficace. Leur idée de base est de supprimer la contrainte de complémentarité de (2.20) et de résoudre le sous-problème linéaire résultant. A chaque itération, on vérifie si la contrainte de complémentarité est satisfaite. Si oui, le point correspondant se trouve dans le domaine induit,

d'où il présente une solution potentielle de PBL, sinon on utilise un procédé de branch and bound pour examiner toutes les combinaisons des contraintes de complémentarité *ralaxées* [63]. On trouvera dans [61] un algorithme dit *prolongé* utilisant la méthode de brunch and bound.

- **Méthode des fonctions de pénalité :** L'idée des algorithmes basés sur cette méthode, proposée d'abord par Bard et Falk, est de placer la contrainte de complémentarité dans la fonction objectif du problème (2.20) pondérée par une constante positive K suffisamment grande. L'idée est reprise par Anandalingram et Wild [2], en utilisant une fonction de pénalité qui est l'écart entre les valeurs des fonctions objectifs primale et duale du suiveur. Dans sa première version l'algorithme ne permettait de donner qu'une solution optimale locale, par la suite il a été amélioré et permet de donner un solution globale. Différents algorithmes utilisant une *méthode de pénalité extérieure* sont proposés par Shimizu et Lu [64] qui exigent simplement la convexité des fonctions objectifs et des fonctions contraintes pour garantir la convergence vers une solution optimale globale. Citons que Campêlo et Scheimberg [22] proposent, via des fonctions de pénalité, un algorithme pour la recherche d'une solution optimale locale. Le lecteur est renvoyé à [63, 33, 2] au sujet de cette méthode.
- **Méthode du pivot complémentaire :** L'idée de l'algorithme utilisant cette méthode revient à Balias et Karwan [14], en écrivant le problème (2.20) comme un *problème du minimum linéaire de complémentarité* MLPC ¹,

$$\left\{ \begin{array}{l} \min c_1x + d_1y, \\ w = b_2 - A_2x - B_2y, \\ v = c_2 + uA_2, \\ x \geq 0, y \geq 0, u \geq 0, v \geq 0, w \geq 0, uw = vx = 0. \end{array} \right. \quad (2.21)$$

Ils utilisent une base d'entrée connue sous le nom de *pivot complémentaire* pour trouver une solution au problème MLPC (2.21). Cette méthode peut être, néanmoins, considérée comme une heuristique puisqu'elle ne converge pas toujours vers la solution optimale. Par la suite Júdice et Faustino [45] mettent quelques modifications pour garantir la convergence globale.

¹En anglais : minimum linear complementarity problem

Approche basée sur les points extrêmes

Cette approche s'appuie sur le théorème 2.8 qui énonce que la solution optimale d'un PBL est atteinte à un point extrême du domaine des contraintes, et sur des résultats connus de la programmation linéaire. Les méthodes basées sur cette approche procèdent par énumération des points extrêmes d'une manière analogue au procédé utilisé avec la méthode du simplexe.

Candley et Townsley sont les premiers à proposer une méthode utilisant cette approche (1982). Leur méthode permet de trouver une solution optimale globale pour les problèmes PBL sans contraintes du niveau supérieur et avec une unique solution du problème du niveau inférieur pour tout paramètre x . Le procédé itératif qu'utilise cette méthode opère de la manière suivante : Il résout deux programmes linéaires, le premier pour le leader avec la variable x et un sous-ensemble de variables y associées à une base optimale du problème du suiveur. Le deuxième problème pour le suiveur en fixant les variables x . Systématiquement, on explore les bases optimales du problème du suiveur pour x fixé, puis on revient au problème du leader avec les bases correspondantes aux variables x , en diminuant le nombre de bases à explorer. Toutefois, le constat est fait, avec des tests numériques, que cette méthode est relativement lente [50]. Voir [21] pour plus de détails.

Toujours en se basant sur l'approche des points extrêmes, Balias et Karwan (1984), proposèrent une méthode appelée *K^{ième} meilleur algorithme*. La méthode permet de trouver une solution optimale globale sous la condition que le domaine induit DI soit borné et l'ensemble des réactions rationnelles $M(x)$ soit un singleton pour tout y [15].

Approche basée sur le gradient

Les algorithmes utilisant cette approche se basent sur le calcul de dérivées directionnelles et de sous-gradients de la fonction objectif du leader F en fonction d'informations tirées du gradient des fonctions du problème du suiveur. Pour ces algorithmes, les hypothèses que $M(x)$ soit uniquement déterminée localement pour tout x , et que $M(x)$ soit continûment différentiables sont exigées. En outre d'autres conditions de régularité sont nécessaires. La nature relativement restrictive de toutes ces hypothèses font que ces méthodes ne sont pas généralement commodes pour la résolution des problèmes linéaires à deux niveaux [63].

2.2.6 Relation PBL- Problèmes classiques

Dans cette section, nous montrons que de simples transformations permettent de formuler certains problèmes classiques comme des programmes bi-niveaux et vice versa. L'intérêt de ces reformulations repose sur l'étude de la complexité. Tout algorithme développé pour résoudre l'un peut être adapté à la résolution de l'autre et vice versa.

PBL et MIP₀₋₁

La formulation du problème de programmation mixte (MIP₀₋₁)² est la suivante :

$$\begin{cases} \max_{x,u} cx + eu, \\ Ax + Eu \leq b, \\ x \geq 0, u \in \{0, 1\}^n, \end{cases} \quad (2.22)$$

où $c \in \mathbb{R}^p$, $e \in \mathbb{R}^n$, A une $m \times n$ -matrice, $b \in \mathbb{R}^m$. Il faudrait noter que la contrainte binaire est équivalente à

$$\begin{cases} 0 \leq u \leq \mathbf{1}, \\ 0 = \min\{u, \mathbf{1}\}, \end{cases}$$

où $\mathbf{1}$ désigne le vecteur de composantes égales à 1.

En introduisant une variable du niveau supérieur y , et en définissant un problème du niveau inférieur tel que la solution optimale correspond à ce minimum, nous obtenons la formulation équivalente à un problème bi-niveaux,

$$\begin{cases} \begin{cases} \min_{x,y,u} cx + eu, \\ Ax + Eu \leq b, \\ 0 \leq u \leq \mathbf{1}, \\ x \geq 0, \\ y = 0, \end{cases} \\ \begin{cases} y \in \arg \max_w \sum_{i=1}^{n_u} w_i, \\ w \leq u, \\ w \leq \mathbf{1} - u, \end{cases} \end{cases} \quad (2.23)$$

où $y, w \in \mathbb{R}^{n_u}$. Dans cette formulation, les contraintes d'intégrité ne sont pas nécessaires; en effet elles sont satisfaites par le biais des contraintes du

²En anglais : Mixed 0-1 programming problem

niveau supérieur $y = 0$ combinées aux conditions d'optimalité du problème du niveau inférieur.

En général les contraintes du niveau supérieur rendent le problème plus difficile à résoudre. Quelques algorithmes seulement sont destinés à la résolution des instances pour lesquelles de telles contraintes sont absentes. Cependant, comme suggéré par Vicente et al. [68], la contrainte $y = 0$ peut être satisfaite en introduisant une pénalité exacte à la fonction objectif du niveau supérieur, i.e. il existe une valeur M^* telle que toute valeur M supérieure à M^* , la solution du programme bi-niveaux ci-après satisfait la contrainte $y = 0$.

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \min_{x,y,u} cx + eu + M\mathbf{1}y, \\ Ax + Eu \leq b, \\ 0 \leq u \leq \mathbf{1}, \\ x \geq 0, \end{array} \right. \\ \left\{ \begin{array}{l} y \in \arg \max_w \sum_{i=1}^{n_u} w_i, \\ w \leq u, \\ w \leq \mathbf{1} - u. \end{array} \right. \end{array} \right. \quad (2.24)$$

Inversement, il existe une réduction polynômiale de PBL en MIP_{0-1} . Remplaçons tout d'abord le problème du niveau inférieur par ses conditions d'optimalité, conduisant à un problème à un seul niveau, comportant les contraintes de complémentarité

$$\left\{ \begin{array}{l} \lambda(b_2 - A_2x - B_2y) = 0, \\ (\lambda B_2 - d_2)y = 0. \end{array} \right.$$

Pour la deuxième transformation, on linéarise les contraintes de complémentarité en introduisant deux vecteurs binaires u et v et un coefficient $L > 0$ suffisamment grand dont l'existence a été établi dans [68].

$$\left\{ \begin{array}{l} b_2 - A_2x - B_2y \leq L(\mathbf{1} - u), \quad \lambda \leq Lu, \\ y \leq L(\mathbf{1} - v), \quad \lambda B_2 - d_2 \leq Lv, \end{array} \right.$$

ce qui conduit à la formulation du MIP_{0-1} équivalente au problème PBL, de

la manière suivante :

$$\left\{ \begin{array}{l} \max_{x,y,u,v,\lambda} c_1x + d_1y, \\ A_1x + B_1y \leq b_1, \\ A_2x + B_2y \leq b_2, \quad -\lambda B_2 \leq -d_2, \\ y \geq 0, \quad \lambda \geq 0, \\ -A_2x - B_2y + Lu \leq L\mathbf{1} - b_2, \quad \lambda - Lu \leq 0, \\ y + Lu \leq L\mathbf{1}, \quad \lambda B_2 - Lv \leq d_2, \\ u \in \{0, 1\}^n, \quad v \in \{0, 1\}^n. \end{array} \right. \quad (2.25)$$

PBL et BILP

Le concept de programmation bilinéaire disjointe a été introduit par Konno [47] en 1971 pour le calcul de l'équilibre de Nash des jeux bi-matriciels. La formulation d'un problème de programmation bilinéaire disjointe (BILP)³ est la suivante :

$$\left\{ \begin{array}{l} \max_{x,u} cx - uQx + ud, \\ Ax \leq b_1, \\ uB \leq b_2, \\ x \geq 0, \quad u \geq 0, \end{array} \right. \quad (2.26)$$

où $c \in \mathbb{R}^{n_x}$, $d \in \mathbb{R}^{n_u}$, Q une $n_u \times n_x$ -matrice, A une $n_v \times n_x$ -matrice, B une $n_u \times n_y$ -matrice, $b_1 \in \mathbb{R}^{n_v}$, $b_2 \in \mathbb{R}^{n_y}$. En exploitant la relation entre un problème PBL et un problème BILP, Audet et al. [5, 4] et Alarie et al. [1] ont été en mesure de concevoir des algorithmes performants de résolution de PBL basée sur une méthode de *coupes et séparation*⁴. Leur approche repose sur la séparabilité des vecteurs x et u de l'ensemble des solutions réalisables de BILP.

Considérons les deux ensembles

$$X = \{x \geq 0 : Ax \leq b_1\} \text{ et } U = \{u \geq 0 : uB \leq Bb_2\}.$$

Si on suppose que ces ensembles sont tout les deux non vides et la solution optimale de BILP est bornée, il est possible de reformuler le BILP comme suit :

$$\max_{x \in X} cx + \max_{u \in U} u(d - Qx). \quad (2.27)$$

³En anglais : Disjoint bilinear programming problem

⁴En anglais : brunch and cut

Pour $x \in X$ fixé, on peut remplacer le second niveau par son dual, ce qui conduit à :

$$\begin{cases} \max_{x \in X} cx + \min_y b_2y, \\ Qx + By \geq d, \\ y \geq 0. \end{cases} \quad (2.28)$$

Sous les hypothèses ci-dessus, le dual du second niveau est borné et admet une solution pour tout $x \in X$. De manière symétrique, on pourrait permuter les rôles de x et u pour obtenir la formulation équivalente suivante :

$$\begin{cases} \max_{u \in U} cx + \min_v vb_1, \\ uQ + vA \geq c, \\ v \geq 0. \end{cases} \quad (2.29)$$

Alors la solution de BILP peut être obtenue en résolvant l'un des deux programmes bi-niveaux symétriques suivant :

$$\begin{cases} \begin{cases} \max_{x,y} cx + b_2y, \\ Ax \leq b_1, \\ x \geq 0, \end{cases} \\ \begin{cases} y \in \arg \min_y b_2y, \\ Qx + By \geq d, \\ y \geq 0. \end{cases} \end{cases} \quad (2.30)$$

$$\begin{cases} \begin{cases} \max_{u,v} ud + vb_1, \\ uB \leq b_2, \\ u \geq 0, \end{cases} \\ \begin{cases} v \in \arg \min_v vb_1, \\ uQ + vA \geq c, \\ v \geq 0. \end{cases} \end{cases} \quad (2.31)$$

Si le BILP est non borné les transformations ci-dessus ne sont plus valables. Audet et al. [5] ont montré que déterminer s'il existe $x \in X$ tel que l'ensemble

$$Y(x) = \{y \geq 0 : By \geq d - Qx\}$$

est vide est fortement NP-complet, de même, déterminer si le BILP est borné est fortement NP-complet.

PBL et Problème Max-Min

D'une manière générale, la formulation du problème Max-Min linéaire est la suivante :

$$\max_x \{ \min_y cx + dy : Ax + By \leq b, x \geq 0, y \geq 0 \}, \quad (2.32)$$

qui est clairement un cas spécial d'un problème PBL, où le problème du niveau inférieur est l'opposé du problème du niveau supérieur. Le problème Min-Max linéaire précédent peut être formulé d'une manière équivalente comme un problème de programmation bi-niveaux linéaire suivant :

$$\left\{ \begin{array}{l} \min_x -cx - dy, \\ \left\{ \begin{array}{l} \min_y cx + dy, \\ Ax + By \leq b, \\ x \geq 0, y \geq 0. \end{array} \right. \end{array} \right. \quad (2.33)$$

Chapitre 3

Contrôle optimal bi-niveaux

Ce chapitre est composé de deux parties : La première partie est consacrée à la présentation des concepts de bases du problème de contrôle. Dans la seconde partie on a fait une étude des problèmes de contrôle optimal bi-niveaux.

3.1 Contrôle Optimal

La théorie du contrôle analyse les propriétés des systèmes commandés, c'est-à-dire des systèmes dynamiques sur lesquels on peut agir au moyen d'un contrôle (commande). Le but alors, est d'amener le système d'un état initial donné à un certain état final prescrit, en respectant éventuellement certains critères.

3.1.1 Position du problème

Le problème général de contrôle optimal est constitué des données suivantes :

Objet de la commande

Il est décrit par l'équation différentielle :

$$\dot{x} = f(t, x(t), u(t)), \quad x(t_0) = x_0,$$

où $x(t) \in \mathbb{R}$ est la position de l'objet à l'instant t , $x(t_0) \in \mathbb{R}$ la position initiale du système et $u(t)$ la commande.

Le but de la commande

Dans un problème de contrôle, le but de la commande consiste à ramener l'objet considéré de la position initiale $x_0 = x(t_0)$, ($x_0 \in M_0$) à la position finale $x_f = x(t_f)$ ($x_f \in M$), où M_0 est l'ensemble de départ et M l'ensemble d'arrivée (accessibilité).

La classe des commandes admissibles

La classe des commandes admissibles U est constituée de fonctions mesurables $u(t)$:

$$U = \{u(t), t \in T = [t_0, t_f]\}.$$

Chaque commande transfère l'objet du point de départ x_0 en un point de l'ensemble d'arrivée M .

Commande Bang-bang

Soit U l'ensemble des commandes et U_e l'ensemble des points extrémaux de U .

La commande admissible u est dite de Bang-bang si $u \in U_e$, c'est-à-dire u prend les valeurs extrémales de U .

Critère de qualité

Le critère de qualité appelé aussi coût ou fonction objectif, est généralement décrit par la formule :

$$J(x, u) = g(t_f, x_f) + \int_{t_0}^{t_f} f^0(t, x(t), u(t))dt.$$

Cette fonctionnelle comporte deux parties : $g(t_f, x_f)$ est le coût terminal. C'est une sorte de pénalité liée à la fin de l'évolution du système au temps final t_f ; il a son importance lorsque t_f est libre, sinon il est constant. Le second terme intervenant dans la fonction objectif $\int_{t_0}^{t_f} f^0(t, x(t), u(t))dt$ dépend de l'état du système tout au long de la trajectoire de la solution, définie par les variables de contrôle u . C'est une fonction d'efficacité de chaque commande sur l'intervalle T .

Ensemble Accessible

Considérons le système dynamique suivant :

$$\dot{x} = f(t, x(t), u(t)) \quad x(0) = x_0, \quad t \in T. \quad (3.1)$$

Définition 3.1. L'ensemble des points accessibles à partir de x_0 en temps T est :

$$Acc(x_0, T) = \{x_u(T), u \in U\},$$

où $x_u(\cdot)$ est la solution du système (3.1) associée au contrôle u .

$Acc(x_0, T)$ est l'ensemble des extrémités des solutions du système (3.1), en temps T lorsque le contrôle u varie.

3.2 Contrôlabilité

La commandabilité du système consiste à trouver une commande qui permet à la trajectoire de passer d'un état initial x_0 à un état final x_f en un temps fini.

Pour ce problème de contrôlabilité, Kalman a donné dès 1949 une caractérisation des systèmes linéaires autonomes, contrôlable en dimension finie. Concernant les systèmes non linéaires, le problème mathématique de contrôlabilité est beaucoup plus compliqué. Il constitue un domaine de recherche actif.

Définition 3.2. Un processus défini par l'équation :

$$\dot{x} = f(t, x, u)$$

est dit *non autonome*. Et il est dit *autonome* s'il ne dépend pas de t i.e :

$$\dot{x} = f(x, u).$$

3.2.1 Contrôlabilité des systèmes linéaires

La formulation mathématique d'un système de contrôle linéaire est la suivante :

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + r(t), \quad x(t_0) = x_0, \quad t \in T, \quad (3.2)$$

où T est un intervalle de \mathbb{R} , A , B et r sont trois applications localement intégrables sur T à valeurs respectivement dans $M_n(\mathbb{R})$, $M_{n,m}(\mathbb{R})$ et \mathbb{R}^m . $M_n(\mathbb{R})$ est l'ensemble des matrices réelles de dimension n , et $M_{n,m}(\mathbb{R})$ est l'ensemble des matrices de n lignes et m colonnes.

Soit $F(\cdot) : T \rightarrow M_n(\mathbb{R})$ la résolvante du système linéaire homogène :

$$\dot{x}(t) = A(t)x(t)$$

définie par :

$$\begin{cases} \dot{F} = A(t)F(t) \\ F(t_0) = Id, \end{cases}$$

où Id : désigne la matrice identité.

Pour tout contrôle u le système (3.2) admet une unique solution $x(\cdot) : I \rightarrow \mathbb{R}^n$ absolument continue donnée par :

$$x(t) = F(t)x_0 + \int_{t_0}^{t^*} F(t)F(s)^{-1}(B(s)u(s) + r(s))ds,$$

pour tout $t \in T$.

Le théorème suivant donne une condition générale pour la contrôlabilité des systèmes linéaires.

Théorème 3.1. [67] *Le système (3.2) est dit contrôlable en temps t^* si et seulement si la matrice*

$$C(t^*) = \int_{t_0}^{t^*} F(t)^{-1}B(t)B(t)^T(F(t)^{-1})^T dt$$

est inversible.

$C(t^*)$ est appelée matrice de contrôlabilité.

Cette condition ne dépend pas de x_0 , c'est-à-dire que si un système linéaire est contrôlable en temps t^* depuis x_0 , alors il est contrôlable en temps t^* depuis tout point.

3.2.2 Contrôlabilité des systèmes linéaires autonomes

Soit le système linéaire autonome suivant :

$$\dot{x}(t) = Ax(t) + Bu(t) + r(t), \quad x(t_0) = x_0. \quad (3.3)$$

Dans ce cas, la matrice $F(t) = e^{At}$ est la solution du système associée au contrôle u s'écrit, pour tout $t \in T$:

$$x(t) = e^{At}(x_0 + \int_{t_0}^{t^*} e^{-As}(B(s)u(s) + r(s))ds).$$

Le théorème suivant donne une condition nécessaire et suffisante de commandabilité du système (3.3).

Théorème 3.2. [67] *Le système $\dot{x}(t) = Ax(t) + Bu(t) + r(t)$ est contrôlable en temps t^* si et seulement si la matrice*

$$C = [B, AB, A^2B, \dots, A^{n-1}B]$$

est de rang n .

La matrice C est appelée *matrice de Kalman*, et la condition $\text{rang}C = n$ est appelée *condition de Kalman*.

Remarque 3.1. La condition de Kalman ne dépend ni de t^* ni de x_0 . Autrement dit, si un système linéaire autonome est contrôlable en temps t^* depuis x_0 , alors il est contrôlable en tout temps depuis tout point.

Exemple 3.1. Soit le système :

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u. \end{cases}$$

Dans ce cas :

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Nous sommes dans une situation où $n = 2$ états et $m = 1$ entrée. La matrice de commandabilité associée :

$$C = \begin{pmatrix} B & AB \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Nous avons le rang $C = 2$, donc le système est commandable.

Exemple 3.2. Considérons maintenant ce système :

$$\begin{cases} \dot{x}_1 = u \\ \dot{x}_2 = u \end{cases}$$

On a $n = 2$ états et $m = 1$ entrée. Dans ce cas :

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

La matrice de commandabilité du système est :

$$C = \begin{pmatrix} B & AB \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

$\text{Rang}C = 1 \neq 2$ donc le système n'est pas contrôlable.

Note : $\dot{x}_1 - \dot{x}_2 = 0$, soit $x_1 - x_2 = \text{constante}$, donc c'est une relation indépendante de u , alors pas de commandabilité.

3.3 Stabilisation

Un *contrôle en boucle ouverte* est une application $t \rightarrow u(t)$ d'un intervalle de temps dans l'espace des contrôles. Un *contrôle en boucle fermée*, appelé aussi un *bouclage*, ou une *rétroaction* (ou en anglais *feed back*), est une application $u \rightarrow g(t)$ définie sur les variables d'état du système. Un des objectifs de la théorie du contrôle est de déterminer des rétroactions qui stabilisent le système en un état particulier.

3.3.1 Bouclage statique

Définition 3.3. On dit que u est un *bouclage statique* du système $\dot{x} = f(x(t), u(t))$, ($x(t) \in M, u(\cdot) \in U$) si sa valeur $u(t)$ à l'instant t ne dépend que de $x(t)$, c'est-à-dire $u = g(x)$ où g est une fonction.

Ce système s'écrit tout simplement

$$\dot{x} = f(x, g(x)). \tag{3.4}$$

Le problème de la *stabilisation* (ou *régulation*) consiste à maintenir le système près d'un équilibre x^* .

Il s'agit donc de construire une loi de commande telle que x^* soit un équilibre asymptotiquement stable du système en boucle fermée (3.4).

3.3.2 concepts de stabilité

On se donne un système

$$\dot{x}(t) = f(x) \tag{3.5}$$

tel que $f(0) = 0$, admettant $x = 0$ comme équilibre (noter que par un changement de variable on peut toujours ramener l'équilibre à l'origine).

Définition 3.4. L'équilibre $x = 0$ du système (3.4) est dit *stable* si pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que pour toute solution $x(t)$ de (3.4), on ait :

$$\|x(0)\| < \eta \Rightarrow \forall t \geq 0, \|x(t)\| < \epsilon.$$

Si l'équilibre n'est pas stable, on dit qu'il est *instable*.

Définition 3.5. L'équilibre $x = 0$ du système (3.5) est dit *attractif* s'il existe $r > 0$ tel que pour toute solution $x(t)$ de (3.5) on ait :

$$\|x(t)\| < r \Rightarrow \lim_{t \rightarrow \infty} x(t) = 0.$$

L'équilibre $x = 0$ du système (3.5) est dit *globalement attractif* si pour toute solution $x(t)$ de (3.5) on a $\lim_{t \rightarrow \infty} x(t) = 0$.

L'ensemble \mathcal{B} défini par la propriété :

$$x(0) \in \mathcal{B} \Rightarrow \lim_{t \rightarrow \infty} x(t) = 0,$$

s'appelle *le bassin d'attraction* de l'origine. Ainsi $x = 0$ est attractif si \mathcal{B} est un voisinage de 0. Il est globalement attractif si $\mathcal{B} = \mathbb{R}^n$.

Définition 3.6. L'équilibre $x = 0$ du système (3.5) est dit *asymptotiquement stable* s'il est stable et attractif. Il est dit *globalement asymptotiquement stable* s'il est stable et globalement attractif.

Définition 3.7. L'équilibre $x = 0$ du système (3.5) est dit *exponentiellement stable* s'il existe $r > 0$, $M > 0$ et $\alpha > 0$ tels que pour toute solution $x(t)$ on ait :

$$\|x(t)\| < r \Rightarrow \|x(t)\| \leq M\|x(0)\|e^{-\alpha t}, \text{ pour tout } t > 0.$$

L'équilibre $x = 0$ du système (3.5) est dit *globalement exponentiellement stable* s'il existe $M > 0$ et $\alpha > 0$ tels que pour toute solution $x(t)$ de (3.5) on a :

$$x(t) \leq M\|x(0)\|e^{-\alpha t}, \text{ pour tout } t > 0.$$

3.4 Principe du maximum de Pontryaguin

La formulation d'un problème de contrôle optimal est la suivante :

$$\begin{cases} \min_u J(x, u) = g(t_f, x(t_f)) + \int_0^{t_f} f^0(t, x(t), u(t))dt, \\ \dot{x} = f(t, x(t), u(t)), \\ x(0) = x_0 \in M_0, \\ x(t_f) = x_f \in M, \\ u \in U, t \in T = [0, t_f]. \end{cases} \quad (3.6)$$

Avant d'annoncer le principe du maximum, introduisons certaines définitions et propriétés essentielles.

Définition 3.8. Le contrôle u est dit *extrémal* sur $[0, t]$ si la trajectoire du système $\dot{x} = f(t, x(t), u(t))$ du problème de contrôle (3.6) associée à u vérifie :

$$x(t) \in \partial Acc(x_0, t), \quad t \in T = [0, t_f].$$

Définition 3.9. Un contrôle $u^*(t), t \in [0, t_f]$ est dit *optimal* si $u^*(.)$ est extrémal et $J(u^*(t)) < J(u(t))$ pour tout contrôle extrémal $u(t), t \in [0, t_f]$.

Théorème 3.3. [67] *Considérons le système :*

$$\forall t \in T, \quad \dot{x} = A(t)x(t) + B(t)u(t), \quad x(0) = x_0.$$

Supposons que le domaine des contraintes noté Ω est compact. Soit $t_f > 0$. Le contrôle u est extrémal sur $T = [0, t_f]$ si et seulement s'il existe une solution non triviale $p(t), t \in T$, de l'équation $\dot{p}(t) = -p(t)A(t)$ telle que :

$$p(t)B(t)u(t) = \max_{v \in \Omega} p(t)B(t)v \quad (3.7)$$

pour presque tout $t \in [0, t_f]$.

Définition 3.10. Le vecteur $p(t) \in \mathbb{R}^n$ est appelé *vecteur adjoint*.

Remarque 3.2. si $\Omega = [-a, a], a \in \mathbb{R}$ positif, la condition (3.7) signifie que

$$u(t) = a \operatorname{signe}(p(t)B(t)).$$

On obtient une commande bang-bang. Dans ce cas, la fonction $p(t)B(t)$ est appelée fonction de commutation.

Théorème 3.4. [67] *Ce théorème est l'énoncé général du principe du maximum de Pontryaguin.*

Considérons le système de contrôle dans \mathbb{R}^n

$$\dot{x} = f(t, x(t), u(t)) \quad (3.8)$$

où $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ de classe C^1 , les contrôles sont des applications mesurables bornées à valeur dans $\Omega \subset \mathbb{R}^m$. Soient M_0 et M deux sous-ensembles de \mathbb{R}^n . Notons par U l'ensemble des contrôles admissibles u dont les trajectoires associées relient un point initial de M_0 à un point final de M en temps t .

On définit le coût

$$J(x, u) = g(t_f, x(t_f)) + \int_0^{t_f} f^0(t, x(t), u(t)) dt,$$

où $f^0 : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ et $g : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ sont de classe C^1 , $x(\cdot)$ est la solution de (3.8) associée au contrôle u .

On considère le problème de contrôle optimal suivant : déterminer une trajectoire reliant M_0 à M en minimisant le coût J . Le temps final peut être fixé ou non.

Si le contrôle $u \in U$ associé à la trajectoire $x(\cdot)$ est optimal sur $[0, t_f]$, alors il existe une application $p(\cdot) : [0, t_f] \rightarrow \mathbb{R}^n$ absolument continue, appelée vecteur adjoint, et un réel $p^0 \leq 0$ tel que le couple $(p(\cdot), p^0)$ est non trivial et tel que pour presque tout $t \in [0, t_f]$:

$$\begin{aligned} \dot{x}(t) &= \frac{\partial H}{\partial p}(t, x(t), p(t), p^0, u(t)), \\ \dot{p}(t) &= \frac{-\partial H}{\partial x}(t, x(t), p(t), p^0, u(t)), \end{aligned}$$

où

$$H(t, x, p, p^0, u) = p'(t)f(t, x, u) + p^0 f^0(t, x, u)$$

est le Hamiltonien du système, et on a la condition de maximisation presque partout sur $[0, t_f]$:

$$H(t, x(t), p(t), p^0, u(t)) = \max_{v \in U} H(t, x(t), p(t), p^0, v).$$

Si de plus le temps final pour joindre M n'est pas fixé, on a la condition au temps final t_f

$$\max_{v \in U} H(t, x(t), p(t), p^0, v) = -p^0 \frac{\partial g}{\partial t}(t_f, x(t_f)).$$

Si de plus M_0 et M (ou juste l'un des deux ensembles) sont des variétés de \mathbb{R}^n ayant des espaces tangents en $x(0) = x_0 \in M_0$ et $x(t_f) = x_f \in M$, alors le vecteur adjoint peut être construit de manière à vérifier les conditions de transversalités aux deux extrémités (ou juste l'un des deux)

$$p(0) \perp (t_f)_{x(0)} M_0,$$

$$p(t_f) - p^0 \frac{\partial g}{\partial t}(t_f, x(t_f)) \perp (t_f)_{x(t_f)} M.$$

3.5 Eléments d'analyse

Nous allons présenter quelques notions d'analyse qui seront utiles pour la suite. L'étude que nous présentons dans ce document requiert des outils d'analyse non régulière (analyse non lisse)¹, i.e. en considérant des fonctions non nécessairement différentiables au sens usuel. Voir [24, 25, 26, 60] à propos de ce sujet.

Commençons par définir les fonctions lipschitziennes, qui forment une classe de fonctions (non nécessairement différentiables) très utilisée avec d'importantes propriétés.

Définition 3.11. Soit la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, et x un point de \mathbb{R}^n . La fonction f est dite *lipschitzienne* au voisinage de x s'il existe un scalaire K tel que, pour tous points y et z appartenant au voisinage de x , on ait :

$$|f(y) - f(z)| \leq K \|y - z\|.$$

Le plus petit scalaire K pour qui cette inégalité est réalisée dans un ensemble donné est appelé le *rang lipschitzien* de f relatif à cet ensemble.

¹En anglais : Nonsmooth analysis

3.5.1 La dérivée directionnelle généralisée

Considérons la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ lipschitzienne au voisinage de x , et soit v un vecteur quelconque de \mathbb{R}^n . La *dérivée directionnelle généralisée* de f en x dans la direction de v , notée $f^\circ(x; v)$ est définie par

$$f^\circ(x; v) = \limsup_{\substack{y \rightarrow x \\ t \rightarrow 0}} \frac{f(y + tv) - f(y)}{t}$$

où y est un vecteur de \mathbb{R}^n et t un scalaire positif.

L'utilité de la dérivée directionnelle généralisée (appelée aussi *dérivée directionnelle au sens de Clarke*) apparaît dans la proposition suivante.

Proposition 3.5. [25] *Soit f une fonction lipschitzienne de rang K au voisinage de x . Alors*

i. La fonction $v \rightarrow f^\circ(x; v)$ est homogène positivement, i.e.

$$f^\circ(x; \lambda v) = \lambda f^\circ(x; v) \quad \forall \lambda > 0,$$

elle est sous-additive, i.e.

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

et vérifie

$$|f^\circ(x; v)| \leq K \|v\|.$$

ii. La fonction $(x, v) \rightarrow f^\circ(x; v)$ est semi-continue supérieurement, et la fonction $v \rightarrow f^\circ(x; v)$ est lipschitzienne de rang K sur \mathbb{R}^n .

iii. $f^\circ(x; -v) = (-f)^\circ(x; v)$.

D'après le théorème de Hahn-Banach [18], et sous les conditions de la proposition 3.5, il existe au moins une forme linéaire $\zeta : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que $f^\circ(x; w) \geq \zeta(w)$ pour tout $w \in \mathbb{R}^n$. Il s'ensuit aussi que ζ est bornée. Notons par $\mathfrak{L}(\mathbb{R}^n)$ l'ensemble des formes linéaires sur \mathbb{R}^n (i.e. le dual de \mathbb{R}^n), et notons par $\|\zeta\|_*$ la norme sur $\mathfrak{L}(\mathbb{R}^n)$.

$$\|\zeta\|_* = \sup\{|\zeta(v)|, v \in \mathbb{R}^n, \|v\| \leq 1\}.$$

Définition 3.12. On appelle le *gradient généralisé* de f au point x , qu'on note $\partial f(x)$, le sous-ensemble de $\mathfrak{L}(\mathbb{R}^n)$ donné par

$$\partial f(x) = \{\zeta \in \mathfrak{L}(\mathbb{R}^n) : f^\circ(x; v) \geq \zeta(v) \quad \text{pour tout } v \in \mathbb{R}^n\}.$$

Pour une fonction convexe non partout différentiable, la notion de sous-gradient généralise celle du gradient.

Définition 3.13. (Sous-différentiabilité) On appelle *sous-gradient* de la fonction convexe f au point x^0 tout vecteur γ vérifiant :

$$f(x) \geq f(x^0) + \gamma^T(x - x^0). \quad (3.9)$$

On appelle *sous-différentiel* de f au point x^0 l'ensemble de tous les sous-gradients de f en x^0 .

L'hyperplan d'équation $z = f(x^0) + \gamma^T(x - x^0)$ qui rencontre l'épigraphe de la fonction f au point $(x^0, f(x^0))$ est appelé un *hyperplan d'appui* de l'épigraphe de f au point x^0 .

Proposition 3.6. [25] *Si f est convexe sur un ensemble convexe $U \subset \mathbb{R}^n$, et lipschitzienne au voisinage de x , alors le gradient généralisé de f en x , $\partial f(x)$, coïncide avec le sous-différentiel de f en x au sens de la définition (3.13).*

Comme il est établi que les fonctions convexes sont lipschitziennes [25], nous allons adopter la même notation $\partial f(x)$ pour le sous-gradient d'une fonction convexe f au point x .

Remarque 3.3. Si f est différentiable en x^0 , le sous-différentiel en x^0 est réduit à un seul point, le gradient de f en x^0 , i.e. $\partial f(x^0) = \{\nabla f(x^0)\}$.

Définition 3.14. Une fonction $x : [0, T] \rightarrow \mathbb{R}^n$ est *absolument continue* si pour une certaine fonction v intégrable au sens de Lebesgue, x peut s'exprimer comme $x(t) = \int_0^t v(\tau) d\tau$. Il s'ensuit que $\dot{x} = v$ presque partout. On appellera *arc* toute fonction absolument continue.

3.5.2 Cône normal de Clarke

En analyse convexe, on appelle *cône* un sous-ensemble $K \subset \mathbb{R}^n$ stable pour la multiplication par un scalaire positif, i.e. $\lambda x \in K$ quand $x \in K$ et $\lambda > 0$.

Soit C un sous-ensemble non vide de \mathbb{R}^n . Nous allons considérer une fonction relative à C qui est lipschitzienne sur \mathbb{R}^n , appelée *fonction distance*, notée d_C , qui est définie par :

$$d_C(x) = \inf\{\|x - c\| : c \in C\}.$$

On dit que $w \in \mathbb{R}^n$ est *perpendiculaire* à C au point c si :

$$d_C(c + w) = \|w\| > 0.$$

Définition 3.15. Soit $C \subset \mathbb{R}^n$ et $x \in C$. Un vecteur v de \mathbb{R}^n est dit *tangent* à C en x si $d_C^\circ(x; v) = 0$. L'ensemble de tous les points tangents à C en x est noté $T_C(x)$ et est appelé *cône tangent* à C au point x .

$$T_C(x) = \{v \in \mathbb{R}^n : d_C^\circ(x; v) = 0\}.$$

Un point est dit *proximal normal* de l'ensemble C en $c \in C$ s'il est un multiple non négatif d'un point perpendiculaire à C en c . D'une manière équivalente : $\xi \in \mathbb{R}^n$ est proximal normal de C au point c si pour tout $t > 0$, le point de C le plus proche (en norme euclidienne) de $c + t\xi$ est c . Notons $\Pi_C(c)$ l'ensemble des points qui sont proximaux normaux de C en c .

Définition 3.16. Soit $c \in C$. Le *cône proximal normal* de C en c est l'ensemble

$$\hat{N}_C(c) = \left\{ \lim_{i \rightarrow \infty} \xi_i, \xi_i \in \Pi_C(c_i), c_i \rightarrow c \right\}.$$

On appelle le *cône normal de Clarke* de C en c , qu'on note $N_C(c)$, l'adhérence de l'enveloppe convexe de $\hat{N}_C(c)$. C'est-à-dire

$$N_C(c) = cl \left(conv(\hat{N}_C(c)) \right) = clco \left(\hat{N}_C(c) \right).$$

Notons par $\langle \cdot, \cdot \rangle$ le produit scalaire dans \mathbb{R}^n et par B la boule unité fermée de \mathbb{R}^n . Considérons une fonction semi-continue inférieurement

$$\phi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\},$$

et un point $\bar{x} \in \mathbb{R}^n$ où ϕ est finie. un vecteur $\xi \in \mathbb{R}^n$ est appelé *sous-gradient proximal* de ϕ en \bar{x} s'il existe $M > 0$, $\delta > 0$ tels que :

$$\langle \xi, x - \bar{x} \rangle \leq \phi(x) - \phi(\bar{x}) + M\|x - \bar{x}\|^2 \quad \forall x \in \bar{x} + \delta B.$$

L'ensemble de tous les sous-gradients proximaux de $\phi(\cdot)$ en \bar{x} est noté $\partial^\pi \phi(\bar{x})$.

On définit le *pseudo-différentiel* de ϕ en \bar{x} , noté $\hat{\partial} \phi(\bar{x})$, par

$$\hat{\partial} \phi(\bar{x}) = \left\{ \lim_{k \rightarrow \infty} \xi_k : \xi_k \in \partial^\pi \phi(x_k), x_k \rightarrow \bar{x}, \phi(x_k) \rightarrow \phi(\bar{x}) \right\}.$$

Le *pré-sous-gradient* de ϕ en \bar{x} , noté $\hat{\partial}^\infty \phi(\bar{x})$, est l'ensemble

$$\hat{\partial}^\infty \phi(\bar{x}) = \left\{ \lim_{k \rightarrow \infty} t_k \xi_k : \xi_k \in \partial^\pi \phi(x_k), x_k \rightarrow \bar{x}, \phi(x_k) \rightarrow \phi(\bar{x}), t_k \xrightarrow{>} 0 \right\}.$$

La notion de pseudo-différentiel est moins forte que celle du gradient généralisé. En effet, si ϕ est lipschitzienne et continue au voisinage de x , on a $\partial\phi(x) = \text{conv}(\hat{\partial}\phi(x))$. Pour des détails sur ces notions, le lecteur peut consulter les travaux de Clarke [24, 25, 26], et Rockafellar [55].

On définit la *fonction indicatrice* d'un ensemble donné C , qu'on note Ψ_C , de la manière suivante :

$$\Psi_C(x) = \begin{cases} 0 & \text{si } x \in C, \\ +\infty & \text{sinon.} \end{cases}$$

Notons que pour un point $x \in C$, nous avons la formule : $\partial\Psi_C(x) = N_C(x)$.

La proposition qui suit résume quelques propriétés des notions exposées.

Proposition 3.7. [24, 25]

i. Si C est un ensemble convexe et non vide, alors le cône proximal normal de C coïncide avec le cône normal dans le sens de l'analyse convexe, i.e. tout vecteur $\xi \in \hat{N}_C(\bar{x})$ si et seulement si

$$\langle \xi, x + \bar{x} \rangle \leq 0 \quad \forall x \in C.$$

ii. La fonction ϕ est lipschitzienne si et seulement si $\hat{\partial}^\infty\phi(x) = \emptyset$.

iii. Pour tout $\lambda \geq 0$, on a $\hat{\partial}(\lambda\phi)(x) = \lambda\hat{\partial}\phi(x)$.

iv. Soient ϕ et ψ deux fonctions de \mathbb{R}^n dans $\mathbb{R} \cup \{\infty\}$, semi-continues inférieurement et finies en x , avec $\hat{\partial}^\infty\phi(x) \cap (-\hat{\partial}^\infty\psi(x)) = \emptyset$. Alors

$$\hat{\partial}^\infty(\phi + \psi)(x) \subset \hat{\partial}^\infty\phi(x) + \hat{\partial}^\infty\psi(x).$$

v. Soit Ψ_C la fonction indicatrice de l'ensemble C . Alors

$$\hat{N}_C(x) = \hat{\partial}\Psi_C(x) = \hat{\partial}^\infty\Psi_C(x).$$

vi. Soient S_1 et S_2 deux sous-ensembles fermés de \mathbb{R}^n , et soit $\bar{x} \in S_1 \cap S_2$. Si $\hat{N}_{S_1}(\bar{x}) \cap (-\hat{N}_{S_2}(\bar{x})) = \{0\}$, alors on a :

$$\hat{N}_{S_1 \cap S_2}(\bar{x}) \subset \hat{N}_{S_1}(\bar{x}) + \hat{N}_{S_2}(\bar{x}).$$

vii. Soit $\phi(x) = f(F(x))$, où $F : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction lipschitzienne dans un voisinage de \bar{x} , et $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ est semi-continue inférieurement avec $F(\bar{x}) \in \text{dom}f = \{y : f(y) \neq +\infty\}$. Alors

$$0 \notin \hat{\partial}\lambda F \quad \text{pour tout vecteur non nul } \lambda \in \hat{\partial}^\infty f(F(\bar{x})),$$

et on a

$$\hat{\partial}\phi(x) \subset \bigcup \{ \hat{\partial}\lambda F(\bar{x}) : \lambda \in \hat{\partial}^\infty f(F(\bar{x})) \}.$$

3.6 Problème du contrôle optimal bi-niveaux

Considérons un problème hiérarchique à deux niveaux où deux décideurs cherchent à prendre la meilleure décision pour atteindre leurs objectifs qui sont généralement différents. De plus, les deux décideurs ne peuvent pas prendre leurs décisions indépendamment l'un de l'autre, mais qu'ils sont dans un système d'interdépendance hiérarchique, par lequel la stratégie optimale choisie par le niveau inférieur (*suiveur*) dépend de la stratégie choisie par le niveau supérieur (*leader*). En revanche, la fonction objectif du *leader* ne dépend pas seulement de sa propre décision mais aussi de la réaction du *suiveur*. Alors, ayant le premier choix, le *leader* pourra évaluer la vraie valeur de sa propre sélection, seulement après la connaissance des réactions possibles du *suiveur*. Admettons que le jeu est coopératif, c'est-à-dire que si pour une décision donnée du *leader*, le *suiveur* possède plusieurs choix, il choisira toujours la fonction qui convient le mieux de point de vu du *leader*.

En particulier, on considère un système dynamique hiérarchique où la fonction d'état $x(t) \in \mathbb{R}^d$ est déterminée à la fois par les décisions $u(\cdot)$ et $v(\cdot)$ du *leader* et du *suiveur* respectivement. La fonction d'état est d'écrite par :

$$\begin{cases} \dot{x}(t) = \phi(t, x(t), u(t), v(t)) & \text{p.p. } t \in [t_0, t_1] \\ x(t_0) = x_0, \end{cases} \quad (3.10)$$

où $u(t) \in \mathbb{R}^n$ et $v(t) \in V(t) \subset \mathbb{R}^m$.

Mathématiquement, le problème du niveau inférieur pour un choix $u(\cdot)$ du leader est le problème du contrôle optimal ordinaire (à un seul niveau) suivant, paramétrie en u :

$$\begin{cases} \min J_2(x, u, v) = \int_{t_0}^{t_1} G(t, x(t), u(t), v(t))dt + g(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), u(t), v(t)) \\ x(t_0) = x_0, \\ v(t) \in V(t) \quad \text{p.p.} \end{cases} \quad (3.11)$$

Le leader va faire face alors, à un problème du contrôle optimal bi-niveaux suivant :

$$\begin{cases} \min J_1(x, u, v) = \int_{t_0}^{t_1} F(t, x(t), u(t), v(t))dt + f(x(t_1)) \\ u \in L^2([t_0, t_1], \mathbb{R}^n) \\ (x, v) \text{ sont des solutions optimales du problème (3.11).} \end{cases} \quad (3.12)$$

Une formulation plus générale d'un problème de contrôle bi-niveaux est donnée comme suit :

$$\left\{ \begin{array}{l} \min J_1(x, u, v) = \int_{t_0}^{t_1} F(t, x(t), u(t), v(t))dt + f(x(t_1)) \\ u(t) \in U \subset \mathbb{R}^n \\ \left\{ \begin{array}{l} \min J_2(x, u, v) = \int_{t_0}^{t_1} G(t, x(t), u(t), v(t))dt + g(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), u(t), v(t)) \quad \text{p.p.} \\ x(t_0) = x_0, t \in [t_0, t_1] \\ v(t) \in V(t) \subset \mathbb{R}^m. \end{array} \right. \end{array} \right. \quad (3.13)$$

3.6.1 Etude du problème

L'approche de résolution du problème de contrôle optimal bi-niveaux (3.13) consiste à le reformuler comme un problème à un seul niveau.

Pour cela définissons d'abord la *fonction valeur* du niveau inférieur :

$$\mathcal{V}(u) : L^2([t_0, t_1], \mathbb{R}^n) \rightarrow \overline{\mathbb{R}}, \text{ avec}$$

$$\mathcal{V}(u) = \inf_v \left\{ \begin{array}{l} J_2(x, u, v) = \int_{t_0}^{t_1} G(t, x(t), u(t), v(t))dt + g(x(t_1)) : \\ \dot{x}(t) = \phi(t, x(t), u(t), v(t)) \quad \text{p.p.} \\ v(t) \in V(t) \\ x(t_0) = x_0, \end{array} \right\}$$

où $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ est la droite linéaire achevée.

Par suite, en utilisant la fonction valeur, le problème (3.13) devient :

$$\left\{ \begin{array}{l} \min J_1(u, v) = \int_{t_0}^{t_1} F(t, x(t), u(t), v(t))dt + f(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), u(t), v(t)) \quad \text{p.p.}, \\ x(t_0) = x_0, \\ u(\cdot) \in L^2([t_0, t_1], \mathbb{R}^n), v(t) \in V(t) \quad \text{p.p.}, \\ \mathcal{V}(u) \geq \int_{t_0}^{t_1} G(t, x(t), u(t), v(t))dt + g(x(t_1)), \end{array} \right. \quad (3.14)$$

qui est un problème à un seul niveau.

Définition 3.17.

- i. Une fonction contrôle du problème (3.11) est une fonction choisie $v(\cdot)$, mesurable au sens de Lebesgue et satisfaisant $v(t) \in V(t)$ p.p., $t \in [t_0, t_1]$.
- ii. Un *arc* est une fonction absolument continue.
- iii. On appelle *paire admissible* du problème (3.11) un couple de fonction $(x(\cdot), v(\cdot))$ sur $[t_0, t_1]$ où $v(\cdot)$ est une fonction contrôle du problème (3.11) et $x(\cdot) : [t_0, t_1] \rightarrow \mathbb{R}^d$ est un arc satisfaisant l'équation différentielle $\dot{x} = \phi(t, x(t), u(t), v(t))$ p.p, avec la condition initiale $x(t_0) = x_0$.
La première composante d'une paire admissible est appelée *trajectoire admissible*, et la deuxième composante est appelée *contrôle admissible*.
- iv. Une *solution* du problème (3.11) est une paire admissible qui minimise la valeur de la fonction coût $J_2(x, u, v)$ sur l'ensemble de toutes les paires admissibles.
- v. On appelle *stratégie admissible* de (3.12), un couple (u, v) où $u \in L^2([t_0, t_1], \mathbb{R})$ et v un contrôle optimal de (3.11).
- vi. La stratégie (u, v) et la trajectoire correspondante x sont *optimales* pour le problème dynamique bi-niveaux (3.12) si (x, u, v) minimise la valeur de la fonction coût $J_1(x, u, v)$ sur l'ensemble des stratégies admissibles et les trajectoires correspondantes du problème (3.12).

Pour toute la suite de cette partie, nous posons les hypothèses suivantes :

- A1.** $V : [t_0, t_1] \longrightarrow \mathbb{R}^m$ est une fonction multivoque dont l'ensemble des valeurs est non vide et compact. Le graphe de $V(t)$ est défini par l'ensemble

$$GrV = \{(s, r) : s \in [t_0, t_1], r \in V(s)\},$$

qui est $\mathcal{L} \times \mathcal{B}$ -mesurable, où $\mathcal{L} \times \mathcal{B}$ est une σ -algèbre de sous-ensembles de $[t_0, t_1] \times \mathbb{R}^n$ engendrée par $M \times N$ où M est une partie mesurable au sens de Lebesgue, et N est une partie borélienne de \mathbb{R}^m .

- A2.** La fonction $F(t, x, u, v) : [t_0, t_1] \times \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ est $\mathcal{L} \times \mathcal{B}$ -mesurable en (t, v) et continûment différentiable en x et u .
Les fonctions $\phi(t, x, u, v) : [t_0, t_1] \times \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^d$,
 $G(t, x, u, v) : [t_0, t_1] \times \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ sont mesurables en t ,

continûment différentiables en x et u et semi-continues inférieurement en v .

A3. Il existe une fonction intégrable $\psi : [t_0, t_1] \rightarrow \mathbb{R}$ telle que :

$$|\nabla_{(x;u)}F| + |\nabla_{(x;u)}G| + |\nabla_{(x;u)}\phi| \leq \psi(t).$$

A4. La fonction $f(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ est continue, localement lipschitzienne et la fonction $g(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ lipschitzienne, continue, de rang L_g .

A5. Pour tout $u \in L^2([t_0, t_1], \mathbb{R}^n)$, le problème (3.11) a une paire admissible.

3.6.2 Différentiabilité de la fonction valeur

Pour l'étude de la différentiabilité généralisée de la fonction valeur $V(u)$, on aura besoin des hypothèses suivantes :

A6. Il existe $k(t) \in L^2([t_0, t_1], \mathbb{R})$ telle que :

$$|\phi| + |\nabla_x \phi| + |G| + |\nabla_x G| + |\nabla_u G| \leq k(t).$$

A7. Pour chaque $(t, x, u) \in [t_0, t_1] \times \mathbb{R}^d \times \mathbb{R}^n$, l'ensemble :

$$\{(\phi(t, x, u, v), G(t, x, u, v) + r) : v \in V(t), r \geq 0\}$$

est convexe.

A8. La matrice Jacobienne $\nabla_u \phi$ est inversible et :

$$|(\nabla_u \phi)^{-1}| + |\nabla_u \phi| \leq M,$$

où $M > 0$ est une constante.

Définissons l'hamiltonien du problème (3.11) par la fonction suivante :

$$H_2(t, x, u, v, p_2) = p_2 \cdot \phi(t, x, u, v) - G(t, x, u, v).$$

En appliquant le principe du maximum de Pontryaguin, on obtient :

$$H_2(t, x, u, v^0, p_2) = \sup_{v \in V(t)} H_2(t, x, u, v, p_2).$$

Y_u est l'ensemble de toute les trajectoires admissibles x du problème (3.11). Le résultat suivant donne la continuité Lipschitzienne et la caractérisation de la différentiabilité généralisée de la fonction valeur :

Théorème 3.8. [71] *Supposons que les hypothèses (A1)-(A8) sont satisfaites. Alors \mathcal{V} est continûment Lipschitzienne au voisinage de u et on a :*

$$\partial V(u) \subset \text{clco} \cup_{x \in Y_u} \left\{ \zeta : \exists \text{ un arc } p_2 \text{ t.q.} : \begin{array}{l} (-p_2, \zeta, \dot{x}) \in \partial H_2(t, x, u, p_2) \quad \text{p.p.} \\ -p_2(t_1) \in \hat{\partial} g(x(t_1)), \end{array} \right\},$$

où ∂H_2 est le gradient généralisé de Clarke suivant (x, u, p_2) .

La démonstration de ce théorème s'appuie sur le lemme suivant :

Lemme 3.9. [71] *Soit u_i une suite convergente (dans L^2) vers u et soit (x_i, v_i) un paire admissible du problème (3.11). Alors, il existe une sous suite de $\{x_i\}$ convergente uniformément vers un arc x et un contrôle v avec (x, v) est une paire admissible de problème (3.11) telle que :*

$$J_2(x, u, v) \leq \limsup J_2(x_i, u_i, v_i).$$

3.6.3 Conditions nécessaire d'optimalité

Dans cette partie, l'approche est de reformuler le problème original comme un problème d'optimisation de dimension infinie et de tirer le résultat désiré de la condition nécessaire d'optimalité de ce dernier.

Définissons le pseudo- hamiltonien [71] du problème (3.14) comme

$$H_1(t, x, u, v, p_1, \lambda, r) = p_1 \cdot \phi(t, x, u, v) - rG(t, x, u, v) - \lambda F(t, x, u, v),$$

pour tout $t \in [t_0, t_1]$, $x, p_1 \in \mathbb{R}^d$; $u \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, $\lambda, r \in \mathbb{R}$.

Théorème 3.10. [71] *Supposons que les hypothèses (A1) – (A4) sont satisfaites. Soit $(u^*(t), v^*(t))$ une stratégie optimale pour le problème dynamique bi-niveaux (3.13) et $x^*(t)$ la trajectoire correspondante. Supposons que la fonction valeur \mathcal{V} pour le problème de second niveau est continue, localement Lipschitzienne au voisinage de u^* . Alors il existe $\lambda \geq 0, r \geq 0$ et un arc p_1 tels que :*

$$-\dot{p}_1(t) = \nabla_x H_1(t, x^*(t), u^*(t), v^*(t), \lambda, r) \quad \text{p.p}$$

$$\max_{v \in V(t)} H_1(t, x^*(t), u^*(t), v, p_1(t), \lambda, r) = H_1(t, x^*(t), u^*(t), v^*(t), p_1(t), \lambda, r) \quad \text{p.p}$$

$$-p_1(t_1) \in \lambda \partial f(x^*(t_1)) + r \partial g(x^*(t_1))$$

$$\nabla_u H_1(t, x^*(t), u^*(t), v^*(t), \lambda, r) \in r\partial\mathcal{V}(u^*) \quad \text{p.p}$$

$$\|p_1\|_\infty + \lambda + r > 0,$$

où $\|\cdot\|_\infty$ est la norme infinie.

Combinons le théorème 3.8 et le théorème 3.10, on obtient la condition nécessaire d'optimalité de problème général du contrôle optimal bi-niveaux.

Théorème 3.11. [71] *Supposons que les hypothèses (A1) – (A8) sont satisfaites. Soit $(u^*(t), v^*(t))$ une stratégie optimale du problème dynamique bi-niveaux (3.13) et $x^*(t)$ la trajectoire correspondante.*

Alors, il existe $\lambda \geq 0$, $r \geq 0$ et un arc p_1 tels que :

$$-\dot{p}_1(t) = \nabla_x H_1(t, x^*(t), u^*(t), v^*(t), \lambda, r) \quad \text{p.p}$$

$$\max_{v \in V(t)} H_1(t, x^*(t), u^*(t), v, p_1(t), \lambda, r) = H_1(t, x^*(t), u^*(t), v^*(t), p_1(t), \lambda, r) \quad \text{p.p}$$

$$-p_1(t_1) \in \lambda\partial f(x^*(t_1)) + r\partial g(x^*(t_1))$$

$$\nabla_u H_1(t, x^*(t), u^*(t), v^*(t), \lambda, r) \in \text{clco} \cup_{x \in Y_{u^*}} \{\zeta : \exists \text{ arc } p_2 \text{ t.q.} : (-\dot{p}_2, \zeta, \dot{x}) \in \partial H_2(t, x, u^*, p_2) \quad \text{p.p, } -p_2(t_1) \in \hat{\partial}g(x(t_1))\},$$

$$\|p_1\|_\infty + \lambda + r > 0.$$

3.6.4 Cas particulier

Dans ce cas, on considère un système hiérarchique à deux niveaux où le niveau supérieur (*leader*) et le niveau inférieur (*suiveur*) cherchent à trouver, respectivement un vecteur $z \in Z$ et une fonction de contrôle u qui permettent de minimiser leurs fonctions coûts respectives $J_1(z, u)$ et $J_2(z, u)$. Admettons que le *leader* prend sa décision le premier en choisissant un vecteur $z \in Z$ et que par la suite le *suiveur* choisit un fonction de contrôle $u \in \mathcal{U}$, où Z est une partie non vide de \mathbb{R}^n et \mathcal{U} est l'ensemble des contrôles admissibles. Avec cette hypothèse sur l'ordre de prise de décisions, le jeu va se dérouler de la manière suivante : étant donnée une décision $z \in Z$ choisie par le *leader*, le *suiveur* va choisir sa fonction contrôle $u \in \mathcal{U}$, qui dépend du vecteur z choisi

par le *leader*, afin de minimiser sa fonction coût $J_2(z, u_z)$. Admettons aussi que le jeu est coopératif, i.e. que si pour un paramètre z le *suiveur* possède plusieurs choix, il choisira toujours la fonction qui convient le mieux du point de vue du *leader*.

La fonction d'état $x(t) \in \mathbb{R}^d$ est déterminée à la fois par les décisions z et $u(\cdot)$ du *leader* et du *suiveur* respectivement. La fonction d'état est décrite par :

$$\begin{cases} \dot{x} = \phi(t, x(t), z, u(t)) & \text{p.p.} \\ x(t_0) = x_0, x(t_1) \in C_1, t \in [t_0, t_1] \end{cases}$$

Selon ces hypothèses, le problème du niveau inférieur pour un choix $z \in Z$ du leader est donné comme suit :

$$\begin{cases} \min_u J_2(z, u) = \int_{t_0}^{t_1} G(t, x(t), z, u(t)) dt + g(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), z, u(t)) & \text{p.p.} \\ x(t_0) = x_0, x(t_1) \in C_1 \\ u(t) \in U(t) & \text{p.p.} \end{cases} \quad (3.15)$$

qui est un problème du contrôle optimal ordinaire (à un seul niveau).

Le problème du leader est donné par le problème d'optimisation dynamique bi-niveaux suivant :

$$\begin{cases} \min_z J_1(z, u_z) = \int_{t_0}^{t_1} F(t, x_z(t), z, u_z(t)) dt + f(x_z(t_1)) \\ z \in Z \\ (x_z, u_z) \text{ sont des solutions optimales du problème (3.15)} \end{cases} \quad (3.16)$$

Remarque 3.4. Une situation pratique dont le problème de contrôle optimal bi-niveaux peut servir de modèle est la suivante : Le leader est le gouvernement qui impose une taxe z , le suiveur est une entreprise qui cherche une politique optimale $u_z(t)$ en réaction à la politique fiscale du gouvernement.

Étude du problème

Définissons la *fonction valeur* du niveau inférieur du problème du contrôle optimal comme une fonction $\mathcal{V} : Z \longrightarrow \overline{\mathbb{R}}$ avec

$$\mathcal{V}(z) = \inf_u \left\{ \begin{array}{l} J_2(z, u) = \int_{t_0}^{t_1} G(t, x(t), z, u(t)) dt + g(x(t_1)) : \\ \dot{x}(t) = \phi(t, x(t), z, u(t)) \quad \text{p.p.} \\ u(t) \in U(t) \\ x(t_0) = x_0, x(t_1) \in C_1. \end{array} \right\},$$

où $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ est la droite linéaire achevée.

D'abord, commençons par formuler le problème de contrôle bi-niveaux (3.16) comme un problème de contrôle optimal à un seul niveau. C'est le leader qui choisit en premier sa décision z , le suiveur va choisir ensuite sa décision d'après le le vecteur z qui est donc une constante, i.e. ne dépendant pas de t . Définissons la fonction constante :

$$\begin{aligned} z &: [t_0, t_1] \longrightarrow Z \\ t &\longmapsto z(t) = z. \end{aligned}$$

En utilisant la fonction valeur \mathcal{V} et la fonction z , le problème de contrôle optimal bi-niveaux (3.16) peut être formulé d'une manière équivalente comme le problème de contrôle optimal à un seul niveau suivant :

$$\left\{ \begin{array}{l} \min_z J_1(z, t) = \int_{t_0}^{t_1} F(t, x(t), z(t), u(t))dt + f(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), z(t), u(t)) \quad \text{p.p,} \\ \dot{z}(t) = 0, \\ x(t_0) = x_0, x(t_1) \in C_1, \\ u(t) \in U(t) \quad \text{p.p,} \\ \mathcal{V}(z(t_1)) \geq \int_{t_0}^{t_1} G(t, x(t), u(t))dt + g(x(t_1)). \end{array} \right. \quad (3.17)$$

Définition 3.18.

- i. Une *fonction contrôle* est une fonction choisie de U , mesurable au sens de Lebesgue et satisfaisant $u(t) \in U(t)$ p.p., $t \in [t_0, t_1]$.
- ii. Un *arc* est une fonction absolument continue.
- iii. On appelle *paire réalisable* du problème (3.15) un couple de fonction $(x(\cdot), u(\cdot))$ sur $[t_0, t_1]$ tel que u est une fonction contrôle et $x : [t_0, t_1] \rightarrow \mathbf{R}^d$ satisfaisant l'équation $\dot{x} = \phi(t, x(t), z, u(t))$ p.p, avec la condition initiale $x(t_0) = x_0$ et la condition $x(t_1) \in C_1$. La première composante d'une paire admissible est appelée *trajectoire réalisable*, et la deuxième composante est appelée *contrôle réalisable*.

- iv. Une *solution* du problème (3.15) est une paire réalisable qui minimise la valeur de la fonction coût $J_2(z, u)$ sur l'ensemble de toutes les paires réalisables.
- v. On appelle *stratégie réalisable* de (3.16), un couple (z, u_z) où $z \in Z$ et u_z un contrôle optimal de (3.15). La stratégie (z^*, u_z^*) est *optimale* pour (3.16) si elle minimise la valeur de la fonction coût $J_1(z, u)$ sur l'ensemble des stratégies réalisables de (3.16).

Pour toute la suite de ce chapitre, nous posons les hypothèses suivantes :

- H1.** $Z \subset \mathbb{R}^n$ et \mathcal{C}_1 sont fermés.
- H2.** $U : [t_0, t_1] \longrightarrow \mathbb{R}^n$ est une multifonction dont l'ensemble des valeurs est non vide et compact. Le graphe de U , i.e. l'ensemble

$$\{(s, r) : s \in [t_0, t_1], r \in U(s)\},$$

qu'on notera par GrU , est $\mathcal{L} \times \mathcal{B}$ -mesurable, où $\mathcal{L} \times \mathcal{B}$ est une σ -algèbre de sous-ensembles de $[t_0, t_1] \times \mathbb{R}^n$ engendrée par $M \times N$ où M est une partie mesurable au sens de Lebesgue de $[t_0, t_1]$ et N est une partie borelienne de \mathbb{R}^m .

- H3.** Il existe une fonction à valeurs réelles k , $\mathcal{L} \times \mathcal{B}$ -mesurable définie sur GrU telle que : pour tout couple $(t, u) \in GrU$, les fonctions $\phi(t, \dots, u)$, $F(t, \dots, u)$, $G(t, \dots, u)$ sont localement lipschitziennes de rang $k(t)$, où $k(t)$ est une fonction intégrable.
Pour tout couple $(x, y) \in \mathbb{R}^d \times \mathbb{R}^n$, les fonctions suivantes :
 $\phi(\cdot, x, y, \cdot) : [t_0, t_1] \times \mathbb{R}^m \rightarrow \mathbb{R}^d$, $F(\cdot, x, y, \cdot) : [t_0, t_1] \times \mathbb{R}^m \rightarrow \mathbb{R}$,
 $G(\cdot, x, y, \cdot) : [t_0, t_1] \times \mathbb{R}^m \rightarrow \mathbb{R}$ sont $\mathcal{L} \times \mathcal{B}$ -mesurables.
- H4.** Les fonctions $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$ sont continues, localement lipschitziennes.
- H5.** Pour tout $z \in Z$, le problème (3.15) admet une paire réalisable.

Différentiabilité de la fonction valeur

Pour l'étude de la différentiabilité généralisée de la fonction valeur V , on aura besoin des hypothèses suivantes :

- B1.** Pour $z \in Z$, $(t, u) \in GrU$, la fonction $x \mapsto (\phi(t, x, z, u), G(t, x, z, u))$ satisfait la condition d'accroissement suivante : Pour $\alpha \geq 0$, $\beta \geq 0$, on a

$$|(\phi(t, x, z, u), G(t, x, z, u))| \leq \alpha|x| + \beta.$$

A'6. Les fonction ϕ et G sont continues, différentiables en x et en z . Ces fonctions sont aussi semi-continues inférieurement en u .

Il existe une fonction k intégrable telle que :

$$|\phi(t, x, z, u)| + |\nabla_x \phi(t, x, z, u)| + |G(t, x, z, u)| + |\nabla_x G(t, x, z, u)| \leq k(t).$$

B2. Pour chaque $(t, x, z) \in [t_0, t_1] \times \mathbb{R}^d \times \mathbb{R}^n$, l'ensemble

$$\{(\phi(t, x, z, u), G(t, x, z, u)) : u \in U(t)\}$$

est convexe.

A'7. Pour chaque $(t, x, z) \in [t_0, t_1] \times \mathbb{R}^d \times \mathbb{R}^n$, l'ensemble

$$\{(\phi(t, x, z, u), G(t, x, z, u) + r) : u \in U(t), r \geq 0\}$$

est convexe.

Définissons l'hamiltonien du problème (3.15) par la fonction suivante :

$$H_2(t, x, z, u, p_2, \lambda) = p_2 \phi(t, x, z, u) - \lambda G(t, x, z, u).$$

En appliquant le principe de maximum, on obtient :

$$H_2(t, x, z, u^0, p_2, \lambda) = \sup_{u \in U(t)} H_2(t, x, z, u, p_2, \lambda).$$

Définition 3.19. On appelle *multiplicateur d'indice* λ correspondant à une trajectoire réalisable x du problème (3.15) un couple d'arcs (p_2, q) tel que :

$$\left\{ \begin{array}{l} (-p_2(t), -\dot{q}(t), \dot{x}(t)) \in \partial_{(x,z,p_2)} H_2((t, x(t), z, p_2(t), \lambda)) \text{ p.p.}, \\ -p_2(t) \in \lambda \hat{\partial} g(x(t_1)) + \hat{N}_{C_1}(x(t_1)), \\ q(t_1) = 0. \end{array} \right.$$

L'ensemble de tous les multiplicateurs d'indice λ correspondant à x est noté par $M^\lambda(x)$.

Soit Y l'ensemble de toutes les trajectoires optimales du problème (3.15). Posons

$$M^\lambda(Y) = \bigcup_{x \in Y} M^\lambda(x).$$

Pour tout multiplicateur d'indice λ , $(p_2, q) \in M^\lambda(x)$, définissons

$$Q(p_2, q) = -q(t_0).$$

Désignons par $QM^\lambda(x)$ l'ensemble de toutes les valeurs possibles de $-q(t_0)$ obtenues de cette manière, et posons

$$Q(M^\lambda(Y)) = \bigcup_{x \in Y} Q(M^\lambda(x)).$$

Lemme 3.12. [25, 72] *Soit $(\alpha_i)_i$ une suite d'éléments de Z , et soit (x_i, u_i) une paire réalisable du problème qui s'obtient en remplaçant z par α_i dans (3.15), i.e. le problème*

$$\begin{cases} \min_u J_2(\alpha_i, u) = \int_{t_0}^{t_1} G(t, x(t), \alpha_i, u(t)) dt + g(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), \alpha_i, u(t)) \\ x(t_0) = x_0, x(t_1) \in \mathcal{C}_1 \\ u(t) \in U(t) \end{cases} \quad (3.18)$$

Alors il existe une sous-suite de $(x_i)_i$ qui converge uniformément vers x , et il existe un contrôle u tels que (x, u) est une paire réalisable du problème

$$\begin{cases} \min_u J_2(\alpha, u) = \int_{t_0}^{t_1} G(t, x(t), \alpha, u(t)) dt + g(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), \alpha, u(t)) \\ x(t_0) = x_0, x(t_1) \in \mathcal{C}_1 \\ u(t) \in U(t) \end{cases} \quad (3.19)$$

avec $J_2(x, u) \leq \liminf J_2(x_i, u_i)$

Le résultat qui va suivre fait ressortir les propriétés de la différentiabilité de la fonction valeur \mathcal{V} reliée aux notions introduites ci-dessus.

Théorème 3.13. [70] *En plus des hypothèses **H1.** à **H5.**, supposons que l'on a au moins un des deux cas suivant :*

(a) *Les hypothèses **B1.** et **B2.** sont réalisées,*

ou

(b) *Les hypothèses **A'6.** et **A'7.** sont réalisées.*

Si de plus $QM^0(Y) = \{0\}$, alors \mathcal{V} est lipschitzienne et continue au voisinage de z . Et on a

$$\hat{\partial}\mathcal{V}(z) \subset QM^1(Y).$$

Commentaire : Une démonstration de ce théorème sous les hypothèses citées, s'appuie sur le lemme précédent et sur des outils de l'analyse non régulière. Une démonstration assez longue est donnée dans [72] en faisant référence à des résultats de [25] et [27].

Condition nécessaire d'optimalité

Définissons le pseudo- hamiltonien du problème (3.17) comme

$$H_1(t, x, z, p_1; \lambda, r) = p_1 \cdot \phi(t, x, z, u) - rG(t, x, z, u) - \lambda F(t, x, z, u),$$

avec $t \in [t_0, t_1]$, $x, p_1 \in \mathbb{R}^d$, $z \in Z$, $r \in \mathbb{R}$.

Théorème 3.14. [70] *Supposons que :*

- i. Les hypothèses **H1.** à **H5.** sont satisfaites.*
- ii. $(z, u(t))$ une stratégie optimale du problème (3.16).*
- iii. $\hat{\partial}g(x(t_1)) \cap \hat{N}_{C_1}(x(t_1)) = \{0\}$, et $x(t)$ la trajectoire correspondante.*
- iv. La fonction valeur \mathcal{V} du problème du niveau inférieur (3.15) est lipschitzienne et continue.*

Alors il existe $\lambda \geq 0$, $r \geq 0$ et des arcs p_1 , η tels que :

$$\left\{ \begin{array}{l} (\dot{p}_1(t), \dot{\eta}(t)) \in \partial_{(x,z)} H_1(t, x(t), z, p_1(t), u(t); \lambda, r) \quad \text{p.p.}, \quad (1) \\ \max_{u \in U(t)} H_1(t, x(t), z, p_1(t), u; \lambda, r) = H_1(t, x(t), z, p_1(t), u(t); \lambda, r) \quad \text{p.p.}, \quad (2) \\ \eta(t_0) = 0, \quad (3) \\ -p_1(t_1) \in \lambda \hat{\partial}f(x(t_1)) + r \hat{\partial}g(x(t_1)) + \hat{N}_{C_1}(x(t_1)), \quad (4) \\ \eta(t_1) \in r \partial \mathcal{V}(z) \quad (5), \\ \|p_1\|_\infty + \|\eta\|_\infty + \lambda + r > 0, \quad (6) \end{array} \right. \quad (3.20)$$

avec

$$\|\eta\|_\infty = \max_{t_0 \leq t \leq t_1} |\eta(t)|.$$

Pour la démonstration de ce théorème, on aura besoin des lemmes qui suivent.

Lemme 3.15. [70] *Soit $C = \{x : \psi(x) \leq 0\}$, où $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction lipschitzienne et continue en un certain voisinage de $\bar{x} \in C$. Supposons que $0 \notin \hat{\partial}\psi(\bar{x})$. Alors*

$$\hat{N}_C(\bar{x}) \subset \bigcup_{r \geq 0} r \hat{\partial}\psi(\bar{x}).$$

Preuve. Si \bar{x} est un point intérieur de C , alors $\hat{N}_C(\bar{x}) = \{0\}$ et le résultat s'ensuit d'une manière triviale. Supposons que \bar{x} appartient à la frontière de C . En vertu des résultats *i. iii.* et *iv.* de la proposition (3.7), $0 \notin \hat{\partial}\psi(\bar{x})$ implique

$$0 \notin \hat{\partial}r\psi(\bar{x}) \text{ pour tout vecteur non nul } r \in \hat{\partial}^\infty \Psi_{\mathbb{R}_-}(\psi(\bar{x})) = \hat{N}_{\mathbb{R}_-}(\psi(\bar{x})) = \mathbb{R}_+,$$

avec Ψ_C est la fonction indicatrice de C , et

$$\mathbb{R}_- = \{s \in \mathbb{R} : s \leq 0\}, \mathbb{R}_+ = \{s \in \mathbb{R} : s \geq 0\}.$$

Puisque $\Psi_C(\bar{x}) = \Psi_{\mathbb{R}_-}(\psi(\bar{x}))$, du résultat *vii.* de la proposition (3.7), on a

$$\hat{\partial}\Psi_C(\bar{x}) \subset \bigcup \{\hat{\partial}r\psi(\bar{x}) : r \in \hat{\partial}\Psi_{\mathbb{R}_-}(\psi(\bar{x}))\},$$

mais

$$\Psi_{\mathbb{R}_-}(\psi(\bar{x})) = \hat{N}_{\mathbb{R}_-}(\psi(\bar{x})) = \mathbb{R}_+ \quad \text{et} \quad \hat{\partial}\Psi_C(\bar{x}) = \hat{N}_C(\bar{x}),$$

ce qui donne le résultat recherché

$$\hat{N}_C(\bar{x}) \subset \bigcup_{r \geq 0} r \hat{\partial}\psi(\bar{x}).$$

□

Lemme 3.16. [25] *Soit*

$$F(x, y, z) : \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$$

une fonction semi-continue inférieurement, et $(\bar{x}, \bar{y}, \bar{z}) \in \text{dom}F$. Supposons que : $F(x, y, z) = F_1(x) + F_2(y) + F_3(z)$. Alors

$$\hat{\partial}F(\bar{x}, \bar{y}, \bar{z}) \subset \hat{\partial}F_1(\bar{x}) \times \hat{\partial}F_2(\bar{y}) \times \hat{\partial}F_3(\bar{z}).$$

Preuve du théorème. Réécrivons le problème de contrôle optimal (3.17) d'une manière équivalente comme suit :

$$\left\{ \begin{array}{l} \min_z J_1(z, t) = \int_{t_0}^{t_1} F(t, x(t), z(t), u(t))dt + f(x(t_1)) \\ \dot{x}(t) = \phi(t, x(t), z(t), u(t)) \quad \text{p.p.} \\ \dot{y}(t) = G(t, x(t), z(t), u(t)) \quad \text{p.p.} \\ \dot{z}(t) = 0 \\ u(t) \in U(t) \quad \text{p.p.} \\ (x, y, z)(t_0) \in \{x_0\} \times \{0\} \times \mathbb{R} \\ (x, y, z)(t_1) \in S = \{(x, y, z) : g(x) + y - \mathcal{V}(z) \leq 0, x \in C_1\}. \end{array} \right. \quad (3.21)$$

Le pseudo-hamiltonien du problème (3.21) est

$$H(t, x, y, z, p_1, p_2, \eta, u, \lambda) = p_1 \cdot \phi(t, x, z, u) + p_2 G(t, x, z, u) - \lambda F(t, x, z, u),$$

pour $t \in [t_0, t_1]$, $x, p_1 \in \mathbb{R}^d$, $y, p_2, \eta, \lambda \in \mathbb{R}$, $z \in Z$. On applique le principe de maximum de Pontryaguin, en remplaçant le gradient généralisé par le sous-différentiel dans les conditions de transversalité, ce qui se traduit par l'existence d'un scalaire λ et d'un arc (p_1, p_2, η) tels que :

$$(R1) -(\dot{p}_1(t), \dot{p}_2(t), \dot{\eta}(t)) \in \partial_{(x,y,z)} H(t, x(t), y(t), z(t), p_1(t), p_2(t), \eta(t), u(t), \lambda) \quad \text{p.p.}$$

$$(R2) \max_{u \in U(t)} H(t, x(t), y(t), z(t), p_1(t), p_2(t), \eta(t), u, \lambda) \\ = H(t, x(t), y(t), z(t), p_1(t), p_2(t), \eta(t), u(t), \lambda) \quad \text{p.p.}$$

$$(R3) (p_1(t_0), p_2(t_0), \eta(t_0)) \in \hat{N}_{\{x_0\} \times \{0\} \times \mathbb{R}}(x(t_0), y(t_0), z(t_0))$$

$$(R4) -(p_1(t_1), p_2(t_1), \eta(t_1)) \in \lambda \hat{\partial} \hat{f}(x(t_1), y(t_1), z(t_1)) + \hat{N}_S((x(t_1), y(t_1), z(t_1)))$$

$$(R5) \|p_1\|_\infty + \|p_2\|_\infty + \|\eta\|_\infty + \lambda > 0,$$

$$\text{où } \hat{f}(x, y, z) = f(x).$$

Posons $\hat{F}(x, y, z) = g(x) + y - \mathcal{V}(z)$. Alors, d'après le lemme (3.16), on a

$$\hat{\partial}\hat{F}(x, y, z) \subset \hat{\partial}g(x) \times \{1\} \times \hat{\partial}(-\mathcal{V}(z)). \quad (3.22)$$

Ainsi $0 \notin \hat{\partial}\hat{F}(x, y, z)$. Posons $S_1 = \{(x, y, z) : g(x) + y - \mathcal{V}(z) \leq 0\}$ et $S_2 = C_1 \times \mathbb{R} \times \mathbb{R}$. D'après le lemme (3.15) et l'inclusion (3.22), on a

$$\begin{aligned} \hat{N}_{S_1}(x, y, z) &\subset \bigcup_{r \geq 0} r \hat{\partial}\hat{F}(x, y, z) \\ &\subset \bigcup_{r \geq 0} r \hat{\partial}g(x) \times \{1\} \times \hat{\partial}(-\mathcal{V})(z). \end{aligned}$$

Puisque $\Psi_{S_2}(x, y, z) = \Psi_{C_1}(x) + \Psi_{\mathbb{R}}(y) + \Psi_{\mathbb{R}}(z)$, et d'après le lemme (3.16) et *v.* de la proposition (3.7), on a

$$\hat{N}_{S_2}(x, y, z) \subset \hat{N}_{C_1} \times \{0\} \times \{0\}, \quad \forall (x, y, z) \in C_1 \times \mathbb{R} \times \mathbb{R}.$$

Comme par hypothèse $\hat{\partial}g(x(t_1)) \cap \hat{N}_{C_1}(x) = \{0\}$, ainsi

$$\hat{N}_{S_1}(x(t_1), y(t_1), z(t_1)) \cap \left(-\hat{N}_{S_2}(x(t_1), y(t_1), z(t_1)) \right) = \{0\}.$$

En appliquant *vi.* de la proposition (3.7), on aura

$$\begin{aligned} \hat{N}_S(x(t_1), y(t_1), z(t_1)) &\subset \hat{N}_{S_1}(x(t_1), y(t_1), z(t_1)) + \left(-\hat{N}_{S_2}(x(t_1), y(t_1), z(t_1)) \right) \\ &\subset \bigcup_{r \geq 0} \hat{\partial}g(x(t_1)) \times \{1\} \times \hat{\partial}(-\mathcal{V})(z) \\ &\quad + \hat{N}_{C_1}(x(t_1)) \times \{0\} \times \{0\}. \end{aligned}$$

D'après le lemme (3.16), on a

$$\hat{\partial}\hat{f}(x(t_1), y(t_1), z(t_1)) \subset \hat{\partial}f(x(t_1)) \times \{0\} \times \{0\}.$$

Ainsi du résultat (R4), on a

$$\begin{aligned} -(p_1(t_1), p_2(t_1), \eta(t_1)) &\in \lambda \hat{\partial}f(x(t_1)) \times \{0\} \times \{0\} \\ &\quad + \bigcup_{r \geq 0} r \hat{\partial}g(x(t_1)) \times \{1\} \times \hat{\partial}(-\mathcal{V})(z) \\ &\quad + \hat{N}_{C_1}(x(t_1)) \times \{0\} \times \{0\} \\ &\subset \lambda \hat{\partial}f(x(t_1)) \times \{0\} \times \{0\} \\ &\quad + \bigcup_{r \geq 0} r \hat{\partial}g(x(t_1)) \times \{1\} \times (-\partial\mathcal{V}(z)) \\ &\quad + \hat{N}_{C_1}(x(t_1)) \times \{0\} \times \{0\}. \end{aligned}$$

De là, les conditions de transversalité (4) et (5) du théorème (3.14) s'en suivent, et on a $p_2(t_1) = -r$, où $r \geq 0$. Puisque le hamiltonien H est indépendant de y , le résultat (R1) implique que $\dot{p}_2(t) = 0$ et

$$(R6) - (\dot{p}_1(t), \dot{\eta}(t)) \in \partial_{(x,z)} H(t, x(t), y(t), z(t), p_1(t), p_2(t), \eta(t), u(t), \lambda) \text{ p.p.}$$

Ainsi $p_2 = -r$, $r \geq 0$, les conclusions (1), (2) et (6) du théorème (3.14) résultent de (R6), (R2) et (R5) respectivement. De (R3) on a $\eta(t_0) = 0$, ce qui achève la démonstration. \square

Il existe des situations où le suiveur possède un *coût favorable*. C'est le cas où le suiveur ne participe que si le coût optimal est inférieur ou égal au coût favorable $L \geq 0$; au dessus de ce seuil L , il ne court aucun risque de *perte*. Dans ce cas, on obtient un nouveau modèle du problème de contrôle optimal bi-niveaux qui prend en considération cette nouvelle donnée :

$$\left\{ \begin{array}{l} \min_z J_1(z, u) = \int_{t_0}^{t_1} F(t, x(t), z(t), u(t))dt + f(x(t_1)), \\ \dot{x}(t)\phi(t, x(t)z(t), u(t)) \quad \text{p.p.}, \\ \dot{z}(t) = 0 \\ x(t_0) = x_0, \quad x(t_1) \in C_1, \\ u(t) \in U(t) \quad \text{p.p.}, \\ \mathcal{V}(z) \geq \int_{t_0}^{t_1} G(t, x(t), z(t), u(t))dt + g(x(t_1)), \\ L \geq \int_{t_0}^{t_1} G(t, x(t), z(t), u(t))dt + g(x(t_1)). \end{array} \right. \quad (3.23)$$

On obtient une condition nécessaire d'optimalité pour le problème (3.23) en utilisant les mêmes techniques décrites dans la démonstration du théorème 3.14.

Théorème 3.17. [70] *Supposons que :*

- i. Les hypothèses **H1.** à **H5.** sont satisfaites.*
- ii. On a au moins l'une des situations suivantes : les hypothèses **B1.** et **B2.** sont satisfaites ou les hypothèses **A'6.** et **A'7.** sont satisfaites.*

Soit (z, u) une stratégie optimale pour le problème (3.23), et $x(t)$ la trajectoire correspondante. Supposons que $\hat{\partial}g(x(t_1)) \cap \hat{N}_{C_1}(x(t_1)) = \{0\}$ et $QM^0(Y) = \{0\}$. Alors il existe :

- i. des scalaires $\lambda \geq 0$, $r \geq 0$, $0 \leq \hat{r} \leq r$,
- ii. des entiers I, J , λ_{ij} avec $\sum_{i=1}^I \sum_{j=1}^J \lambda_{ij} = 1$,
- iii. des trajectoires optimales $x_i(t)$ pour le problème du niveau inférieur (3.15),
- iv. des arcs p, η, p_{ij}, q_{ij}

tels que :

$$-(\dot{p}(t), \dot{\eta}(t)) \in \partial_{(x,z)} H_1(t, x(t), z, p(t), u(t), \lambda, r) \quad \text{p.p.},$$

$$\max_{u \in U(t)} H_1(t, x(t), z, p(t), u, \lambda, r) = H_1(t, x(t), z, p(t), u(t), \lambda, r) \quad \text{p.p.},$$

$$\eta(t_0) = 0,$$

$$-p(t_1) \in \hat{\partial}f(x(t_1)) + r\hat{\partial}g(x(t_1)) + \hat{N}_{C_1}(x(t_1)),$$

$$\eta(t_1) = \hat{r} \sum_{ij} \lambda_{ij} q_{ij}(t_0)$$

$$(-\dot{p}_{ij}(t), -\dot{q}_{ij}(t), \dot{x}(t)) \in \partial_{(x,z,p)} H_2(t, x_i(t), z, p_{ij}(t), 1) \quad \text{p.p.},$$

$$q_{ij}(t_1) = 0,$$

$$-p_{ij}(t_1) \in \hat{\partial}g(x_i(t_1)) + \hat{N}_{C_1}(x_i(t_1)),$$

$$\|p\|_\infty + \|\eta\|_\infty + \lambda + r > 0.$$

Exemple

L'exemple que nous allons traiter dans cette section est un problème de la régulation de la pêche ; formulé et résolu pour la première fois par Clarke et Murno (cité dans [70]).

Il est généralement reconnu que les ressources halieutiques se trouvant à une distance ne dépassant pas 200 miles² des côtes sont la propriété des

²Mile est une mesure anglo-saxonne de longueur équivalant à 1609 mètres

États côtiers adjacents. Pour ces États choisissant d'autoriser la présence dans les eaux éloignées de leurs côtes de 200 miles de flottes de pêche, un des problèmes auquel ils font face est de trouver les conditions générales optimales de l'accès à leurs zones d'économie exclusives que vont imposer aux flottes pêchant dans ces eaux.

Supposons qu'en temps t , la population de poissons est régie par le système dynamique suivant

$$\dot{x}(t) = F(x) - qE(t)x(t),$$

avec :

$x(t)$: est la taille de la population de poissons au temps t ,

$F(x)$: est le taux naturel de sa croissance,

$E(t)$: est la quantité de la pêche en temps t ,

q est une constante positive.

Admettons que $F(x) > 0$ pour $0 < x < \bar{x}$, $F(0) = F(\bar{x}) = 0$ et $F''(x) < 0$ pour tout $x > 0$, où \bar{x} est la charge utile de la ressource. Nous supposons aussi que

$$0 \leq E(t) \leq E_{max},$$

où E_{max} est la borne supérieure de $E(t)$.

Supposons que l'État côtier impose la condition que la population de poissons ne devra pas descendre au dessous du seuil $\tilde{x} \geq 0$ au temps terminal T_1 .

L'état côtier établit une stratégie (ensemble de taxes : n :=taxe sur la récolte et m :=taxe sur l'effort) pour laquelle la flotte de pêche doit répondre par une action $E(t)$. Cette flotte, va jouer le rôle du décideur du niveau inférieur (suiveur). Afin de maximiser son gain, ce suiveur aura à résoudre le problème de contrôle optimal suivant :

$$\left\{ \begin{array}{l} \max_E \int_0^{T_1} e^{-\delta t} [(p_0 - n)qx(t) - (c_0 + m)]E(t)dt, \\ \dot{x}(t) = F(x(t)) - qE(t)x(t), \\ x(0) = x_0, \quad x(T_1) \geq \tilde{x}, \\ E(t) \in [0, E_{max}], \end{array} \right. \quad (3.24)$$

où p_0 représente le prix unitaire d'une prise de poissons, c_0 le coût des charges, et $\delta > 0$ le taux d'actualisation.

Ce problème de contrôle optimal est linéaire. L'hamiltonien est égal :

$$H_2(t, x, n, m, E, p_2) = p_2(t)[F(x(t)) - qEx(t)] + e^{-\delta t}[(p_0 - n)qx(t) - (c_0 + m)]E,$$

avec $p_2(T_1) \geq 0$

En appliquant le principe du maximum, on aura :

$$\max_{E(t) \in [0, E_{max}]} H_2(t, x, n, m, E, p_2) = H(t, x, n, m, E(t), p_2),$$

avec

$$H_2(t, x, n, m, E(t), p_2) = p_2(t)[F(x(t)) - qE(t)x(t)] + e^{-\delta t}[(p_0 - n)qx(t) - (c_0 + m)]E(t),$$

$$p_2(T_1) \geq 0.$$

La condition nécessaire pour qu'un couple (x, E) soit solution de (3.24) est l'existence d'un arc p_2 tel que :

$$\dot{p}_2(t) = -\frac{\partial H_2}{\partial x} = -\{p_2(t)[F'(x) - qE] + e^{-\delta t}((p_0 - n)qE)\},$$

d'où

$$-\dot{p}_2(t) = p_2(t)[F'(x) - qE] + e^{-\delta t}(p_0 - n)qE, \quad (3.25)$$

Comme l'objectif est que $E(t)$ maximise l'hamiltonien, $E(t)$ doit être ou bien le contrôle singulier, ou bien $E(t) = 0$ ou $E(t) = E_{max}$. Le contrôle singulier apparaît quand le coefficient de E dans l'hamiltonien est égal à zéro, c'est à dire :

$$H_2(t, x, n, m, E(t), p_2) = E(t)\{-p_2(t)qx(t) + e^{-\delta t}[(p_0 - n)qx(t) - (c_0 + m)]\} + p_2(t)F(x(t)),$$

alors : $-p_2(t)qx(t) + e^{-\delta t}[(p_0 - n)qx(t) - (c_0 + m)] = 0$, ce qui implique :

$$p_2(t) = e^{-\delta t}\left[(p_0 - n) - \frac{c_0 + m}{qx}\right].$$

On aura :

$$\dot{p}_2(t) = e^{-\delta t} \left[-\delta \left[(p_0 - n) - \frac{c_0 + m}{qx} \right] + \frac{c_0 + m}{qx^2} \frac{dx}{dt} \right]. \quad (3.26)$$

Pour l'équation adjointe (3.25), on a

$$\begin{aligned} \dot{p}_2(t) &= -p_2[F'(x) - qE] - e^{-\delta t}[(p_0 - n)qE] \\ &= -e^{-\delta t} \left\{ \left[(p_0 - n) - \frac{c_0 + m}{qx} \right] [F'(x) - qE] + (p_0 - n)qE \right\}, \end{aligned}$$

L'expression précédente de $\dot{p}_2(t)$ est égale à (3.26), le contrôle E disparaît et on obtient l'égalité suivante

$$\delta = F'(x) + \frac{F(x)(c_0 + m)/qx^2}{p_0 - n - (c_0 + m)/qx}. \quad (3.27)$$

Pour n et m fixés, cette équation admet une solution unique x_* représentant la biomasse optimale (la masse totale de poissons peuplant le milieu). La trajectoire optimale est celle qui va le plus vite vers la biomasse optimale x_* .

Soit $\mathcal{V}(n, m)$ la valeur optimale du problème ci-dessus. La participation d'une flotte de pêche est imposée uniquement dans le cas où $\mathcal{V}(n, m) \geq L$, où L désigne la rémunération alternative des autres états côtiers.

L'État côtier va jouer le rôle du décideur du niveau supérieur (le leader), et aura à résoudre le problème d'optimisation dynamique bi-niveaux suivant :

$$\left\{ \begin{array}{l} \max_{(n,m)} \int_0^{T_1} e^{-\delta t} (nqx(t) + m)E(t)dt, \\ \dot{x}(t) = F(x(t)) - qE(t)x(t), \\ x(0) = x_0, \quad x(T_1) \geq \tilde{x}, \\ E(t) \in [0, E_{max}] \quad \text{p.p.}, \\ \mathcal{V}(n, m) \leq \int_0^{T_1} e^{-\delta t} [(p_0 - n)qx(t) - (c_0 + m)]E(t)dt, \\ \mathcal{V}(n, m) \geq L. \end{array} \right. \quad (3.28)$$

Le hamiltonien de ce problème est donné par :

$$H_1(t, x, n, m, E, p_1, \lambda, r) = p_1(t)[F(x(t)) - qEx(t)] + e^{-\delta t} [r [(p_0 - n)qx(t) - (c_0 + m)] + \lambda(nqx(t) + m)] E$$

On vérifie facilement que les conditions du théorème 3.17 sont satisfaites. Notons que le problème du suiveur a une solution unique. En vertu de ce théorème on a : si (n, m, x, E) est une solution optimale du problème (3.28), alors il existe des arcs $p_1, p_2, \eta_1, \eta_2, q_1, q_2$ et des scalaires $\lambda \geq 0, r \geq 0$,

$0 \leq \hat{r} \leq r$ tels que :

$$-\dot{p}_1 = \frac{\partial H_1}{\partial x} = p_1[F'(x) - qE] + e^{-\delta t}[r(p_0 - n) + \lambda n]qE, \quad (1)$$

$$\dot{\eta}_1 = \frac{\partial H_1}{\partial n} = (r - \lambda)e^{-\delta t}qx E, \quad (2)$$

$$\dot{\eta}_2 = \frac{\partial H_1}{\partial m} = (r - \lambda)e^{-\delta t}E, \quad (3)$$

$$\max_{E \in [0, E_{max}]} H_1(t, x, n, m, E, p_1, \lambda, r) = H_1(t, x, n, m, E(t), p_1, \lambda, r), \quad (4)$$

$$(\eta_1, \eta_2)(0) = (0, 0), \quad (5)$$

$$p_1(T_1) \geq 0, \quad (6)$$

$$(\eta_1, \eta_2)(T_1) = \hat{r}\lambda q(0), \quad (7)$$

$$-\dot{p}_2 = \frac{\partial H_2}{\partial x} = p_2[F'(x) - qE] + e^{-\delta t}(p_0 - n)qE, \quad (8)$$

$$\dot{q}_1 = \frac{\partial H_2}{\partial n} = e^{-\delta t}qx E, \quad (9)$$

$$\dot{q}_2 = \frac{\partial H_2}{\partial m} = e^{-\delta t}E, \quad (10)$$

$$\max_{E(t) \in [0, E_{max}]} H_2(t, x, n, m, E, p_2) = H(t, x, n, m, E(t), p_2), \quad (11)$$

$$(q_1, q_2)(0) = (0, 0), \quad (12)$$

$$p_2(T_1) \geq 0, \quad (13)$$

$$\|p_1\|_\infty + \|\eta\|_\infty + \lambda + r > 0. \quad (14)$$

Sans perte de généralité, posons $\lambda = 1$. Des résultats (1) et (4) précédents, il apparaît que l'état stable (n, m, x_*) du problème (3.28) est une solution de l'équation suivante :

$$\delta = F'(x_*) + \frac{F(x_*)(r(c_0 + m) - m)/qx_*^2}{r(p_0 - n) + n - (r(c_0 + m) - m)/qx_*}. \quad (3.29)$$

Puisque (n, m, x, E) est une solution optimale du problème (3.28), (x, E) est alors la solution optimale du problème (3.24) et x_* est la biomasse optimale associée avec (n, m) est définie par l'équation (3.29). En combinant les

équations (3.29) et (3.27) on aura :

$$n = \rho p_0,$$

$$m = -\rho c_0,$$

où ρ est une constante qui provient de la résolution du système des deux équations (3.29) et (3.27) aux deux inconnues n et m . Clarke et Murno ont montré que pour la flotte de pêche, le couple taxe-charge (n, m) optimal doit être atteint pour $\mathcal{V}(n, m) = L$, ce qui donne $\rho = (\mathcal{V}_0 - L)/\mathcal{V}_0$, où \mathcal{V}_0 est le revenu global net pour la flotte,

$$\mathcal{V}_0 = \max \left(\int_0^{T_1} e^{-\delta t} (p_0 q x(t) - c_0) E(t) dt \right),$$

et que la politique de pêche optimale $E(t)$ permettra de maximiser le revenu global net pour les flottes de pêche. La condition nécessaire d'optimalité ci-dessus est satisfaite pour $\lambda = 1$, $r = 1$ et $\hat{r} = 0$.

CONCLUSION

L'objet de ce travail est l'étude théorique des problèmes de programmation à deux niveaux et leurs applications au problème de contrôle optimal.

Pour l'optimisation bi-niveaux, différents algorithmes ont été mis au point pour la résolution du cas linéaire. Dans notre cas, nous avons transformé le problème initial bi-niveaux à un problème d'optimisation classique (à un seul niveau). Cette classe de problèmes d'optimisation est une classe NP-difficile. Il ressort aussi, que le traitement de la position pessimiste s'affronte à plus de difficultés, et mériterait d'être abordée. Les reformulations de ces problèmes à d'autres classes ne doivent pas être simplement interprétées comme des outils pour transformer un problème en un autre. Elles procurent un point de vue qui permet une meilleure compréhension des caractéristiques inhérentes à ces problèmes, et permet parfois d'obtenir des algorithmes d'une classe en les adaptant à une autre.

Quant aux problèmes de contrôle optimal, nous avons d'abord développé certains éléments de base de cette théorie, en portant un intérêt particulier à la contrôlabilité des systèmes linéaires et non linéaires. À la fin nous avons étudié les problèmes de contrôle optimal bi-niveaux. Notre démarche a été inspiré des notions de l'analyse non lisse (non régulière). Nous avons considéré deux cas : Dans le premier cas, les variables de décision du leader et du suiveur sont des fonctions contrôles, alors que, dans le deuxième cas, la variable de décision du leader est un vecteur et celle du suiveur est une fonction contrôle. Pour obtenir une condition d'optimalité, le problème a été reformulé en un problème à un seul niveau, en faisant intervenir la fonction valeur du niveau inférieur.

Ce domaine offre de nombreuses perspectives :

- ✓ Etude des problèmes de contrôle optimal bi-niveaux sans coopération entre le leader et le suiveur.
- ✓ Etude des systèmes avec des commandes prédictives.
- ✓ Systèmes perturbés.

Bibliographie

- [1] S. Alarie, C. Audet, B. Jaumard, G. Savard. *Concavity cuts for Disjoint Bilinear Programming*. Center of Research on Parallel Computation, Rice University, Houston, 1999.
- [2] G. Anandalingram and D. j. White. *A solution Method for the Linear Stackelberg Problem Using Penalty Functions*. IEEE Trans. Automatic Control, Vol. 35, No. 10, 1990.
- [3] A. Anzi. *Résolution d'un problème de programmation bi-niveaux linéaire par la méthode D.C.*. Mémoire de Magister, Université de Bejaïa, 2009.
- [4] C. Audet, G. Savard, W. Zghal. *New brunch-and-cut algorithm for bi-level linear programming problem*. Journal of Optimization, Theory and Application, 2007.
- [5] C. Audet, P. Hansen, B. Jaumard, G. Savard. *A Symmetrical Linear Maximum Approach to Disjoint Bilinear Programming*. Mathematical Programming, 1999.
- [6] C. Audet, Hansen, B. Jaumard, G. Savard. *Links between Linear Bilevel and Mixed 0-1 Programming Problems*. Journal of Optimization, Theory and Application, 1997.
- [7] C. Audet. *Optimisation globale structurée, propriétés, équivalences et résolution*. Thèse Ph.D, Université de Montreal, Département de Mathématiques et de Génie Industriel, 1997.
- [8] B. Bank, J. Guddat, D. Klatte, B. Kummer, K. Temmar. *Non-linear parametric optimization*. Akademie-Verlag, Berlin, 1982.
- [9] J. F. Bard. *Practical Bilevel Optimization : Algorithms and Application*. Kluwer Academic Publisher, Dordrecht, 1998.
- [10] J. F. Bard and J. E. Falk. *An explicit Solution to the Multi-Level Programming Problem*. Computers and Operation Research, VI. 9, 1982.

- [11] C. Berge. *Espaces topologiques, Fonctions multivoques*. Collection Universitaire de Mathématiques, Dunod, Paris, 1966.
- [12] D. P. Bertsekas. *Convex Analysis and Optimization*. Athena Scientific USA, 2003.
- [13] D. P. Bertsekas. *Nonlinear Programming*. 2nd Edition, Athena Scientific, Belmont, 1999.
- [14] W. F. Bialas and M. H. Karwan. *Two-Level Programming*. Management Science, Vol. 30, 1984.
- [15] W. F. Bialas and M. H. Karwan. *On Two-Level Optimization*. IEEE Trans. Automatic Control, Vol. AC-27, No. 1, pp.211-214, 1982.
- [16] J. M. Borwein, A. S. Lewis. *Convex analysis and nonlinear optimization, theory and examples*. Auto-edition, 1999.
- [17] S. Boyd, L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [18] H. Brezis. *Analyse fonctionnelle : Théorie et applications*. Masson, Paris, 1983.
- [19] L. Brotcorne, P. Marcotte, G. Savard. *Bilevel Programming : The Montreal School*. University of Montreal, 2008.
- [20] A. Cambini, L. Martein. *Generalized Convexity and Optimization, theory and applications*. Springer-Verlag, Berlin, Heidelberg, 2009.
- [21] W. Candler and R. Townsley. *A Linear Two-Level Programming Problem*. Computers and Operations Research, Vol. 9, 1982.
- [22] M. Campêlo, S. Scheimberg. *A Study of Local Solution in Linear Bilevel Programming*. Journal of Optimization Theory and Application, Vol. 125, No. 1, 2005.
- [23] C. I. Chen and J. B. Cruz Jr. *Stackelberg solution for two-person games with biased information pattern*. IEEE Trans. Automatic Control, No. 6, 1972.
- [24] F. H. Clarke et al. *Nonsmooth Analysis and Control Theory*. Springer-Verlag, New York, Inc. 1998.
- [25] F. H. Clarke. *Optimization and nonsmooth Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1990.
- [26] F. H. Clarke. *Methods of Dynamic and Nonsmooth Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1989.

- [27] F. H. Clarke, P. D. Loewen. *The value function in optimal control : sensitivity, controllability, and time-optimality*. SIAM, Control and Optimization, No. 24, 1986.
- [28] B. Colson, P. Marcotte and G. Savard. *Bilevel programming : A survey*. 4OR A Quarterly Journal of Operations Research, 2007.
- [29] J. Dattorro. *Convex Optimization and Euclidean Distance Geometry*. Meboo Publishing USA, 2005.
- [30] S. Dempe, J. Dutta, B. S Mordukhovich. *Variational Analysis in Bilevel Programming*. In *Mathematical programming and Theory of Decision Making*, pp. 257-278, S. K Neogy, R. B. Bapat, A. K. Das, T. Parthasarathy (Eds), World Scientific Publishing Co. Pte. Ltd, 2008. World Scientific Publishing, 2008.
- [31] S. Dempe et V. Kalashnikov. *Optimization with multivalued mappings*. pp.3-20. Springer Science + Business Media, LLC, 2006.
- [32] S. Dempe. *Bilevel Programming*. In *Essays and Surveys in global optimization*, C. Audet, G. Savard, H. Hansen (Eds), pp. 165-194, Springer Science+ Business Media, Inc, 2005.
- [33] S. Dempe. *Foundations of bilevel programming*. Dordrecht : Kluwer Academic Publishers, 2002.
- [34] J. B. Etoa Etoa. *Optimisation hiérarchique : théorie, algorithmes et applications*. Editions Publibook, 2008.
- [35] N. Faísca, V. Dua, B. Rustem, p. M. Saravia, E. N. Pistikopoulos. *Parametric Global Optimization for Bilevel Programming*. Springer Science +Business Media B.V. 2006.
- [36] C. Fricke. *An Introduction to Bilevel Programming*. Departement of Mathematics and Statistics, University of Melbourne.
- [37] P. Gács, L. Lovász. *Complexity of Algorithms : Lecture Notes*, Spring 1999.
- [38] M. R. Garey and D.S Johnson. *Computers and Intractability : A guide to the theory of NP-completeness* . W. H. Freeman and Co.,San Francisco, 1979.
- [39] M. Gendreau, P. Marcotte and G. Savard. *A Hybrid Tabu-Ascent Algorithm for the linear Bilevel Programming Problem*. Journal of global optimizaton. Kluwer Academic Publishers, Boston, 1996.

- [40] G. Gilles. *Introduction à l'algorithmique, notes de cours*. Année Académique 2007-2008.
- [41] G. Giorgi, A. Guerraggio, J. Thierfelder. *Mathematics of Optimization : Smooth and Nonsmooth Case*. Elsevier Science and Technology Books, 2004.
- [42] J. Guddat, F. Guerra Vasquez, H. Th. Jongen. *Parametric Optimization : Singularities, Pathfollowing and Jumps*. John Wiley & Sons, Chichester and B. F. Teubner, Stuttgart, 1990.
- [43] K. Hamid. *Programmation mathématique bi-niveaux et contrôle optimal bi-niveaux*. Mémoire de Magister, Université Mouloud Mammeri de Tizi-Ouzou, 2010.
- [44] S. R. Hedjazi, A. Memariani, G. Jahanshanloo, M. M. Sepehri. *Bilevel Programming Solution by Genetic Algorithms*. Proceedings of the First National Industrial Engineering Conference, 2001.
- [45] J. J. Júdice and A. M. Faustino. *A Sequential LCP Method for Bilevel Linear Programming*. Annals of Operations Research, Vol. 34, 1992.
- [46] A. Kolmogorov, S. Fomine. *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*. Éditions Mir-Moscou, 1973.
- [47] H. Konno. *Bilinear Programming*. Technical Report No. 71-10, Operations Research House, Department of Operations Research, Stanford University, Stanford, 1971.
- [48] C. Kunapuli, K. P. Bennetti, J. Hu, J. S. Pang. *Classification model selection via bilevel programming*. Optimization Methods and software, 2007.
- [49] R. Lucchetti. *Convexity and well-Posed Problems*. Springer Science+Business Media, Inc, 2006
- [50] P. Marcotte. G. Savard. *Bilevel Programming : A Combinatorial Perspective*. Chapter 1.
- [51] A. G. Mersha and S. Dempe. *Linear Bilevel Programming with Upper Level Constraints Depending on the Lower Level Solution*. Applied Mathematics and Computation, 2006.
- [52] M. Minoux. *Programmation mathématique, théorie et algorithmes*. Bordas et C.N.E.T-E.N.S.T, Paris 1983.

- [53] M. S. Osman, W. F. Abd El Wahed, M. M. K El Shafei and H. B. Abd El Wahab. *A solution Methodology of Bilevel Linear Programming base on Genetic Algorithm*.
- [54] L. Pontriaguine, V. Boltianski, R. Gamkrelidze, E. Michtchenko. *Théorie Mathématique des processus optimaux*. Edition Mir, Moscou, 1974.
- [55] R. T. Rockafellar. *Extensions of subgradient calculus wiht applications to optimization*. Nonlinear Analysis, Theory and Applications, No 9, 1985.
- [56] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [57] C. D. Rodrigues, D. Quadri, P. Michelon. *Un modèle bi-niveau pour le problème de la recherche d'une cible dynamique*. MajecSTIC, Avignon, France, 2009.
- [58] Jh. Ryu, E. N. Pistikopoulos. *A bilevel programming framework for enterprise-wide supply chain planning problem under uncertainty*. Centre of Systems Engineering, Department of Chemical Engineering, Imperial College, London, U.K, 2004.
- [59] K. Schittkowski, Ch. Zillober. *Nonlinear programming*. Department of mathematics, University of Bayreuth, D-95440 Bayreuth, Germany.
- [60] W. Schirotzek. *Nonsmooth Analysis*. Springer-Verlag, Berlin Heidelberg, 2007.
- [61] C. Shi, J. Lu, G. Zhang. *An Extended Branch and Bound Algorithm for Linear Bilevel Programming* . Faculty of Information Technology, University of Technology, Sydney, Australian, 2007.
- [62] C. Shi, G. Zhang, and J. Lu. *On the definition of linear bilevel programming solution*. Application Mathematics Compututer, 160, pp 169-176, 2005.
- [63] K. Shimuzu, Y. Ishizuka, J. F. Bard. *Nondifferentiable and two level mathematical programming*. Kluwer Academic Publishers, 1997.
- [64] K. Shimizu and M. Lu. *A Global Optimization Method for the Stackelberg Problem with Convex Functions via Problem Transformation and Concave Programming*. IEEE Trans. Systems, Man and Cybernetics, Vol. SMC-25, No. 12, 1995.
- [65] S. M. Sinha. *Mathematical programming, theory and methods*. Elsevier Science and Technology Books, 2006.

- [66] J. A. Snyman. *Practical Mathematical Optimization : An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms* . Springer Science+Business Media, Inc., 2005.
- [67] E. Trélat. *Contrôle optimal : Théorie et applications*. Vuibert, Collection "Mathématiques Concrètes", 2005.
- [68] L. N. Vicente, G. Savard, J. J. Júdice. *The Discrete Linear Bilevel Programming Problem*. Journal of Optimization, Theory and Application, 1996.
- [69] W. Xiandjia, H. Yuan and Z. Wuyue. *Bilevel Programming Model of Earthwork Allocation System*. Y. Shi et al. (Eds) : MCDM 2009, CCIS 35, pp. 275-281, Springer-Verlag, Berlin, Heidelberg, 2009.
- [70] J. J. Ye. *Necessary conditions for bilevel dynamic optimization problems*. Department of Mathematics and Statistics, University of Victoria, Canada, 1993.
- [71] J. J. Ye. *Optimal Strategies for Bilevel Dynamic Problems*. Department of Mathematics and Statistics, University of Victoria, Canada, 1993.
- [72] Y. Yin. *Genetic Algorithms Based Approach for Bilevel Programming Models*. Journal of Transportation Engineering, No. 126(2), 2000.
- [73] M. I. Zelikin. *Control theory and Optimization*. Springer, 1999.
- [74] Z. Zheng, G. Zhang, Q. He, J. Lu, Z. Shi. *Rule sets Based Bilevel Decision Model*. Australian Computer Society, Inc, 2006.

Application de la Programmation bi-niveaux au Problème de contrôle optimal

Résumé : On s'intéresse ici à l'étude du problème d'optimisation bi-niveaux qui est un problème hiérarchique de deux niveaux de décision : Un niveau supérieur appelé *leader* et un niveau inférieur appelé *suiveur*. La solution optimale dépend évidemment des deux niveaux qui sont en interdépendance hiérarchique.

La seconde partie de ce mémoire traite le problème de contrôle optimal bi-niveaux. Pour obtenir une condition d'optimalité, on reformule le problème à deux niveaux en un problème de contrôle optimal à un seul niveau, et on fait intervenir la fonction valeur du problème du niveau inférieur. L'étude réalisée dans ce mémoire utilise les outils de l'analyse non régulière.

Mots clés : programmation mathématique bi-niveaux, contrôle optimal bi-niveaux, fonction valeur , analyse non régulière.

Application of bilevel programming to optimal control problem

Abstract : We are interested here in the study of bilevel optimization problem which is an hierarchical problem at two levels of decision : an upper level called *leader* and a lower level called *follower*. The optimal solution depends of course on the two levels who are in an hierarchical interdependence.

The second part of this paper treats the bilevel optimal control problem. To obtain an optimality condition, we reformulate the bilevel problem as a single level optimal control problem, and we involve the value function of the lower level problem. The study realized in this paper uses the tools of nonsmooth analysis.

Key words : bilevel mathematical programming, bilevel optimal control, value function, nonsmooth analysis.