

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université de Batna 2
Faculté des Mathématiques et de l'Informatique
Département de Mathématiques

Thèse

Présentée en vue de l'obtention du Diplôme de

DOCTORAT EN SCIENCES

Option : Mathématiques appliquées

Par

Bounibane Bachir

Thème

**Extension de quelques méthodes de points intérieurs
pour un problème d'optimisation**

Soutenue le : **03/10/2019**

Devant le jury

Président :	Dr. Said Guedjiba	Pr. Université de Batna 2
Rapporteur :	Dr. El Amir Djefal	M.C.A. Université De Batna 2
Examineurs :	Dr. Menkad Safa	M.C.A. Université de Batna 2
	Dr. Zouhir Mokhtari	Pr. Université Mohamed Kheidar. Biskra
	Dr. Yahia Djabrane	M.C.A. Université Mohamed Kheidar. Biskra
	Dr. Nacer Khelil	M.C.A. Université Mohamed Kheidar. Biskra

Table des matières

LIST OF TABLES	5
LIST OF FIGURES	6
Introduction	7
I PRÉLIMINAIRES ET NOTIONS FONDAMENTALES	18
1.1 Notations, définitions et résultats utiles	18
1.1.1 Produit scalaire et normes	18
1.2 Eléments d'analyse convexe	20
1.2.1 Ensembles affines	20
1.2.2 Fonctions affines	20
1.2.3 Ensembles convexes	21
1.2.4 Fonctions convexes	21
1.2.5 Fonction mid-convexe	22
1.2.6 Caractérisation d'une fonction convexe différentiable	22
1.3 Programmation mathématique (<i>PM</i>)	23
1.3.1 Classification et résolution d'un programme mathématique	24
1.3.2 Qualification des contraintes	25
1.3.3 Contraintes et Lagrangien	25
1.3.4 Existence et unicité d'une solution optimale d'un programme mathématique	26
1.3.5 Conditions d'optimalité	27
II PROGRAMMATION LINÉAIRE ET MÉTHODES DE POINTS INTÉRIEURS	30
2.1 Programmation linéaire (<i>PL</i>)	30
2.1.1 Représentation matricielle	31
2.1.2 Formes usuelles d'un programme linéaire	32
2.1.3 Dualité Lagrangienne et conditions K.K.T	33
2.1.4 Propriétés de la dualité	35

2.1.5	La méthode de Newton	37
2.2	Les méthodes de barrières (<i>Chemin central</i>)	38
2.2.1	Barrière primal-dual en programmation linéaire	39
2.2.2	La fonction de Lagrange associée au problème (PPL_μ)	41
2.2.3	Comparaison des conditions du problème (PPL_μ) et (PPL)	42
2.2.4	Le chemin central primal-dual	45
2.2.5	Concept de proximité	45
2.3	Différentes méthodes de points intérieures	48
2.3.1	Méthodes de suivi de chemin primale-duale	48
2.3.2	Méthode de suivi de chemin à pas court	51
2.3.3	Méthode de suivi de chemin primale-duale à pas longs.	54
2.4	Méthode de Trajectoire Centrale basée sur les fonctions noyaux pour PL	55
2.5	Fonctions noyau et propriétés	61
2.5.1	Qualification d'une fonction noyau	62
2.5.2	Propriétés et relation entre les conditions de qualification	64
2.5.3	Une borne supérieure de $\Psi(v)$ après chaque itération externe	65
2.5.4	Analyse de la décroissance de la fonction barrière de proximité Ψ	67
2.5.5	Borne d'itérations et analyse de la complexité	75
III MÉTHODES DE LA TRAJECTOIRE CENTRALE BASÉE SUR UNE NOUVELLE FONCTION NOYAU POUR PL		79
3.1	Fonction noyau paramétrique avec double barrière	79
3.1.1	Introduction	79
3.1.2	Nouvelles directions cherchées	81
3.1.3	Propriétés de fonction noyau	83
3.1.4	Eligibilité (Qualification) de la nouvelle fonction noyau	84
3.1.5	Analyse de la complexité	90
Conclusions et perspectives		95

Bibliographie 96

Liste des tableaux

Table des figures

Notations et terminologie

$(\mathcal{P})_I$:	Le problème non linéaire. initial
(PM)	:	Programmation mathématique.
(PL)	:	Programmation linéaire.
(PSD)	:	Programmation semi-définie.
(PPL)	:	Le problème linéaire primal.
(PDL)	:	Le problème dual de (PPL) .
$(CQSDP)$:	Le problème quadratique convexe semi-défini
$(MPIs)$:	Les méthodes de points intérieurs
(Tc)	:	Trajectoire centrale.
<i>Algo</i>	:	Algorithme.
(IPC)	:	Condition de points intérieurs.
<i>K.K.T</i>	:	Karush-Kuhn-Tucker

Opérations

\mathbb{R}^n : L'ensemble des vecteurs avec n composantes réelles.

\mathbb{R}^n : L'orthant positif de l'espace \mathbb{R}^n .

$x \geq 0$: Les composantes de x , $x_i \geq 0$ pour tout i .

$x > 0$: Les composantes de x , $x_i > 0$ pour tout i .

x^t : Transposé du vecteur x de \mathbb{R}^n .

$X \succeq 0$: X est une matrice symétrique semi-définie positive.

$X \succ 0$: X est une matrice symétrique définie positive.

xs = $(x_1s_1, x_2s_2, \dots, x_ns_n)$ (produit d'Hadamard).

$\frac{x}{s}$ = $(\frac{x_1}{s_1}, \frac{x_2}{s_2}, \dots, \frac{x_n}{s_n})^T : (s \neq 0)$.

\sqrt{x} = $(\sqrt{x_1}, \dots, \sqrt{x_n})^T : (x \geq 0)$.

x^{-1} = $(\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n})^T : (x \neq 0)$.

$\langle x, s \rangle$ = $x^T s = \sum_{k=1}^n x_k s_k$: (le produit scalaire de deux vecteurs).

$\|x\|_\infty$ = $\max_{1 \leq k \leq n} |x_k|$: (norme infinie).

$$\|x\|_2 = \sqrt{\sum_{k=1}^n x_k^2} : (\text{norme euclidienne}).$$

Vecteurs

Les vecteurs sont désignés par des lettres minuscules. Si x est un vecteur de \mathbb{R}^n , on désigne par :

x^T : le vecteur transposé de x .

x_i : la i -ème composante de x .

x_k : le k -ème vecteur d'une suite de vecteurs.

e : le vecteur de \mathbb{R}^n , dont toutes les composantes sont égales à 1.

Des vecteurs sont libres c.à.d. sont linéairement indépendants.

Matrices

Les matrices sont désignées par des lettres majuscules. Si A est une matrice, on désigne par.

A^T : La matrice transposée de A .

X = $diag(x)$, la matrice diagonale X avec $X_{ii} = x_i$.

I = $diag(e)$, la matrice identité d'ordre n .

X^{-1} = $diag(x^{-1})$, l'inverse de la matrice X avec $X_{ii}^{-1} = \frac{1}{x_i} : (x > 0)$

A est de plein rang : A est de rang $m : (m \leq n) : si ses lignes sont libres.$

$\mathbb{R}^{n \times n}$:, est Symétrique : si $A = A^T$, c.à.d: $a_{ij} = a_{ji} : pour tout i = 1, \dots, n$ et $j = 1, \dots, n$

Fonctions

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, une fonction différentiable à plusieurs variables x_1, \dots, x_n .

Alors :

$\nabla f(x)$ = $(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_n}(x))$, : (gradient de f au point x).

$\frac{\partial f}{\partial x_i}(x)$: les dérivées partielles au point x .

$\nabla^2 f(x)$ = $(\frac{\partial^2 f}{\partial x_i \partial x_j}(x))_{1 \leq i, j \leq n}$, la matrice Hessienne.

Introduction

Les méthodes de points intérieurs sont apparues dans les années cinquante. Ces méthodes reposent sur la fonction barrière logarithmique définie, en 1955, par Frisch [16], pour résoudre des problèmes non linéaires. Cette fonction barrière logarithmique sera largement étudiée dans les années soixante, notamment dans le livre de Fiacco et McCormick [15]. C'est dans ce livre que le terme de points intérieurs sera introduit. Les itérés générés par les méthodes basées sur cette fonction sont strictement réalisables : ils restent à l'intérieur du domaine réalisable, d'où la terminologie de ces méthodes.

L'idée générale de ces méthodes consiste à approcher d'une manière itérative la solution des problèmes d'optimisation avec contraintes d'inégalité de la forme

$$\begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ g(x) \geq 0, \end{cases} \quad (\mathcal{P}_I)$$

à partir de l'intérieur de l'ensemble réalisable $\mathcal{K} = \{x : g(x) \geq 0\}$. Ces méthodes sont basées sur les méthodes barrières. Dans la littérature, quelques exemples de fonctions ont été proposés (par exemple la fonction inverse, la fonction quadratique inverse), mais la fonction barrière logarithmique, proposée par Frisch [16], reste l'exemple le plus répandu.

La fonction barrière logarithmique associée au problème (\mathcal{P}_I) est donnée par

$$B(x, \mu) = f(x) - \mu \sum_{i=1}^n \log(g_i(x))$$

où $\mu > 0$ est le paramètre barrière. La méthode barrière logarithmique consiste à trouver une solution du problème (\mathcal{P}_I) en résolvant une suite de problèmes sans contrainte

$$\min_{x \in \mathbb{R}^n} B(x, \mu), \quad (1)$$

en conduisant le paramètre barrière vers zéro. Les conditions d'optimalité du premier ordre de ce problème sont données par

$$\nabla f(x) - \sum \frac{\mu}{g_i(x)} \nabla g_i(x) = 0.$$

Une étude extensive de cette méthode a été établie par Fiacco et McCormick [15]. Ces derniers ont montré que si x^* est une solution du problème (\mathcal{P}_I) et $x^*(\mu)$ est une solution du problème (1), alors sous des hypothèses standards, on a

$$\lim_{\mu \rightarrow 0} x^*(\mu) = x^* \text{ et } \lim_{\mu \rightarrow 0} \frac{\mu}{g_i(x^*(\mu))} = s^*,$$

ou s^* est le multiplicateur de Lagrange optimal associé aux contraintes d'inégalité.

Avant 1984, tout problème d'optimisation linéaire se résolvait par la méthode du simplexe développé par Dantzig [13] ou par une variante de celle-ci. Des recherches ont été menées pour mettre au point une autre méthode mais aucune de celles proposées n'améliorait celle du simplexe. Aussi, pendant une quarantaine d'années, cette méthode domina l'optimisation linéaire. Puis, dans les années 70, la théorie de la complexité devint une partie intégrante de l'optimisation linéaire, si bien qu'on demanda aux méthodes développées de converger en un temps polynômial, c'est à dire de résoudre le problème en un nombre d'opérations qui doit être borné par un polynôme en la taille du problème. Mais la méthode du simplexe n'a pas cette propriété, comme l'ont montré Klee et Minty [27]. On se demanda alors si un algorithme d'optimisation linéaire avait cette propriété.

En 1979, le mathématicien soviétique L. G. Khachiyan [25] a appliqué la méthode de l'ellipsoïde (initialement introduite par N. Shor en 1970 [45]) à la programmation linéaire et a prouvé que sa complexité algorithmique est de type polynomial. Cette découverte représente un succès théorique, cependant en pratique l'algorithme est loin de rivaliser la méthode du simplexe.

La révolution des méthodes de points intérieurs a démarré en 1984. Karmarkar

[24] découvre une méthode de points intérieurs (méthode projective) de complexité polynomiale plus efficace en pratique que celle de Khachiyan. Il annonce également des performances supérieures à celles de l'algorithme simplexe dans le cas où les matrices sont creuses et de grandes dimensions. Depuis, les méthodes de points intérieurs ont connu un développement exceptionnel donnant lieu à un changement radical dans l'étude théorique et la résolution numérique des programmes mathématiques linéaires. On parvient même à résoudre convenablement des problèmes non traités auparavant.

Ces programmes rivalisant n'échappent pas tout de même aux inconvénients théoriques, algorithmique et numérique, représentés par le problème d'initialisation qui devient de plus en plus difficile et la complexité algorithmique, devenue très importante pour l'étude de la vitesse de convergence des algorithmes développés.

Le succès des méthodes de points intérieurs pour la résolution des problèmes de programmation linéaire entraîna son extension aux problèmes de programmation non linéaire [50]. Et en particulier aux problèmes de programmation quadratique convexes [33, 48] ainsi qu'aux problèmes non linéaires convexes [3, 49, 56]. Ces études ne contredirent pas l'efficacité des méthodes de points intérieurs du type primal-dual observée pour le cas linéaire. Il en découla l'envie de généraliser cette approche aux problèmes non linéaires et non convexes. El-Bakry et al. [14], McCormick et Falk [31], Tits et al. [47] et Yamashita [57] ont développé des algorithmes du type primal-dual à convergence globale pour les problèmes non linéaires non convexes. Enfin, on peut encore classer les méthodes de point intérieur selon le type d'algorithme qu'elles emploient. A cet effet, on distingue quatre catégories :

- a) **Les méthodes de mise à l'échelle affine.**
- b) **Les méthodes projectives**
- c) **Les méthodes de réduction de potentiel.**
- e) **Les méthodes de suivi de chemin.**

Les méthodes de points intérieurs ont fait récemment l'objet de plusieurs monographies dont Roos, Terlaky et Vial [42], Wright [52], Ye [58] J. Renegar [43] et Nesterov et Nemirovsky [36]. Un excellent survol des méthodes de points intérieurs pour la programmation linéaire est celui de Gonzaga [20].

Depuis l'apparition des méthodes de points intérieurs initialisées par Karmarkar en 1984 [24], de nombreuses méthodes ont été proposées pour la programmation linéaire et ont ensuite été généralisées pour la programmation semi-définie.

La programmation semi-définie (*SDP*) est l'un des problèmes d'optimisation. L'étude de ce genre de problèmes a connu un fantastique regain d'intérêt et la plupart des documents sur (*SDP*) ont été écrits dans les années 90, bien que ses racines remontent à quelques décennies de plus, (voir par exemple Bellman et Fan [9]).

Les premiers travaux de (*MPIs*) pour (*SDP*) ont été développés indépendamment par Alizadeh [1] et Nesterov et Nemirovsky [36]. En 1995, Alizadeh [1] a proposé une méthode de réduction du potentiel de Ye pour (*SDP*). Presque en même temps, en 1994, Nesterov et al [36] ont prouvé que (*MPIs*) sont capables de résoudre les problèmes d'optimisation conique généraux, en particulier les problèmes (*SDP*), en un temps polynomial. En 2002, Halicka et al [21] ont proposé une étude sur la convergence de la méthode de trajectoire centrale en programmation semi-définie. En 1997, Monteiro [34], a proposé une méthode de trajectoire centrale primale-duale. En 1999, Monteiro et al [32, 35] ont proposé une méthode primale-duale de type trajectoire centrale et ont prouvé que la complexité de son algorithme est polynomiale.

Les fonctions noyaux jouent un rôle important dans la conception de nouveaux algorithmes de points intérieurs de type primal-dual. Les méthodes de points intérieurs basées sur la notion des fonctions noyaux ont fait récemment l'objet de plusieurs études citons, [4, 7, 10, 12, 17, 18, 19, 26, 38, 39, 40, 41]

Notre travail apporte sur des contributions théoriques, et algorithmiques pour résoudre des problèmes d'optimisation linéaire. On s'intéresse plus particulièrement,

aux méthodes de problèmes, linéaire basée sur une nouvelle fonction noyau qualifiée.

Premièrement, Peng et al. [39] ont présenté des méthodes de points intérieurs pour (PL) et (SDP) basées sur les fonctions barrières auto-régulière, puis, Bai et al. [4, 6, 7], ont proposé une classe des méthodes de points intérieurs de type primal-dual pour (PL) basé sur une variété de fonctions noyaux non-auto-régulière et ils ont obtenu la même complexité favorable pour les algorithmes à grand et petit-pas que [39]. De plus, Wang et all ont fait l'extension des résultats ci-dessus obtenus pour (PL) à (SDP) [52] et $(CQSDP)$ [53].

Pour d'autre algorithmes de points intérieurs basé sur les fonctions noyaux on cite [8, 15, 51, 52, 60, 62]. Notons que la meilleure borne d'itérations pour les algorithmes à grand pas qui est d'ordre $\mathcal{O}(\sqrt{n} \log n(\log \frac{n}{\varepsilon}))$, a été obtenue jusqu'à présent

L'objectif de cette thèse est :

1. Proposer une méthode de résolution de la programmation lineaire (PL) par la méthode de trajectoire centrale, via une nouvelle fonction noyau et on prend en considération que notre fonction est qualifiée.
2. L'introduction d'une nouvelle fonction noyau paramétrée définie par

$$\psi(t) = t^2 - 1 - \log(t) + \frac{t^{-p} - 1}{p}, \quad p > 0$$

Nous avons montré que l'algorithme correspondant converge après

$$\mathcal{O}\left(pn^{\frac{p+1}{2p}} \log \frac{n}{\varepsilon}\right) \text{ itérations pour les méthodes à long-pas.}$$

Et $\mathcal{O}(p\sqrt{n} \log \frac{n}{\varepsilon})$ itérations pour les méthodes à court-pas.

3. Sous un choix spécial du paramètre p ($p = \frac{\log n}{2} - 1$) on montre que la complexité d'un algorithme à grand pas qui est basée sur cette fonction noyau se réduit par un facteur de $\mathcal{O}(\log n)$ c'est-à-dire on obtient la meilleure complexité connue jusqu'à présent pour ces algorithmes à grand pas qui est de l'ordre $\mathcal{O}(\sqrt{n} \log n(\log \frac{n}{\varepsilon}))$, (voir par exemple [37]). Nous montrons aussi que

pour $p = 1$, nous obtenons la meilleure complexité algorithmique, obtenue jusqu'à présent, pour les méthodes à court-pas à savoir $\mathcal{O}(\sqrt{n} \log \frac{n}{\varepsilon})$.

4. On peut faire l'extension des résultats obtenus pour (PL) à plusieurs problèmes d'optimisation tels que le problème (SDP), le problème quadratique convexe semi-défini ($CQSDP$) et les problèmes de complémentarité linéaire. On note qu'ils y a des algorithmes similaires qui sont prolongés avec succès à un problème (SDP). On cite par exemple [1, 2, 22, 23, 28, 34, 38, 49, 51, 52, 53, 54, 55, 59, 60, 61]. Notons que la meilleure borne d'itérations pour les algorithmes à grand pas qui est d'ordre $\mathcal{O}(\sqrt{n} \log n (\log \frac{n}{\varepsilon}))$, a été obtenue jusqu'à présent par quelques fonctions noyaux barrières paramétrées dans le tableau suivant :

La fonction noyau	Paramètre	Références
$\frac{t^2 - 1}{2} - \frac{t^{1-q} - 1}{q - 1}, q > 1$	$q = \frac{1}{2} \log n$	[38, 39]
$\frac{t^2 - 1}{2} - \frac{t^{1-q} - 1}{q(q-1)} - \frac{q-1}{q} (t-1), q > 1$	$q = \frac{1}{2} \log n$	[38, 39]
$\frac{t^2 - 1}{2} - \frac{e^{\sigma(1-t)} - 1}{\sigma}, \sigma > 0$	$\sigma = \mathcal{O}(\log n)$	[8]
$\frac{t^2 - 1}{2} + \int_1^t e^{q(\frac{1}{\xi} - 1)} d\xi, q \geq 1$	$q = \mathcal{O}(\log(n+1))$	[6]
$t^2 - 1 + \frac{t^{-p} - 1}{p} - \log t, p > 0$	$p = \frac{\log n}{2} - 1$	[11]

Cette thèse est composée de trois chapitres et est organisée comme suit :

Le chapitre 1 On présente une introduction de certaines notions et résultats qui seront utiles dans la suite, à savoir

- L’analyse convexe.
- La programmation mathématique.
- Rappels de base portant sur l’optimisation, les conditions de qualifications de contraintes ainsi que sur les résultats d’existence d’une solution optimale
- Les conditions d’optimalité

Le chapitre 2 Dans ce chapitre, nous présentons des algorithmes de points intérieurs pour résoudre des problèmes de programmation linéaires, nous allons aussi présenter une synthèse théorique et une étude algorithmique pour la résolution d’un programme linéaire primal-dual par les méthodes de points intérieurs basées sur la méthode de trajectoire centrale. Nous avons introduit et apporté un ensemble de règles d’ordre théorique et algorithmique sur la méthode de trajectoire centrale via une fonction noyau.

La présentation de ces algorithmes donne l’occasion de rappeler des outils utiles pour l’étude des problèmes d’optimisation : dualité Lagrangienne, conditions $K.K.T$, directions de descentes et méthodes de barrière

Le chapitre 3 Dans le troisième chapitre, on développe un algorithme de points intérieurs de type primal-dual basé sur une nouvelle fonction noyau qui doit être qualifiée. En outre, nous formulons un algorithme de points intérieurs primal-dual pour (PL) en utilisant la fonction de proximité et de donner son analyse de la complexité. Enfin, ce manuscrit est achevé par une conclusion et une bibliographie.

Chapitre I

PRÉLIMINAIRES ET NOTIONS FONDAMENTALES

Dans ce chapitre, introductif nous allons introduire quelques notions fondamentales et résultats bien connus qui permet de faciliter la compréhension de cette thèse, notamment les normes, ainsi quelques notions de base de l'analyse convexe et de la programmation mathématique qui seront utilisées par la suite. Nous présentons aussi certains éléments fondamentaux des méthodes de points intérieurs pour la résolution de la programmation linéaire (*PL*) par la méthode de trajectoire centrale

1.1 Notations, définitions et résultats utiles

1.1.1 Produit scalaire et normes

On commence par la définition du produit scalaire de deux vecteurs.

Définition 1 *Étant donné deux vecteurs $x, y \in \mathbb{R}^n$, le produit scalaire euclidien est noté et défini par*

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i.$$

Et de la norme euclidienne associée

$$\|x\| = (x^T x)^{\frac{1}{2}} = \sqrt{\langle x, x \rangle}.$$

Définition 2 *La boule ouverte de centre x et de rayon ε est définie par*

$$\mathcal{B}(x, \varepsilon) = \{y \in \mathbb{R}^n : \|x - y\| < \varepsilon\}$$

Pour étudier les propriétés de convergence des méthodes proposées, basée sur les fonctions noyaux on va utiliser les notations de Landau.

Définition 3 (Notation grand $\mathcal{O}(\cdot)$) Soient f et $g : \mathbb{N} \longrightarrow \mathbb{R}^+$ On note $f = \mathcal{O}(g)$ ssi $\exists c \in \mathbb{R}^+, \exists n_0 \in \mathbb{N}$ tel que:

$$\forall n > n_0, f(n) \leq cg(n).$$

Autrement dit : $f(n)$ est en $\mathcal{O}(g(n))$ s'il existe un seuil n_0 à partir duquel la fonction $f(\cdot)$ est toujours dominée par la fonction $g(\cdot)$, à une constante multiplicative fixée près

Remarque 4 La notation \mathcal{O} est utilisée pour indiquer la borne supérieure asymptotique

De même on définit :

Définition 5 (Notation Θ) Soient f et $g : \mathbb{N} \longrightarrow \mathbb{R}^+$ On note $f = \Theta(g)$ $\exists c_1, c_2 \in \mathbb{R}^+, \exists n_0 \in \mathbb{N}$ tel que :

$$\forall n > n_0, c_1g(n) \leq f(n) \leq c_2g(n).$$

Définition 6 (Notation o) On note $f(n) = o(g(n))$ lorsque pour tout réel c , il existe un entier n_0 tel que pour tout $n \geq n_0$,

$$f(n) \leq cg(n).$$

Définition 7 Le coût d'un algorithme est le nombre d'opérations arithmétiques intervenant dans la réalisation de cet algorithme. La notation $\mathcal{O}(g(N))$, signifie que le nombre d'opérations arithmétiques est borné par la quantité $cg(N)$ où c est une constante positive, c'est à dire

$$\mathcal{O}(g(N)) \leq cg(N).$$

N est la taille de la mémoire nécessaire pour stocker les données, et g est une fonction quelconque. Autrement l'algorithme Algo est de complexité polynômiale c.à.d résoudre le problème en un nombre d'opérations qui doit être borné par un polynôme en fonction de la taille du problème.

1.2 *Éléments d'analyse convexe*

La notion de convexité est un outil mathématique important pour l'étude théorique et numérique des problèmes d'optimisation. On présente dans cette section quelques notions d'analyse convexe d'usage courant. Pour plus de détails voir [44].

1.2.1 Ensembles affines

Définition 8 *Un sous ensemble F de \mathbb{R}^n est dit affine si :*

$$\forall x, y \in F, \forall \lambda \in \mathbb{R} : \lambda x + (1 - \lambda)y \in F.$$

On dit aussi que F est une variété affine (ou linéaire)

Les ensembles affines élémentaires sont : \emptyset , $\{x\}$ ($x \in \mathbb{R}^n$), et chaque sous-espace vectoriel de \mathbb{R}^n .

Définition 9 *On appelle combinaison affine des éléments x_1, \dots, x_m de \mathbb{R}^n tout élément de la forme :*

$$x = \sum_{i=1}^m \lambda_i x_i, \text{ avec } \lambda_i \in \mathbb{R} \text{ et } \sum_{i=1}^m \lambda_i = 1.$$

Théorème 10 *Toute partie affine de \mathbb{R}^n contient ses combinaisons affines :*

$$\forall x_i \in F = \sum_{i=1}^m \lambda_i x_i, \text{ avec } \lambda_i \in \mathbb{R} \text{ et } \sum_{i=1}^m \lambda_i = 1.$$

1.2.2 Fonctions affines

Définition 11 *Une fonction f définie de \mathbb{R}^n dans \mathbb{R}^m est dite affine si*

$$f[\lambda x + (1 - \lambda)y] = \lambda f(x) + (1 - \lambda)f(y), \forall x, y \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}.$$

Théorème 12 *Toute fonction affine f de \mathbb{R}^n dans \mathbb{R}^m est de la forme $f(x) = Ax + b$;*

où

$$A \in \mathcal{M}_{m,n}(\mathbb{R}) \text{ et } b \in \mathbb{R}^m.$$

Proposition 13 Une fonction affine f de \mathbb{R}^n dans \mathbb{R}^m est linéaire si et seulement si $f(0) = 0$, i.e., $b = 0$.

1.2.3 Ensembles convexes

Définition 14 Un ensemble $S \subseteq \mathbb{R}^n$ est appelé convexe si le segment joignant toute paire de points appartenant à S appartient également entièrement à S . Algébriquement, on demande donc que $x \in S$ et $y \in S$ entraîne que,

$$\lambda x + (1 - \lambda)y \in S \text{ pour tout } \lambda \in [0, 1].$$

Définition 15 S est un polyèdre convexe s'il est de la forme :

$$S = \{x \in \mathbb{R}^n : a_i^t x \leq b_i, i = 1, \dots, m\},$$

où a_i est un vecteur non nul de \mathbb{R}^n et b_i un scalaire pour $i = 1, \dots, m$

Remarque 16 S peut s'écrire sous la forme matricielle suivante :

$$S = \{x \in \mathbb{R}^n : Ax \leq b\},$$

où A est une matrice de $\mathbb{R}^{m \times n}$ et b un vecteur de \mathbb{R}^m .

1.2.4 Fonctions convexes

Définition 17 La fonction $f : S \rightarrow \mathbb{R}$ est dite convexe équivaut à dire que

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \forall \lambda \in [0, 1], \quad \forall x, y \in S$$

ou d'une manière équivalente : si $\forall m \in \mathbb{N}^*, \forall \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1, x_i \in S$

$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i),$$

et si l'inégalité au dessus est stricte, alors f est dite strictement convexe $\forall x \neq y$:

1.2.5 Fonction mid-convexe

Définition 18 Soit f une fonction réelle de la variable réelle définie sur un intervalle I non réduit à un point. On dit que « la fonction f est mid-convexe sur S » si on a :

$$\forall (x, y) \in S^2, f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}.$$

► Soit f une fonction réelle de la variable réelle définie sur S non réduit à un point.

Si f est continue et mid-convexe sur S alors f est convexe sur I

1.2.6 Caractérisation d'une fonction convexe différentiable

Si $f \in C^1(S)$, où S est un ensemble convexe alors on a les équivalences suivantes :

1. f est convexe si et seulement si

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle, \forall x, y \in S$$

2. f est convexe si et seulement si

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0, \forall x, y \in S$$

3. De plus, f est dite strictement convexe si l'une ou l'autre des inégalités précédentes sont strictes pour $x \neq y$.

4. f est convexe si et seulement si $\nabla^2 f(x)$ (le Hessien de f) est semi-défini positif sur S (c'est à dire que $y^t \nabla^2 f(x) y \geq 0, \forall x, y \in S$ ou encore toutes les valeurs propres de $\nabla^2 f(x)$ sont positives).

5. f est strictement convexe si et seulement si $\nabla^2 f(x)$ est défini positif sur S (c'est à dire que $y^t \nabla^2 f(x) y > 0, \forall x, y \in S$ et $y \neq 0$ ou encore toutes les valeurs propres de $\nabla^2 f(x)$ sont strictement positives).

1.3 Programmation mathématique (PM)

Débutons par un bref rappel des conditions d'optimalité du premier ordre pour un programme mathématique (PM) général (non linéaire) avec des contraintes d'égalité et d'inégalité.

Considérons à présent un problème ayant des contraintes d'égalité et d'inégalité

$$\begin{cases} \min_{x \in \mathcal{F}} f(x) \\ \text{ou } \mathcal{F} = \{x \in \mathbb{R}^n, g_i(x) \leq 0, \text{ pour } i \in I, h_j(x) = 0, \text{ pour } j \in J\} \end{cases} \quad (PM)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est appelée fonction objectif de (PM) et \mathcal{F} est un sous-ensemble non vide de \mathbb{R}^n défini par des contraintes $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Ainsi la programmation non linéaire se présente comme étant une généralisation de la programmation quadratique et de la programmation linéaire.

L'ensemble \mathcal{F} est appelé l'ensemble des solutions réalisables ou ensemble des contraintes. Tout point $x \in \mathbb{R}^n$ vérifiant : $x \in \mathcal{F}$, est appelé point admissible du problème (PM). Les ensembles finis $I = \{1, \dots, m\}$ et $J = \{1, \dots, p\}$, représentent respectivement les indices des contraintes d'inégalité et égalité. On note par \mathcal{I} , \mathcal{J} l'ensemble d'indices des contraintes d'inégalités (respectivement égalités) actives

$$\mathcal{I} = \{i : g_i(x) \leq 0, i = 1, \dots, m\}, \quad \mathcal{J} = \{j : h_j(x) = 0, j = 1, \dots, p\}.$$

Pour le problème (PM), on dit que x^* est

- **Solution réalisable (faisable) :** On appelle solution réalisable de (PM) tout point vérifiant les contraintes (i.e., $x^* \in \mathcal{F}$)
- **Solution optimale globale :** Une solution réalisable qui optimise l'objectif sur \mathcal{F} est dite solution optimale globale de (PM) (i.e., $x^* \in \mathcal{F}$ est un minimum global de f si et seulement si pour tout $x \in \mathcal{F}$, $f(x^*) \leq f(x)$)

L'ensemble des solutions optimales globales est noté :

$$\arg \min f(x) \quad (\text{L'argument du minimum})$$

- **Solution optimale locale :** Un point $x^* \in \mathcal{F}$ est une solution optimale locale de (PM) s'il existe une boule centrée en x^* de rayon $\varepsilon > 0$ telle que pour tout $x \in \mathcal{F} \cap \mathcal{B}(x^*, \varepsilon)$ on a $f(x^*) \leq f(x)$.
- Un minimum local strict de f si et seulement s'il existe une boule centrée en x^* de rayon $\varepsilon > 0$ telle que pour tout $x \in \mathcal{F} \cap \mathcal{B}(x^*, \varepsilon)$ on a $f(x^*) < f(x)$ avec $x \neq x^*$.

L'ensemble des solutions optimales locales de (PM) est noté :

$$\underset{\mathcal{F}}{\text{loc min}} f(x) :$$

1.3.1 Classification et résolution d'un programme mathématique

On classe le problème (PM) à partir de deux propriétés fondamentales à savoir la convexité et la différentiabilité de la fonction objectif et les contraintes.

A ce propos,

1. (PM) est convexe si les fonctions f, g_i sont convexes et les fonctions h_j sont affines.
2. (PM) est différentiable si les fonctions f, g_i, h_j sont différentiables.
3. (PM) est linéaire (f linéaire, g_i, h_j affines).

La convexité et la différentiabilité jouent un rôle très important dans la programmation mathématique.

Définition 19 Soit une contrainte inégalité $g_i(x) \leq 0$ Si x^* satisfait $g_i(x^*) < 0$, on dit que la contrainte est inactive en x^* . Si x^* satisfait $g_i(x^*) = 0$, on dit que la contrainte est active ou saturée en x^*

Remarque 20 Une contrainte d'égalité $h_j(x) = 0$ est par définition est active en tout point $x \in \mathcal{J}$

1.3.2 Qualification des contraintes

La condition de qualification est satisfaite en tout $x^* \in \mathcal{F}$ dans les cas suivants :

1. Les contraintes sont affines.,
2. Les gradients des contraintes saturées en x^* sont libres .
3. \mathcal{F} est convexe et $\text{int}(\mathcal{F}) \neq \emptyset$ (**condition de Slater**).

1.3.3 Contraintes et Lagrangien

Avec ces définitions, nous pouvons énoncer maintenant les conditions d'optimalité du premier ordre ou conditions de Karush, Kuhn et Tucker (*K.K.T*). On associe au problème (*PM*) la fonction lagrangienne

Définition 21 (Lagrangien) *On appelle Lagrangien associé à problème (PM) la fonction*

$$\begin{aligned} \mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m &\longrightarrow \mathbb{R} \\ (x, \lambda, \mu) &\longmapsto \mathcal{L}(x, \lambda, \mu) \end{aligned}$$

$$\text{tel que } \mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{j \in J} \lambda_j h_j(x) + \sum_{i \in I} \mu_i g_i(x)$$

Les vecteurs $\lambda = (\lambda_j)_{j \in J}$, et $\mu = (\mu_i)_{i \in I}$ sont appelés *multiplicateurs de Lagrange*.

- Lorsque f, h_1, \dots, h_p et g_1, \dots, g_m sont différentiables, le vecteur gradient du Lagrangien en $x \in \mathbb{R}^n$

$$\begin{aligned} \nabla \mathcal{L}_x(x, \lambda, \mu) &= \begin{pmatrix} \frac{\partial \mathcal{L}_x(x, \lambda, \mu)}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial \mathcal{L}_x(x, \lambda, \mu)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f(x, \lambda, \mu)}{\partial x_1} + \sum_{j=1}^p \lambda_j \frac{\partial h_j(x)}{\partial x_1} + \sum_{i=1}^m \mu_i \frac{\partial g_i(x)}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial f(x, \lambda, \mu)}{\partial x_n} + \sum_{j=1}^p \lambda_j \frac{\partial h_j(x)}{\partial x_n} + \sum_{i=1}^m \mu_i \frac{\partial g_i(x)}{\partial x_n} \end{pmatrix} \\ &= \nabla f(x) + \sum_{j=1}^p \lambda_j \nabla h_j(x) + \sum_{i=1}^m \mu_i \nabla g_i(x) \end{aligned}$$

La condition de Lagrange s'écrit alors, sous les hypothèses adéquates :

$$x \text{ est un extremum local} \implies \exists! (\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^m, \nabla \mathcal{L}_x(x, \lambda, \mu) = 0$$

1.3.4 Existence et unicité d'une solution optimale d'un programme mathématique

Dans ce paragraphe, nous nous intéresserons aux théorèmes qui permettent d'assurer qu'il existe un minimum à un problème d'optimisation.

Considérons à nouveau le problème (PM) :

$$\min f(x) \text{ sous la contrainte } x \in \mathcal{F}, \mathcal{F} \text{ un ensemble non vide de } \mathbb{R}^n.$$

Il existe principalement deux théorèmes donnant des conditions suffisantes d'existence d'un point de minimum : le premier dans le cas où l'ensemble des contraintes est fermé borné, le second pour un ensemble de contraintes fermé mais non borné.

a) Cas où l'ensemble \mathcal{F} des contraintes est borné

Théorème 22 (Théorème de Weierstrass) *Soit \mathcal{F} un ensemble compact non vide de \mathbb{R}^n et $f : \mathcal{F} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application continue sur \mathcal{F} . Alors f est bornée et atteint ses bornes. Autrement dit, f admet au moins un point $x^* \in \mathcal{F}$ de minimum global de f sur \mathcal{F} :*

$$\forall y \in \mathcal{F}, f(x^*) \leq f(y),$$

et au moins un point $x^ \in \mathcal{F}$ de maximum global de f sur \mathcal{F} :*

$$\forall y \in \mathcal{F}, f(x^*) \geq f(y).$$

b) Cas où l'ensemble \mathcal{F} des contraintes est non borné

Définition 23 (Applications coercives) *Une application $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ continue est coercive, i.e. infinie à l'infini ssi*

$$\forall A \in \mathbb{R}, \exists k > 0 \text{ tel que } \forall x \in \mathcal{F}, \text{ si } \|x\| \geq k \text{ alors } f(x) \geq A.$$

On note : $\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$,
 (souvent $\mathcal{D} = \mathbb{R}^n$)

Théorème 24 Soient F un fermé non vide de \mathbb{R}^n et $f : F \rightarrow \mathbb{R}$, une application continue et coercive sur F , alors (PM) admet une solution optimale globale.

Théorème 25 Si F est convexe et f est strictement convexe, alors il existe au plus une solution optimale de (PM).

c) Convexité

Théorème 26 Soit x^* un point de minimum local d'un problème de minimisation.

a) Si le problème est convexe, alors x^* est un point de minimum global.

b) Si le problème est strictement convexe, alors x^* est l'unique point de minimum global.

1.3.5 Conditions d'optimalité

Dans tout ce qui suit, on suppose que la condition de qualification des contraintes est vérifiée.

Une qualification des contraintes est simplement une hypothèse qui assure l'existence de ces multiplicateurs en x pour que les conditions de Karush-Kuhn-Tucker (K.K.T) soit satisfaite.

Cas de contraintes d'égalité

On suppose il n'y a que des contraintes d'égalité :

$$\mathcal{F} = \{x \in \mathbb{R}^n : h_j(x) = 0, \forall j \in J\}$$

Théorème 27 (Condition de Lagrange) Soit x^* un minimum local du problème

$$\begin{cases} \min f(x) \\ \text{s.c.} \\ h_j(x) = 0 \end{cases} .$$

Alors il existe des réels λ_i tels que

$$\nabla f(x^*) + \sum_{j=1}^p \lambda_j \nabla h_j(x^*) = 0.$$

Condition de K.K.T (Cas de contraintes d'égalité et d'inégalité)

En optimisation non linéaire avec contraintes, les conditions nécessaires de Karush-Kuhn-Tucker (conditions de *K.K.T*) sont les plus utilisées pour définir un minimum local.

Le théorème suivant généralise les conditions de Lagrange (théorème 28) au cas de contraintes égalitaires et inégalitaires.

Théorème 28 (Conditions nécessaires du premier ordre) Soit x^* un point admissible du problème (PM). Supposons f , g et h différentiables en x^* et les contraintes qualifiées au point x^* . Si x^* est un point de minimum local de f sur \mathcal{F} . Alors il existe $\lambda^* = (\lambda_1^*, \dots, \lambda_p^*) \in \mathbb{R}^p$ et $\mu^* = (\mu_1^*, \dots, \mu_m^*) \in \mathbb{R}^m$ tels que :

$$\left\{ \begin{array}{l} \nabla \mathcal{L}_x(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{j=1}^p \lambda_j^* \nabla h_j(x^*) + \sum_{i=1}^m \mu_i^* \nabla g_i(x^*) = 0 \\ \mu_i^* \geq 0, \quad \forall i \in \{1, \dots, m\} \\ h_j(x^*) = 0, \quad j = 1, \dots, p \\ g_i(x^*) \leq 0, \quad i \in 1, \dots, m \\ \mu_i g_i(x^*) = 0 \quad \forall i \in \{1, \dots, m\} \end{array} \right. \quad (2)$$

où μ_i et λ_j sont appelés les multiplicateurs de Lagrange. Ces conditions sont aussi connues sous le nom de conditions de Karush, Kuhn et Tucker (*K.K.T*). La condition $\mu_i g_i(x^*) = 0$, appelée condition de complémentarité. Nous dirons que la condition de complémentarité stricte est satisfaite lorsque

$$\mu_i > 0 \Leftrightarrow g_i(x^*) = 0, \quad \forall i \in \{1, \dots, m\}$$

Nos conditions nécessaires (2) signifient que le gradient de la fonction objectif peut être reconstruit par une combinaison linéaire des gradients des contraintes actives,

c'est-à-dire

$$\nabla f(x^*) = -\lambda \nabla h(x^*) - \mu \nabla g(x^*)$$

Remarque 29 - *Si les contraintes ne sont pas qualifiées en x^* , les conditions de K.K.T ne s'appliquent pas.*

- *Si (PM) est convexe, les conditions de K.K.T sont à la fois nécessaires et suffisantes pour que x soit un minimum global*

Chapitre II

PROGRAMMATION LINÉAIRE ET MÉTHODES DE POINTS INTÉRIEURS

2.1 *Programmation linéaire (PL)*

Nous étudions dans ce chapitre la programmation linéaire, c'est à dire la classe des problèmes d'optimisation où la fonction objectif et les contraintes sont toutes affines. Il s'agit des problèmes d'optimisation les plus célèbres, les plus simples théoriquement et les plus étudiés en recherche opérationnelle.

En termes d'optimisation, un programme linéaire (*PL*) consiste à minimiser (ou maximiser) une forme linéaire ou fonction coût (appelée objectif ou économique) soumise à des contraintes linéaires.

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une application linéaire. Soient $g_1, \dots, g_m : \mathbb{R}^n \rightarrow \mathbb{R}$, m applications linéaires, et $(b_1, \dots, b_m) \in \mathbb{R}^m$. Notons $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, un problème de programmation linéaire s'exprime sous la forme : trouver le minimum (respectivement le maximum) de $f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\min f(x), \text{ où } \max f(x)$$

soumis aux contraintes

$$\left\{ \begin{array}{l} g_1(x_1, \dots, x_n) \leq b_1 \\ g_2(x_1, \dots, x_n) \leq b_2 \\ \cdot \\ \cdot \\ \cdot \\ g_m(x_1, \dots, x_n) \leq b_m \end{array} \right.$$

f, g_1, g_2, \dots, g_m étant des formes linéaires sur \mathbb{R}^n , on peut toujours noter pour certains vecteurs $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, $a_i = (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n$, et $x = (x_1, \dots, x_n) \in \mathbb{R}^n$:

$$f(x) = \sum_{i=1}^n c_i x_i = \langle c, x \rangle, \quad g_i(x) = \sum_{j=1}^n a_{ij} x_j = \langle a_{ij}, x \rangle.$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire usuel de \mathbb{R}^n

2.1.1 Représentation matricielle

En écriture matricielle, un problème sous forme canonique s'écrit donc

$$(PL) \left\{ \begin{array}{l} \min \langle c, x \rangle = \min c^T x \\ \text{s.c. } Ax \leq b \end{array} \right.$$

$$A = (a_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = \begin{pmatrix} a_{11} & a_{1n} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ a_{m1} & a_{mn} \end{pmatrix} \in \mathcal{M}_{m,n}, \quad b = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

où $c = (c_1, \dots, c_n)^T$ et $x = (x_1, \dots, x_n)^T$ sont des vecteurs colonnes à n lignes, $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ est une matrice à n lignes et m colonnes, et $b = (b_1, \dots, b_m)^T$ est un vecteur colonne à m lignes avec $m \leq n$.

2.1.2 Formes usuelles d'un programme linéaire

Un programme linéaire peut être écrit sous l'une des deux formulations suivantes

1. Forme canonique

En programmation linéaire, un problème d'optimisation sous forme canonique est un problème sous la forme

$$\begin{cases} \min c^T x \\ Ax \leq b, \\ x \geq 0. \end{cases}$$

Forme standard

Un problème de programmation linéaire mis sous la forme standard peut être défini comme suit

$$\begin{cases} \min c^T x, \\ Ax = b, \\ x \geq 0, \end{cases}$$

où A est une matrice réelle plein rang de type (m, n) , b un vecteur de \mathbb{R}^m et $x \in \mathbb{R}^n$, tel que x une variable inconnue.

Dans toute la suite, on s'intéresse au problème (PL) sous forme standard suivant

$$(PL) \begin{cases} \min c^T x, \\ Ax = b, \\ x \geq 0, \end{cases} \quad (3)$$

nous appellerons domaine admissible (réalisable) \mathcal{F}_P qui est un polyèdre convexe fermé. L'ensemble des vecteurs x (solutions primales) qui satisfont les contraintes, c'est-à-dire

$$\mathcal{F}_P = \{x \in \mathbb{R}^n : Ax = b \text{ et } x \geq 0\},$$

tandis que l'ensemble associé \mathcal{F}_P^0 sera le sous-ensemble des points \mathcal{F}_P qui satisfont de manière stricte les contraintes de positivité (solutions primales strictement réalisables)

$$\mathcal{F}_P^0 = \{x \in \mathbb{R}^n : Ax = b \text{ et } x > 0\}.$$

Au problème (PL) , (désigné comme primal) on peut toujours lui associer un autre problème linéaire nommé problème dual (PD) . De plus, lorsque le primal cherche à minimiser sa fonction objectif, le dual est un problème de maximisation

2.1.3 Dualité Lagrangienne et conditions K.K.T

Le lagrangien du problème (PL)

On applique les idées des sections précédentes (dualité Lagrangienne et conditions $(K.K.T)$) au cas particulier des problèmes d'optimisation en Programmation Linéaire sous forme standard

$$(PL) \quad \min_{x \in \mathbb{R}^n} c^T x, \quad \text{telque} \quad \begin{cases} Ax = b, \\ x \geq 0, \end{cases}.$$

On rappelle que $x \geq 0$ signifie que $x_j \geq 0$ pour tout $j \in \{1, \dots, n\}$ quand $x = (x_j)$.

On suppose que l'ensemble de contrainte \mathcal{F}_P^0 est non vide. Ainsi la condition de Slater est vérifiée.

Le lagrangien du problème (PL) est la fonction

$$\begin{aligned} \mathcal{L} : \mathbb{R}^{n+m+n} &\longrightarrow \mathbb{R} \\ (x, y, s) &\longmapsto \mathcal{L}(x, y, s) \end{aligned}$$

définie par

$$\mathcal{L}(x, y, s) = c^T x - y^T (Ax - b) - s^T x$$

ou $y^T = (y_1, \dots, y_m)$ est le vecteur multiplicateur de Lagrange et $s = (s_1, \dots, s_n)$ est le vecteur de Lagrange associé à x

Les conditions d'optimalité du premier ordre de $K.K.T$ de ce problème sont données par

$$\begin{cases} c - A^T y - s = 0 \\ Ax = b \\ x \geq 0 \\ s \geq 0 \\ \text{diag}(s) \cdot x = 0 \end{cases}$$

On a

$$\begin{aligned} \mathcal{L}(x, y, s) &= c^T x - y^T (Ax - b) - s^T x = c^T x - y^T Ax + y^T b - s^T x \\ &= y^T b + (c - A^T y - s)^T x \end{aligned}$$

A l'optimum :

$$\mathcal{L}(x^*, y^*, s^*) = \underbrace{(c - A^T y^* - s^*)}_{=0}^T x^* + (y^*)^T b = (y^*)^T b$$

$$\begin{cases} \max_{y \in \mathbb{R}^m} b^T y \\ c - A^T y - s = 0 \\ \text{et } y \geq 0 \end{cases} \Leftrightarrow \begin{cases} \max_{y \in \mathbb{R}^m} b^T y \\ \text{avec } A^T y \leq c \end{cases}$$

Définition 30 Soit le programme linéaire primal écrit sous forme standard :

$$(PPL) \begin{cases} \min c^T x, \\ Ax = b, \\ x \geq 0, \end{cases}$$

On appelle problème dual de (PPL) , le programme linéaire

$$(DPL) \begin{cases} \max b^T y, \\ A^T y \leq c, \\ y \in \mathbb{R}^m, \end{cases}$$

Ce problème est fortement lié à (PPL) et sera pour cette raison appelé le problème dual de (PPL). On vérifie aisément que ce problème peut également être formulé comme

$$\begin{cases} \max b^T y, & y \in \mathbb{R}^m, \text{ et } s \in \mathbb{R}^n, \\ \text{avec } A^T y + s = c. \end{cases}$$

- L'introduction de ce vecteur ligne supplémentaire s , contenant les n variables d'écart du dual, permettra de simplifier grandement nos notations, et nous utiliserons donc principalement cette seconde formulation du dual.
- La dualité joue un rôle fondamental dans l'étude et la résolution de (PM). Entre autre, elle fournit des informations supplémentaires très utiles.

Définissons également les domaines admissible \mathcal{F}_D et strictement admissible \mathcal{F}_D^0 pour le problème dual, de façon tout à fait similaire aux domaines \mathcal{F}_P et \mathcal{F}_P^0 pour le primal

$$\begin{aligned} \mathcal{F}_D &= \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c \text{ et } s \geq 0\}, \\ \mathcal{F}_D^0 &= \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c \text{ et } s > 0\}, \end{aligned}$$

\mathcal{F}^0 l'ensemble des solutions primales-duales strictement réalisables :

$$\mathcal{F}^0 = \mathcal{F}_P^0 \times \mathcal{F}_D^0 = \{(x, y, s) : x \in \mathcal{F}_P^0 \text{ et } (y, s) \in \mathcal{F}_D^0\}$$

2.1.4 Propriétés de la dualité

Citons quelques propriétés élémentaires du couple primal-dual

Théorème 31 (Dualité faible) *Si $w = (x, y, s) \in \mathcal{F}$, alors*

$$c^T x \geq b^T y.$$

Preuve. On a

$$c^T x - b^T y = c^T x - (Ax)^T y = x^T c - x^T A^T y = x^T (c - A^T y) \geq 0$$

Ce résultat affirme que toute solution admissible de (PPL) fournit une borne supérieure pour le problème de maximisation (DPL) , et que toute solution du dual fournit une borne inférieure à l'objectif du problème primal

$$\begin{aligned} p^* &= \inf \{c^T x : x \in \mathcal{F}_P\} \\ d^* &= \sup \{b^T y : y \in \mathcal{F}_D\} \end{aligned}$$

■

Corollaire 32 (Dualité forte) *Si $w^* = (x^*, y^*, s^*) \in \mathcal{F}$ et telle que $c^T x^* = b^T y^*$, alors w^* est une solution optimale \Leftrightarrow Si x^* et (y^*, s^*) sont solutions optimales de leurs problèmes respectifs. Cela entraîne que les problèmes (PPL) et (DPL) ont les mêmes valeurs optimales.*

Saut de dualité

La quantité $c^T x - b^T y$, toujours positive, est appelée le saut de dualité

$$S(x, y) = c^T x - b^T y = x^T s$$

1. Cette dernière quantité est positive puisque $x \geq 0$ et $c - A^T y \geq 0$ découle de $A^T y \leq c$.
2. $x^T s$ est une mesure de l'écart entre la fonction objectif primale $c^T x$ et la fonction dual objectif $b^T y$.

Conditions d'optimalité La théorie de Karush-Kuhn-Tucker permet d'écrire des conditions nécessaires d'optimalité pour tout problème d'optimisation contraint possédant une fonction objectif différentiable. De plus, lorsque le problème est convexe ces conditions sont également suffisantes, ce qui est le cas pour l'optimisation linéaire. En considérant le problème (PPL) , l'écriture de ces conditions mène au système (celui-ci peut également aisément être obtenu à l'aide de la théorie des multiplicateurs de Lagrange)

$$x \text{ est optimal pour } (PPL) \Leftrightarrow \exists(z, t) \text{ telque } \begin{cases} A^T z + t = c \\ Ax = b \\ x_i t_i = 0 \forall i \\ x \text{ et } t \geq 0 \end{cases}$$

On remarque que la seconde équation possède exactement la même structure que les contraintes d'égalité du problème dual (DPL). En fait, si nous identifions z avec y et t avec s , nous trouvons

$$x \text{ est optimal pour } (PPL) \Leftrightarrow \exists(y, s) \text{ telque } \begin{cases} Ax = b, \\ A^T y + s = c, \\ x_i s_i = 0 \forall i, \\ x \text{ et } s \geq 0. \end{cases}$$

Finalement, en utilisant les définitions de (PPL) et (DPL) et le fait que lorsqu'un vecteur ligne a et un vecteur colonne b de même longueur sont positifs,

$$a_i b_i = 0 \forall i \in \{1, \dots, n\} \Leftrightarrow \sum_{i=1}^n a_i b_i = 0 \Leftrightarrow a^T b = 0,$$

on obtient

$$x \text{ est optimal pour } (PPL) \Leftrightarrow \exists(y, s) \text{ telque } \begin{cases} x \in \mathcal{F}_P, \\ (y, s) \in \mathcal{F}_D, \\ x^T s = 0. \end{cases}$$

2.1.5 La méthode de Newton

Parmi les méthodes les plus populaires appliquées pour la résolution du système d'équations non linéaire ($K.K.T$), est la méthode de Newton, dans ce qui suit nous décrivons son principe.

Etant donnée une fonction de $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ une fonction non linéaire différentiable telque

$$F(x) = (f_1(x), f_2(x), \dots, f_n(x))^T$$

Cette méthode est une procédure itérative qui a pour objectif de trouver un point $x \in \mathbb{R}^n$ tel que $F(x) = 0$. Pour chaque itéré $x^{(k)}$, la méthode calcule une approximation du premier ordre (linéaire) de F autour de $x^{(k)}$ et définit l'itéré suivant $x^{(k+1)}$ comme le zéro de cette approximation linéaire. Plus formellement, si J est la matrice jacobienne de F , on écrit

$$F(x^{(k)} + \Delta x^{(k)}) \approx F(x^{(k)}) + J(x^{(k)})\Delta x^{(k)},$$

où

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

et le pas de Newton $\Delta x^{(k)}$ est choisi de telle façon que cette approximation linéaire est égal à zéro on pose donc

$$x^{(k+1)} = x^{(k)} + \Delta x^{(k)} \text{ avec } \Delta x^{(k)} = -J(x^{(k)})^{-1} F(x^{(k)}).$$

Le calcul de $\Delta x^{(k)}$ est généralement effectué en pratique via la résolution du système linéaire

$$J(x^{(k)}) \Delta x^{(k)} = -F(x^{(k)})$$

2.2 Les méthodes de barrières (Chemin central)

Les méthodes de points intérieurs ont initialement été proposées par Frisch en 1955. L'idée générale de ces méthodes consiste à approcher d'une manière itérative la solution des problèmes d'optimisation avec contraintes d'inégalité de la forme

$$\left\{ \begin{array}{l} \min_{x \in \mathbb{R}^n} f(x) \\ \text{s.c.} \\ g(x) \geq 0, \end{array} \right. \quad (\mathcal{P}_I)$$

à partir de l'intérieur de l'ensemble réalisable $\mathcal{K} = \{x : g(x) \geq 0\}$. Ces méthodes sont basées sur l'utilisation de fonctions barrière

Définition 33 Une fonction barrière $\Phi(x)$ pour le programme \mathcal{P}_f est une fonction non négative et continue sur le domaine admissible qui tend vers l'infini lorsque la contrainte d'inégalité $g(x)$ tend vers la frontière à partir de l'intérieur du domaine, c'est-à-dire

$$\lim \Phi(x) = \infty, \text{ quand } g(x) \longrightarrow 0^+ \in Fr(\mathcal{K})$$

Les deux exemples classiques de fonctions barrières sont définies pour tout x par

$$\Phi(x) = - \sum_{i=1}^n \log(g_i(x)) \text{ et } \Phi(x) = \sum_{i=1}^n \frac{1}{g_i(x)}.$$

Pour les contraintes $g(x) \leq 0$, on prend généralement

$$\Phi(x) = - \sum_{i=1}^n \log(-g_i(x))$$

Dans ces conditions, on peut remplacer le problème avec contraintes

$$\begin{cases} \min f(x) \\ \text{s. c.} \\ g_i(x) \geq 0 \text{ pour } i = 1, \dots, m, \end{cases}$$

par le problème non contrainte paramétré

$$B(x, \mu) = \min \left(f(x) - \mu \sum_{i=1}^n \log(g_i(x)) \right) \text{ avec } \mu \in \mathbb{R}^+. \quad (\mathcal{P}_{\mu_k})$$

Les premières méthodes de points intérieurs consistaient à résoudre une succession de problèmes barrières (\mathcal{P}_{μ_k}) , tout en faisant tendre le paramètre barrière μ_k vers 0.

Le rôle du terme barrière ajouté consiste à tenir les itérés, générés par une méthode d'optimisation pour problèmes sans contraintes, éloignés de la zone non admissible.

2.2.1 Barrière primal-dual en programmation linéaire

On se place maintenant dans le cadre des problèmes de programmation linéaire.

On considère alors le problème sous forme standard

$$\min_{x \in \mathbb{R}^n} c^T x, \text{ tel que } Ax = b, x \geq 0 \quad (PPL)$$

On utilise une méthode de barrière logarithmique associée à (PPL) : pour tout $\mu > 0$, on introduit la fonction objectif : pour tout $x \in U$,

$$\phi_P(x) = c^T x - \mu \sum_{i=1}^n \log(x_i),$$

où $U = \{x \in \mathbb{R}^n : x > 0\}$ est un ouvert de \mathbb{R}^n , le terme correspondant $\log(x_i)$ dans la fonction $\phi_P(x)$ tendra vers l'infini, étant donné qu'on souhaite minimiser la fonction $\phi_P(x)$, les itérés vont donc rester loin du bord, à l'intérieur de l'ensemble défini par les contraintes.

On considère le problème d'optimisation :

$$\begin{cases} \min (c^T x - \mu \sum_{i=1}^n \log(x_i)) \\ Ax = b, \\ x > 0. \end{cases} \quad (PPL_\mu)$$

et à son dual (DPL_μ) (puisque'il s'agit d'une maximisation, nous devons soustraire le terme barrière)

$$\max_{y \in \mathbb{R}^m} \left(b^T y + \mu \sum_{i=1}^n \log(s_i) \right) \text{ tel que } \begin{cases} A^T y + s = c \\ s > 0 \text{ et } y \in \mathbb{R}^m \end{cases}. \quad (DPL_\mu)$$

La famille de fonctions barrière proposée, par Frisch (1955), associée aux problèmes primal-dual sont :

$$\phi_P(x) = c^T x - \mu \sum_{i=1}^n \log(x_i), \text{ et } \phi_D(x) = -b^T y - \mu \sum_{i=1}^n \log(s_i)$$

Faisons la somme des deux fonctions barrière logarithmique primal-dual

$$\begin{aligned} \phi_{PD}(x) &= c^T x - b^T y - \mu \sum_{i=1}^n \log(x_i) - \mu \sum_{i=1}^n \log(s_i) \\ &= x^T s - \mu \sum_{i=1}^n \log(x_i s_i). \end{aligned}$$

Cette fonction dépend uniquement du produit $x_i s_i$ ajoutons dans les deux termes une constante $n\mu(\log \mu - 1)$, puis on divise sur le nombre positif 2μ , nous écrivons par

la suite

$$\begin{aligned}\Psi_{lb}(x, s, \mu) &= \frac{\phi_{PD}(x) + n\mu(\log \mu - 1)}{2\mu} = \sum_{i=1}^n \frac{1}{2} \left(\frac{x_i s_i}{\mu} - 1 - \log \frac{x_i s_i}{\mu} \right) \\ &= \sum_{i=1}^n \psi_{lb}(v_i) = \frac{v_1^2 - 1}{2} - \log v_1 + \dots + \frac{v_n^2 - 1}{2} - \log v_n\end{aligned}$$

Où

$$v_i = \sqrt{\frac{x_i s_i}{\mu}}, \quad \psi_{lb}(t) = \frac{t^2 - 1}{2} - \log t, \quad t > 0$$

Cette fonction dite fonction log-barrière, cela montre que la fonction logarithmique barrière associée au problème primal-dual est déterminée par une fonction ψ_{lb} d'une variable dite fonction noyau, cette fonction est strictement convexe admet un minimum au point $t = 1$ et $\psi_{lb}(1) = 0$

2.2.2 La fonction de Lagrange associée au problème (PPL_μ)

La fonction de Lagrange associée au problème (PPL_μ) est

$$L_\mu(x, y) = c^T x - y^T (Ax - b) - \mu \sum_{i=1}^n \log(x_i) \quad (4)$$

où y est le multiplicateur de Lagrange associé à la contrainte " $Ax - b$ ".

Condition de Lagrange

On a pour tout $x \in U$ et $y \in \mathbb{R}^m$

$$\begin{cases} \nabla_x L_\mu(x, y) = c - A^T y - \mu X^{-1} e \\ \nabla_y L_\mu(x, y) = b - Ax = 0, \\ x > 0, \end{cases}$$

où pour tout $x \in U$, $X = \text{diag}(x_1, \dots, x_n)$ et $e = (1, \dots, 1)^T \in \mathbb{R}^n$. En écrivant les conditions *K.K.T*, on que x est solution de (PPL_μ) si et seulement si, il existe $y \in \mathbb{R}^m$ tel que

$$\begin{cases} Ax = b \\ \nabla_x L_\mu(x, y) = 0 \text{ c.à.d } A^T y + \mu X^{-1} e = c \text{ et } x > 0 \end{cases}$$

où $X = \text{diag}((x_1, \dots, x_n))$. En posant $s = \mu X^{-1}e$, on peut réécrire les conditions $K.K.T$ précédentes comme :

$$\begin{cases} Ax = b, \\ A^T y + s = c \\ XSe = \mu e \\ x > 0, \text{ et } s > 0 \end{cases} \quad (KKT_\mu)$$

2.2.3 Comparaison des conditions du problème (PPL_μ) et (PPL)

On peut mettre ces conditions en parallèle avec les conditions $K.K.T$ du problème initial (PPL) (cf. Section 2.2) : il existe s, y tels que

$$(PPL) \begin{cases} Ax = b, & (KKT_1) \\ A^T y + s = c, & (KKT_2) \\ XSe = 0, & (KKT_3) \\ x \geq 0, s \geq 0 & (KKT_4) \end{cases} \quad \text{et } (PPL_\mu) \begin{cases} Ax = b, & (KKT_1) \\ A^T y + s = c & (KKT_2) \\ XSe = \mu e & (KKT_3) \\ x > 0, s > 0 & (KKT_4) \end{cases}$$

1. Lorsque $\mu = 0$ on retrouve les conditions d'optimalité primales-duales du programme lineaire originel.
2. On remarque que la condition de complementarité (KKT_3) est la seule condition $K.K.T$ impactée par l'introduction du paramètre μ dans les problèmes de barrières (PPL_μ) .
3. Les deux problèmes (PPL_μ) et (DPL_μ) admettent exactement les mêmes conditions d'optimalité (ce qui signifie qu'ils sont équivalents)
4. L'écriture des conditions $K.K.T$ permet aussi de voir que quand $\mu \rightarrow 0$, une solution de (PPL_μ) converge vers une solution de (PPL) et que les solutions de (PPL_μ) pour $\mu > 0$ se situent bien à l'intérieur de l'ensemble des contraintes contrairement aux solutions de (PPL) qui vérifient la condition de complementarité (KKT_3) pour lesquelles soit le multiplicateur de Lagrange est nul soit la

contrainte est active.

5. L'originalité de l'approche primal-dual est de regarder les conditions d'égalité de $K.K.T$ comme un problème de recherche de zéro d'une fonction. Maintenant on va résoudre le système (KKT_μ) . Une façon d'y arriver est de faire une itération de Newton.

On va étudier dans cette partie la résolution des problèmes barrières (PPL_μ) à μ fixé.

On doit donc résoudre le système (KKT_μ) . Définissons tout d'abord une fonction $F_\mu : \mathbb{R}^{2n+m} \longrightarrow \mathbb{R}^{2n+m}$ par :

$$F_\mu(x, y, s) = \begin{pmatrix} Ax - b, \\ A^T y + s - c, \\ XSe - \mu e \end{pmatrix},$$

où e représente un vecteur dont toutes les composantes sont égales à 1, tandis que X et S sont des matrices carrées

$$\begin{aligned} X &= \text{diag}(x_1, \dots, x_n), \quad S = \text{diag}(s_1, \dots, s_n), \\ X^{-1} &= \text{diag}\left(\frac{1}{x_1}, \dots, \frac{1}{x_n}\right) \text{ et } s = \mu X^{-1}e > 0. \end{aligned}$$

L'idée de base de toute méthode primale-dual est la recherche d'un point réalisable qui résout l'équation non linéaire $F_\mu(x^*, y^*, s^*) = 0$. Pour cela, on utilise l'algorithme

de Newton. On calcule les directions de descente $p = \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix}$ en un point (x, y, s)

solutionnant l'équation

$$J_{F_\mu(x,y,s)} p = -F_\mu(x, y, s)$$

où $J_{F_\mu(x,y,s)}$ est la Jacobienne de F_μ en (x, y, s) :

$$J_{F_\mu(x,y,s)} = \begin{pmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix}$$

Le pas de Newton

Pour résoudre le système des conditions d'optimalité perturbées (KKT_μ), on utilise la méthode de Newton. Le pas de Newton $(\Delta x, \Delta y, \Delta s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ est la solution de

$$\begin{pmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu e - XSe \end{pmatrix}$$

Par un simple calcul, on arrive à résoudre le système linéaire suivant :

$$\begin{cases} \Delta s = -A^T \Delta y, \\ A \Delta x = 0, \\ AS^{-1}X \Delta s = AS^{-1}(\mu e - XSe), \end{cases} \quad (5)$$

qui donne :

$$\begin{aligned} \Delta x &= S^{-1}(\mu e - XSe) - S^{-1}X \Delta s \\ \Delta y &= \left[(AS^{-1}XA^T)^{-1} AS^{-1} \right] (\mu e - XSe) \\ \Delta s &= -A^T \Delta y. \end{aligned}$$

Après la résolution du système (5), on obtient les directions $(\Delta x, \Delta y, \Delta s)$, et le nouvel itéré de Newton est calculé en prenant $\alpha = 1$, mais ce choix du pas de déplacement ne donne pas toujours la stricte faisabilité. Notre but est de préserver la stricte faisabilité et donc on doit trouver un α tel que le nouvel itéré soit strictement réalisable c.à.d

De manière générale, l'itéré suivant (nouvel itéré) $(x^{(k+1)}, y^{(k+1)}, s^{(k+1)})$ est calculé par une formule

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \\ s^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \\ s^{(k)} \end{pmatrix} + \alpha_k \begin{pmatrix} \Delta x^{(k)} \\ \Delta y^{(k)} \\ \Delta s^{(k)} \end{pmatrix},$$

où le scalaire $\alpha_k \in]0, 1]$.

Remarque 34 1. la (k) ième itération $((x^{(k)}, y^{(k)}, s^{(k)}) \in \mathcal{F}^0) \Leftrightarrow \begin{pmatrix} Ax - b = 0, \\ A^T y + s - c = 0 \end{pmatrix}$.

2. Le pas de déplacement α_k est choisi pour rester dans l'ensemble des points intérieurs \mathcal{F}^0

Remarque 35 Quand le paramètre de barrière μ varie, la solution optimale minimisante du (PPL_μ) décrit une courbe à l'intérieur du domaine réalisable appelée chemin central.

2.2.4 Le chemin central primal-dual

Définition 36 Le chemin central (C) est une courbe située dans

$$\mathcal{F}^0 = \mathcal{F}_P^0 \times \mathcal{F}_D^0 = \{z = (x, y, s) : Ax = b, A^T y + s = c, x > 0, s > 0\},$$

paramétrée par un scalaire $\mu > 0$. C'est donc l'image d'une application

$$\mu \in]0, +\infty] \mapsto z(\mu) := (x(\mu), y(\mu), s(\mu)) \in \mathcal{F}^0,$$

autrement dit

$$(\mathcal{C}_\mu) = \{(x(\mu), y(\mu), s(\mu)) : F(x(\mu), y(\mu), s(\mu)) = 0, (x(\mu), s(\mu)) > 0\}.$$

2.2.5 Concept de proximité

Pour que les itérations produites par l'algorithme restent réalisables et proches de la trajectoire centrale (\mathcal{C}_μ) , on est obligé d'introduire une mesure de proximité.

1. La caractéristique des points du chemin central, outre leur admissibilité, étant l'égalité des produits $x_i s_i$ (on a $x_i s_i = \mu$ pour le point repéré par μ).

Définissons la mesure de dualité suivante :

$$\min_{\mu} \|F_{\mu}(x, y, s)\| = \min_{\mu} \|XSe - \mu e\|_2$$

- La projection orthogonale de XSe sur $\{\mu e : \mu \in \mathbb{R}\}$ est donné par $\frac{1}{n} \langle x, s \rangle e$.

Donc nous avons

$$\min_{\mu} \|F_{\mu}(x, y, s)\| = \min_{\mu} \|XSe - \mu e\| = \left\| XSe - \frac{1}{n} \langle x, s \rangle e \right\|.$$

- Sa solution est la moyenne arithmétique des produits $x_i s_i$

2. Pour $z = (x, y, s) \in \mathcal{F}^0$:

$$\begin{aligned} z \in (\mathcal{C}) &\Leftrightarrow \left\| XSe - \frac{1}{n} \langle x, s \rangle e \right\| = 0 \\ &\Leftrightarrow XSe = \frac{1}{n} \langle x, s \rangle e \end{aligned}$$

Nous définissons la mesure de la dualité $\bar{\mu}$ par

$$\bar{\mu} = \frac{1}{n} \sum_{i=1}^n x_i s_i = \frac{x^T s}{n} = \frac{1}{n} \langle x, s \rangle$$

Donc

- Un point $z \in \mathcal{F}^0$ est donc proche du chemin central (\mathcal{C}) nous avons $\mu \approx \bar{\mu}$
 - Le paramètre μ contrôle la distance à l'optimalité.
3. Ceci equivalent, à dire que un point $z = (x, y, s) \in \mathcal{F}^0$ est proche du chemin central si $\left\| \frac{Xs}{\bar{\mu}(z)} - e \right\|$ est petit devant 1 ou encore

$$\|Xs - \bar{\mu}(z) e\| < \bar{\mu}(z).$$

4. Définissons dans l'espace primal-dual, la notion de proximité au chemin central (quantité mesurant la proximité entre un itéré strictement admissible $(x, y, s) \in \mathcal{F}^0$ et le point du chemin central $(x(\mu), y(\mu), s(\mu))$), par la moyenne des écarts

$$(x_i s_i - \bar{\mu})$$

$$\begin{aligned} \delta(x, y, s) &= \frac{1}{\bar{\mu}} \left\| \begin{pmatrix} x_1 s_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n s_n \end{pmatrix} - \begin{pmatrix} \bar{\mu} \\ \cdot \\ \cdot \\ \cdot \\ \bar{\mu} \end{pmatrix} \right\|_2 \\ &= \frac{1}{\bar{\mu}} \|x^T s - \bar{\mu} e\|_2 = \frac{1}{\bar{\mu}} \|X S e - \bar{\mu} e\|_2 \end{aligned}$$

5. Une autre possibilité pour mesurer la proximité consiste à prendre

$$\delta(x, y, s) = \frac{1}{2} \|v - v^{-1}\|, \quad \text{avec } v = \sqrt{\frac{xs}{\bar{\mu}}}, \quad xs = (x_1 s_1, \dots, x_n s_n)$$

– C'est facile de vérifier que

$$\delta(x, y, s) = 0 \Leftrightarrow v = v^{-1} \Leftrightarrow xs = \bar{\mu} e$$

6. Divers voisinages du chemin central (\mathcal{C}) sont associés à ce concept de proximité

► Voisinage restreint (petit voisinage) avec la norme 2 défini par :

$$\begin{aligned} \mathcal{V}_2(\beta) &= \{(x, y, s) \in \mathcal{F}^0 : \|X S e - \bar{\mu} e\|_2 \leq \bar{\mu} \beta\}, \quad \text{avec } 0 < \beta < 1 \\ &= \{(x, y, s) \in \mathcal{F}^0 : \delta(x, y, s) \leq \beta\} \end{aligned}$$

► Le voisinage $\mathcal{V}_2(\beta)$ exprime que la distance Euclidienne entre

$(x_1 s_1, x_2 s_2, \dots, x_n s_n)$ et $(\bar{\mu}, \bar{\mu}, \dots, \bar{\mu})$ ne peut excéder un certain pourcentage de $\bar{\mu}$.

A ce propos, un point est dit voisin de la trajectoire centrale s'il appartient à l'ensemble $\mathcal{V}_2(\beta)$

► $(\mathcal{C}) = \mathcal{V}_2(0) \subset \mathcal{V}_2(\beta_1) \subset \mathcal{V}_2(\beta_2) \subset \mathcal{F}_0$.

► Considérons par exemple pour $n = 2$

$$\frac{1}{2} (w_1 - w_2)^2 = \|w - \bar{\mu}_w e\|^2 \leq (\beta \bar{\mu}_w)^2 = \beta^2 \frac{(w_1 + w_2)^2}{2}$$

Avec

$$w = (w_1, w_2, \dots, w_n)^T = xs = (x_1s_1, x_2s_2, \dots, x_ns_n)^T.$$

- ▶ Évidemment, pour $\beta = 0$ nous avons $x_1s_1 = x_2s_2$
- ▶ Voisinage large avec la norme $\|\cdot\|_{-\infty}$ défini par :

$$\mathcal{V}_{-\infty}(\gamma) = \{(x, y, s) \in \mathcal{F}^0 / x_i s_i \geq \bar{\mu} \gamma\},$$

avec $0 \leq \gamma < 1$ γ largeur du voisinage

- ▶ Les voisinages les plus utilisés dans les méthodes du point intérieur et plus particulièrement dans les méthodes de la trajectoire centrale sont $\mathcal{V}_2(\beta)$, $\mathcal{V}_{-\infty}(\gamma)$

2.3 Différentes méthodes de points intérieures

On distingue quatre catégories : les méthodes de mise à l'échelle affine, les méthodes de réduction de potentiel les, projectives et les méthodes de suivi de chemin.

Notre objectif n'est pas de fournir une liste exhaustive de toutes les méthodes qui ont été proposées à ce jour, mais plutôt de présenter uniquement les méthodes de trajectoire centrale (suivi de chemin) pour un problème d'optimisation linéaire

2.3.1 Méthodes de suivi de chemin primale-duale

Les méthodes de suivi de chemin sont articulées autour du chemin central décrit plus haut et se caractérisent par un choix du paramètre μ différent de zéro. Leur principe revient à définir un certain voisinage autour du chemin central, et à faire évoluer les itérés à l'intérieur de ce voisinage tout en progressant vers la solution.

On peut se demander comment ces algorithmes choisissent les valeurs de μ_k , chaque choix de μ_k revient à choisir une cible appartenant au chemin central.

Le principe des méthodes de suivi de chemin sera donc de calculer la mesure de dualité de l'itéré courant, soit μ_k , et à prendre pour cible un μ_{k+1} inférieur (plus proche de l'optimum) en appliquant un coefficient de proportionnalité $\sigma_k \in [0, 1]$,

selon la formule

$$\mu_{k+1} = \sigma_k \mu_k = \sigma_k \frac{\left(x_i^{(k)}\right)^T s_i^{(k)}}{n}$$

Le choix de σ_k dépend de la méthode considérée.

La direction de Newton centrée peut être maintenant obtenue en résolvant le système suivant :

$$\begin{pmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S^{(k)} & 0 & X^{(k)} \end{pmatrix} \begin{pmatrix} \Delta x^{(k)} \\ \Delta y^{(k)} \\ \Delta s^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^{(k)} S^{(k)} e + \sigma_k \mu_k e \end{pmatrix} \quad (6)$$

Réglage du paramètre de centrage Le nouveau point doit rester dans le voisinage $\mathcal{V}_2(\beta)$

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \\ s^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \\ s^{(k)} \end{pmatrix} + \alpha_k \begin{pmatrix} \Delta x^{(k)} \\ \Delta y^{(k)} \\ \Delta s^{(k)} \end{pmatrix} \in \mathcal{V}_2(\beta),$$

cela implique que ,

$$\|X^{(k+1)} S^{(k+1)} e - \mu_k e\|_2 \leq \mu_k \beta$$

avec

$$\begin{cases} X^{(k+1)} = X^{(k)} + D_x \\ S^{(k+1)} = S^{(k)} + D_s \\ S^{(k)} \Delta x^{(k)} + X^{(k)} \Delta s^{(k)} = -X^{(k)} S^{(k)} e + \sigma \mu_k e \\ D_x = \text{diag}(\Delta x^{(k)}), \Delta x^{(k)} = D_x e \\ D_s = \text{diag}(\Delta s^{(k)}), \Delta s^{(k)} = D_s e \end{cases}$$

Donc

$$\begin{aligned}
\|X^{(k+1)}S^{(k+1)}e - \mu_k e\|_2 &\leq \mu_k \theta \Rightarrow \|(X^{(k)} + D_x)(S^{(k)} + D_s)e - \mu_k e\|_2 \leq \mu_k \beta \\
&\Rightarrow \|(X^{(k)}S^{(k)}e + S^{(k)}\Delta x^{(k)} + X^{(k)}\Delta s^{(k)}) - \mu_k e\|_2 \leq \mu_k \beta \\
&\Rightarrow \|\sigma \mu_k e - \mu_k e\|_2 \leq \mu_k \beta \\
&\Rightarrow |1 - \sigma| \|e\|_2 \leq \beta \text{ avec } \|e\|_2 = \sqrt{n} \\
&\Rightarrow \sigma \geq 1 - \frac{\beta}{\sqrt{n}}.
\end{aligned}$$

Cette valeurs du σ nous garantit que l'itération de Newton reste dans le voisinage $\mathcal{V}_2(\beta)$. On ne peut pas prendre σ trop petit sous peine de sortir de \mathcal{F}^0 . En prenant par exemple $\sigma = 1 - \frac{\beta}{\sqrt{n}}$ comme facteur de reduction de $\bar{\mu}$

Algorithme générique de points intérieur

Algorithme 37

1-Choisir un itéré initial $(x^{(0)}, y^{(0)}, s^{(0)}) \in \mathcal{F}^0$, possédant une mesure de dualité μ_0 et une constante $0 < \sigma < 1$

2-Répéter pour $k = 0, 1, 2, \dots$

3-Choisir $\sigma \in]0, 1[$, soit $\mu_k = \frac{(x^{(k)})^T s^{(k)}}{n}$,

4-Résoudre le système (6) avec $(x^{(k)}, y^{(k)}, s^{(k)})$ et $\mu_{k+1} = \sigma \mu_k$ pour obtenir $(\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)})$

5-Choisir α_k tel que:

$$x^{(k)} + \alpha_k \Delta x^{(k)} > 0 \text{ et } s^{(k)} + \alpha_k \Delta s^{(k)} > 0.$$

6-Poser

$$\begin{bmatrix} x^{(k+1)} \\ y^{(k+1)} \\ s^{(k+1)} \end{bmatrix} = \begin{bmatrix} x^{(k)} \\ y^{(k)} \\ s^{(k)} \end{bmatrix} + \alpha_k \begin{bmatrix} \Delta x^{(k)} \\ \Delta y^{(k)} \\ \Delta s^{(k)} \end{bmatrix} \text{ et } \mu_{k+1} = \sigma \mu_k$$

Il existe de nombreux types de méthodes de suivi de chemin : les méthodes dites à pas court, les méthodes prédiction-correction (à pas alterné) et les méthodes à pas long.

2.3.2 Méthode de suivi de chemin à pas court

Les méthodes de suivi de chemin à pas courts ont comme objectif de maintenir les itérés successifs au voisinage $\mathcal{V}_2(\beta)$ du chemin central. La stratégie adoptée est de maintenir constant pour tout k le paramètre de centrage, $\sigma_k = \sigma$ et la longueur du pas $\alpha_k = 1$. Lorsque le paramètre σ est lié à la taille du voisinage $\mathcal{V}_2(\beta)$, une suite de mesures de dualité décroissante de façon monotone vers zéro. ($\mu_1 > \mu_2 > \dots > 0$) et $\lim_{k \rightarrow 0} \mu_k = 0$.

Les deux caractéristiques principales de la méthode à pas courts sont

1. La mesure de dualité du point visé pour l'itéré suivant est définie par, $\mu_{k+1} = \sigma \mu_k$ où σ est une constante strictement comprise entre 0 et 1.
2. L'itéré suivant sera calculé en appliquant une seule itération de la méthode de Newton aux conditions primales-duales perturbées (KKT_μ) , avec la valeur $\mu = \sigma \mu_k$ qui définit une cible sur le chemin central

$$\begin{cases} Ax = b, \\ A^T y + s = c \\ x_i s_i = \sigma \mu_k, \forall i \\ x > 0, \text{ et } s > 0 \end{cases}$$

Algorithme 38 (Méthode de suivi de chemin à pas court)

Soit un itéré initial $(x^{(0)} ; y^{(0)} ; s^{(0)}) \in \mathcal{V}_2(\theta)$ possédant une mesure de dualité μ_0 et une constante $0 < \sigma < 1$

Répéter pour $k = 0, 1, 2, \dots$

Calculer $\mu_k = \frac{(x^{(k)})^T s^{(k)}}{n}$.

Calculer le pas de Newton $(\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)})$ à l'aide du système d'équations linéaires (6)

Poser

$$\begin{bmatrix} x^{(k+1)} \\ y^{(k+1)} \\ s^{(k+1)} \end{bmatrix} = \begin{bmatrix} x^{(k)} \\ y^{(k)} \\ s^{(k)} \end{bmatrix} + \begin{bmatrix} \Delta x^{(k)} \\ \Delta y^{(k)} \\ \Delta s^{(k)} \end{bmatrix} \quad \text{et } \mu_{k+1} = \sigma \mu_k$$

Fin

Esquissons à présent une preuve de la correction de cet algorithme. Afin que notre stratégie de suivi de chemin fonctionne, nous devons garantir que nos itérés $(x^{(k)}, y^{(k)}, s^{(k)})$ restent suffisamment proches des points $(x^k(\mu), y^k(\mu), s^k(\mu))$ situés sur le chemin central qui nous guide vers une solution optimale. A cet effet, définissons une quantité mesurant la proximité entre un itéré strictement admissible $(x, y, s) \in \mathcal{F}^0$ et le point du chemin central $(x(\mu), x(\mu), x(\mu))$. Puisque la propriété principale de ce point central est $x_i s_i = \bar{\mu} \quad \forall i$, ou de manière équivalente $x^T s = \bar{\mu} e$, la quantité suivante

$$\delta(x, y, s) = \frac{1}{\bar{\mu}} \|x^T s - \bar{\mu} e\|_2,$$

egale à zéro si et seulement si (x, y, s) est égal à $(x(\mu), x(\mu), x(\mu))$ et augmente au fur et à mesure que l'on s'éloigne de ce point central.

A l'aide de cette mesure de proximité l'analyse de l'algorithme repose sur les étapes suivantes

Admissibilité stricte. Prouver que l'admissibilité stricte est préservée par le pas de Newton :

si

$$(x^{(k)}, y^{(k)}, s^{(k)}) \in \mathcal{F}^0 \text{ alors } (x^{(k+1)}, y^{(k+1)}, s^{(k+1)}) \in \mathcal{F}^0$$

Il faudra être particulièrement attentif aux contraintes de positivité, puisqu'elles ne sont a priori pas prises en compte par la méthode de Newton.

Mesure de dualité. Prouver que la mesure de dualité visée est atteinte après le pas de Newton : si $(x^{(k)}, y^{(k)}, s^{(k)})$ possède une mesure de dualité égale à μ_k , l'itéré suivant $(x^{(k+1)}, y^{(k+1)}, s^{(k+1)})$ a une mesure de dualité égale à $\mu_{k+1} = \sigma\mu_k$

Proximité. Prouver que la proximité au chemin central est préservée : il existe une constante $\tau > 0$ telle que si $\delta(x^{(k)}, y^{(k)}, s^{(k)}) < \tau$, on a $\delta(x^{(k+1)}, y^{(k+1)}, s^{(k+1)}) < \tau$ après le pas de Newton

En ajoutant une hypothèse initiale stipulant que $\delta(x^{(0)}, y^{(0)}, s^{(0)}) < \tau$ on peut alors démontrer que la suite des itérés produit par l'algorithme restera confinée dans le voisinage imposé du chemin central et convergera donc (approximativement) vers son point limite, qui est une solution optimale strictement complémentaire.

Critères d'arrêt Pour l'algorithme primal-dual, on utilise un critère d'arrêt classique : le saut de dualité. On sait qu'une solution est optimale pour le primal et le dual en un point d'annulation du saut de dualité. Pour le problème initial (PL) le saut de dualité est

$$s(x, y) = c^T x - b^T y = x^T s$$

Cette quantité est toujours positive pour des points x et s dans leurs ensemble de contraintes et elle est nulle si et seulement si x est solution du primal et s est solution du dual.

Une idée naturelle est alors d'arrêter l'algorithme quand le saut de dualité est proche de zéro. On a $x^T s = n\mu$. Un critère d'arrêt est donc donner par $n\mu$ petit et $x^T s - n\mu$ petit, nous arrêtons l'algorithme lorsque le saut de dualité devient inférieur à ε (paramètre de précision), ce qui se produit dès que $n\mu_k < \varepsilon$. Nous sommes à présent en mesure de formuler l'algorithme sous sa forme finale :

Algorithme 39

Soit un itéré initial $(x^{(0)}, y^{(0)}, s^{(0)}) \in \mathcal{F}^0$ possédant une mesure de dualité μ_0 , la précision requise ε et des constantes appropriées $0 < \sigma < 1$ et τ telles que

1- $\delta(x^{(0)}, y^{(0)}, s^{(0)}) < \tau$.

2- Répéter pour $k = 0, 1, 2, \dots$

3-Calculer le pas de Newton $(\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)})$ à l'aide du système d'équations linéaires (6)

4-Poser $(x^{(k+1)}, y^{(k+1)}, s^{(k+1)}) = (x^{(k)}, y^{(k)}, s^{(k)}) + (\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)})$ et $\mu_{k+1} = \sigma \mu_k$.

Jusqu'à ce que $n\mu_{k+1} < \varepsilon$

En outre, il est possible de prouver qu'une solution de précision ε est atteinte après un nombre d'itérations K vérifiant

$$K = \mathcal{O}\left(\sqrt{n} \log \frac{n\mu_0}{\varepsilon}\right). \tag{7}$$

Cette borne polynomiale sur le nombre d'itérations, qui varie comme la racine carrée de la taille du problème, est à ce jour la meilleure jamais atteinte pour la programmation linéaire.

2.3.3 Méthode de suivi de chemin primale-duale à pas longs.

Les méthodes de suivi de chemin à pas longs vont générer une séquence d'itérés appartenant au voisinage $\mathcal{V}_{-\infty}(\gamma)$. Rappelons que pour de faibles valeurs de γ (par exemple 10^{-3}), le voisinage $\mathcal{V}_{-\infty}(\gamma)$ occupe une grande partie de l'intérieur de l'espace admissible \mathcal{F}^0 comme voisinage du chemin central devant contrôler les itérés. Dans l'algorithme proposé ci-dessous, on prend un pas α le long de la direction de Newton le plus grand possible tout en respectant l'appartenance au voisinage $\mathcal{V}_{-\infty}(\gamma)$

Algorithme 40

Soit un itéré initial $(x^{(0)}, y^{(0)}, s^{(0)}) \in \mathcal{F}^0$, une mesure de dualité initiale μ_0 , la précision requise ε et des constantes appropriées $0 < \sigma < 1$ et τ telles que $\delta(x^{(0)}, y^{(0)}, s^{(0)}) < \tau$

Répéter pour $k = 0, 1, 2, \dots$

Calculer le pas de Newton $(\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)})$ à l'aide du système d'équations linéaires (6)

Poser $(x^{(k+1)}, y^{(k+1)}, s^{(k+1)}) = (x^{(k)}, y^{(k)}, s^{(k)}) + \alpha_k (\Delta x^{(k)}, \Delta y^{(k)}, \Delta s^{(k)})$ et $\mu_{k+1} = \sigma \mu_k$. avec une longueur de pas α_k choisie de façon à ce que $(x^{(k+1)}, y^{(k+1)}, s^{(k+1)}) \in \mathcal{F}^0$

Si $\delta(x^{(k+1)}, y^{(k+1)}, s^{(k+1)}) < \tau$ **Alors** poser $\mu_{k+1} = \sigma \mu_k$

Sinon poser $\mu_{k+1} = \mu_k$ Jusqu'à ce que $n\mu_{k+1} < \varepsilon$

Typiquement, on prend $\gamma = 0.99$, ce qui permet à $\mathcal{V}_{-\infty}(\gamma)$ d'occuper une grande partie de \mathcal{F}^0 .

2.4 Méthode de Trajectoire Centrale basée sur les fonctions noyau pour PL

Dans ce paragraphe, nous considérons le problème d'optimisation linéaire primal donné par :

$$\min \{c^T x : Ax + b = 0, x \geq 0\}, \quad (P)$$

où $A \in \mathbb{R}^{m \times n}$, $\text{rang}(A) = m$, $b \in \mathbb{R}^m$, et $c, x \in \mathbb{R}^n$, et son problème dual associé :

$$\max \{b^T y : A^T y + s = c, y \in \mathbb{R}^m, s \in \mathbb{R}^n, s \geq 0\}. \quad (D)$$

En 1984, Karmarkar [24] a proposé une nouvelle méthode en temps polynomial pour résoudre des programmes linéaires. Cette méthode, et ses variantes, développées par la suite, s'appellent Méthodes de Points Intérieurs, notées **IPMs**. Les lecteurs peuvent se référer à des références de base traitant le sujet, voir Y. Q. Bai et al. [5], et J. Peng et al. [38], C. Roos et al. [42], Y. Ye [59]. Sans perte de généralité, nous supposons que (P) et (D) satisfont la Condition de Points Intérieurs, notée **IPC**, i.e., il existe (x^0, y^0, s^0) tel que

$$Ax^0 = b, x^0 > 0, A^T y^0 + s^0 = c, s^0 > 0.$$

Il est bien connu que la recherche de solutions optimales de (P) et (D) est équivalente à la résolution du système suivant :

$$\begin{cases} Ax = b, & x \geq 0, \\ A^T y + s = c, & s \geq 0 \\ xs = 0. \end{cases} \quad (8)$$

Le système (8) est obtenu par le théorème de la dualité forte appliqué aux problèmes duaux (P) et (D) . L'idée principale d'**IPMs** est de remplacer la troisième équation du système (8), dite condition de complémentarité, par l'équation paramétrée $xs = \mu e$, avec $\mu > 0$. Nous obtenons le système suivant :

$$\begin{cases} Ax = b, & x \geq 0, \\ A^T y + s = c, & s \geq 0 \\ xs = \mu e. \end{cases} \quad (9)$$

Assez surprenant, si la condition **IPC** est satisfaite, alors il existe une solution $(x(\mu); y(\mu); s(\mu))$ du dernier système, pour chaque $\mu > 0$, et cette solution est unique. Nous appelons $x(\mu)$ le μ -centre de (P) et $(y(\mu), s(\mu))$ le μ -centre de (D) . L'ensemble de μ -centres (avec μ prenant plusieurs valeurs décroissantes des nombres réels positifs qui convergent vers 0) forme un trajet, appelé la trajectoire centrale de (P) et (D) .

L'importance de la trajectoire centrale pour l'optimisation linéaire a été reconnue par plusieurs chercheurs, tels que Sonnevend [46] et Megiddo [30]. Si $\mu \rightarrow 0$, la limite de la trajectoire centrale existe, et étant donné que les points limites satisfont la condition de complémentarité, la limite donne des solutions optimales de (P) et (D) .

D'un point de vue théorique, on peut supposer que l'**IPC** est satisfaite. Pour simplifier les contributions théoriques, et toujours sans perte de généralité, nous pouvons supposer que $x^0 = s^0 = e$. Nous supposons que $(x(\mu), y(\mu), s(\mu))$ est connu pour certaines valeurs positives de μ .

Par exemple, en raison de l'hypothèse ci-dessus, nous pouvons supposer que pour $\mu = 1$, $x(1) = s(1) = e$. Nous réduisons μ à $\mu = (1 - \theta)\mu$ pour un nombre θ fixé

$\theta \in]0, 1[$, et on résout le système de Newton suivant :

$$\begin{cases} A\Delta x = 0, \\ A^T \Delta y + \Delta s = 0, \\ s\Delta x + x\Delta s = \mu e - xs. \end{cases} \quad (10)$$

Le système (10) admet une solution unique désignée par $(\Delta x, \Delta y, \Delta s)$. Cette solution est dite la direction de Newton classique pour l'optimisation linéaire. Ensuite, μ est à nouveau réduite par le facteur $(1 - \theta)$, et nous appliquons de nouveau la méthode de Newton ciblant les nouveaux μ -centres, et ainsi de suite. Ce processus est répété jusqu'à ce que devient assez petit (obtention de la condition $n\mu \leq \varepsilon$), alors dans ce cas nous obtenons une ε -solution des problèmes (P) et (D). Le nouvel itéré de la méthode de Newton avec un pas fixé est donné par :

$$x_+ = x + \alpha\Delta x, \quad y_+ = y + \alpha\Delta y, \quad s_+ = s + \alpha\Delta s, \quad \text{où } \alpha \in]0, 1]. \quad (11)$$

Maintenant, nous introduisons le vecteur réduit v et les directions de recherche réduite d_x et d_s comme suit :

$$v = \sqrt{\frac{xs}{\mu}}, \quad d_x = \frac{v\Delta x}{x}, \quad d_s = \frac{v\Delta s}{s}. \quad (12)$$

Le système (10) peut être réécrit comme suit :

$$\begin{cases} A\Delta x = 0, \\ A^T \Delta y + \Delta s = 0, \\ s\Delta x + x\Delta s = -uv(v - v^{-1}). \end{cases} \quad (13)$$

Qui est équivalent à

$$\begin{cases} \bar{A}d_x = 0, \\ \bar{A}^T \Delta y + d_s = 0, \\ d_x + d_s = v - v^{-1}, \end{cases} \quad (14)$$

où $\bar{A} = \frac{1}{\mu}AV^{-1}X$, $V = \text{diag}(v)$, $X = \text{diag}(x)$. A noter que la partie droite de la troisième équation du système (14) n'est autre que l'opposé du gradient de la fonction barrière logarithmique $\Psi_l(v)$, i.e., $d_x + d_s = -\nabla\Psi_l(v)$, alors le système (14) peut être réécrit comme suit :

$$\begin{cases} \bar{A}d_x = 0, \\ \bar{A}^T\Delta y + d_s = 0, \\ d_x + d_s = -\nabla\Psi_l(v), \end{cases} \quad (15)$$

où la fonction barrière $\Psi_l(v) : \mathbb{R}_{++}^n \longrightarrow \mathbb{R}_+$ est définie comme suit :

$$\Psi_l(v) = \Psi_l(x, y, s) = \sum_{i=1}^n \psi_c(v_i) = \sum_{i=1}^n \psi_c\left(\sqrt{\frac{x_i s_i}{\mu}}\right) \quad (16)$$

avec

$$\psi_c(v_i) = \frac{v_i^2 - 1}{2} - \log(v_i). \quad (17)$$

Lemme 41 *Si $(d_x, \Delta y, d_s)$ est une solution du système (15) alors d_x, d_s sont orthogonaux.*

Car :

$$d_x^T d_s = -d_x^T (\bar{A}^T \Delta y) = -(\bar{A} d_x)^T \Delta y = 0$$

Nous utilisons $\Psi(v)$ la fonction de proximité pour mesurer la distance entre l'itéré courant et la trajectoire centrale pour $\mu > 0$ donné. Nous définissons également la mesure de proximité, basée sur la norme, $\delta(v) : \mathbb{R}_{++}^n \longrightarrow \mathbb{R}_+$, comme suit :

$$\delta(v) = \delta(x, y, s) = \frac{1}{2} \|dx + ds\| = \frac{1}{2} \|\nabla\Psi(v)\| \quad (18)$$

Nous appelons $\psi(t)$ la fonction noyau de la fonction barrière logarithmique $\Psi(v)$. A noter que la paire (x, s) coïncide avec le μ -centre $(x(\mu), s(\mu))$ si et seulement si $v = e$. On peut facilement vérifier que la fonction noyau $\psi_c(t)$, définie par (17), est une fonction strictement convexe, qui est définie pour chaque $t \in \mathbb{R}_{++}$ et qui admet un minimum pour $t = 1$, dont la valeur de sa fonction est égale à 0. Il ressort clairement

de la description précédente que la proximité de (x, s) à $(x(\mu), s(\mu))$ est mesurée par la valeur de $\Psi(v)$ et avec $\tau > 0$ en tant que valeur de seuil. Si $\Psi(v) < \tau$, on commence une nouvelle itération externe en effectuant une mise à jour du paramètre μ via

$$\mu^+ = (1 - \theta)\mu, \quad 0 < \theta < 1.$$

Si non ($\Psi(v) > \tau$), nous entrons dans une itération interne par le calcul des directions de recherche à l'itération en cours par rapport à la valeur actuelle de μ et on applique (11) pour obtenir des nouveaux itérés. Si nécessaire, on va répéter la procédure jusqu'à ce qu'on trouve un itéré au voisinage de $(x(\mu), y(\mu), s(\mu))$. Alors μ est encore réduit par un facteur de $(1 - \theta)$ avec $0 < \theta < 1$, et on applique la méthode de Newton pour trouver les nouveaux μ -centres, et ainsi de suite. Ensuite, est à nouveau réduit par le facteur $(1 - \theta)$ avec $0 < \theta < 1$, et nous appliquons la méthode de Newton ciblant les nouveaux μ -centres, et ainsi de suite. Ce processus est répété jusqu'à ce que μ est suffisamment petit, c'est-à-dire jusqu'à ce que $n\mu < \varepsilon$. A ce stade, nous avons trouvé une solution μ -approximative pour un **PL**. Les paramètres, τ , θ et le pas α décrits dans l'algorithme sont choisis tel que le nombre d'itérations produites par l'algorithme sera minimal que possible.

Algorithme 42 *Algorithme Générique Primal-dual IPMs pour l'optimisation linéaire*

Données :

Une fonction de proximité $\Psi(v)$;

Un paramètre de limite $\tau > 0$;

Un paramètre de précision $\varepsilon > 0$;

Un paramètre de mise à jour fixe ; $0 < \theta < 1$;

début:

Initialisation: soit (x^0, y^0, s^0) vérifie la **IPC** :

$$k = 0, \mu^{(0)} = 1, v^0 = \sqrt{\frac{x^0 s^0}{\mu^0}}$$

Tanque : $n\mu^k > \varepsilon$ **faire** :

début (itération externe)

$$\mu^{k+1} = (1 - \theta)\mu^k ;$$

Tant que : $\Psi(v) > \tau$ **faire** :

début (itération interne)

Résoudre le système (15) pour déterminer $(d_x, \Delta y, d_s)$ puis $(\Delta x, \Delta y, \Delta s)$;

Calculer le pas de déplacement α , et poser

$$x = x + \alpha\Delta x;$$

$$y = y + \alpha\Delta y;$$

$$s = s + \alpha\Delta s;$$

$$v = \sqrt{\frac{xs}{\mu}};$$

fin (itération interne)

fin (itération externe)

fin.

2.5 Fonctions noyau et propriétés

Définition 43 On dit que $\psi : \mathbb{R}_{++} \rightarrow \mathbb{R}_+$ est une fonction noyau si ψ est deux fois différentiable et vérifie les conditions suivantes

1. $\psi(1) = \psi'(1) = 0$,
2. $\psi''(t) > 0, \forall t > 0$,
3. $\lim_{t \rightarrow 0^+} \psi(t) = \lim_{t \rightarrow +\infty} \psi(t) = +\infty$.

Les deux premières conditions désignent que $\psi(t)$ est strictement convexe et admet une valeur minimale, si $t = 1$. On définit $\psi(t)$ à partir de sa dérivée seconde comme suit :

$$\psi(t) = \int_1^t \int_1^y \psi''(x) dx dy.$$

De plus, la troisième condition indique la propriété de barrière pour la fonction $\psi(t)$.

Lemme 44 Soit ψ une fonction noyau alors :

1. $t\psi'(t) \geq \psi(t), \forall t \geq 1$
2. Si $\psi(t_1) = \psi(t_2)$, tel que $t_1 < 1 < t_2$ alors :

$$\psi'(t_1) < 0, \psi'(t_2) > 0$$

$$\psi(\beta t_1) < \psi(\beta t_2), \forall \beta > 1.$$

Preuve. 1. Soit la fonction $h(t) = t\psi'(t) - \psi(t)$, pour $t \geq 1$. On a :

$$h(1) = 0,$$

$$h'(t) = t\psi''(t) \geq 0.$$

Donc on a bien que $t\psi'(t) \geq \psi(t), \forall t \geq 1$.

2. Supposons que $\psi(t_1) = \psi(t_2)$ avec $t_1 < 1 < t_2$ alors :

$$t_2\psi'(t_2) > \psi(t_2),$$

et

$$\psi(1) > \psi(t_1) + (1 - t_1)\psi'(t_1),$$

donc

$$\psi'(t_1) < 0, \psi'(t_2) > 0$$

D'autre part, on définit $f(\beta) = \psi(\beta t_1) - \psi(\beta t_2)$, pour $\beta \geq 1$. On a :

$$f(1) = 0.$$

$$f'(\beta) = t_1\psi'(\beta t_1) - t_2\psi'(\beta t_2).$$

Comme $\psi''(t) \geq 0$, et $\beta t_1 < \beta t_2$, pour $\beta > 1$, donc :

$$\psi'(\beta t_1) \leq \psi'(\beta t_2), \text{ pour } \beta > 1.$$

D'autre part, $\psi'(\beta t_2) > 0$, pour $t_2 > 1$, pour $\beta > 1$, donc :

$$f'(\beta) < (t_1 - t_2)\psi'(\beta t_2) < 0, \text{ pour } \beta > 1.$$

alors

$$f(\beta) = \psi(\beta t_1) - \psi(\beta t_2) < 0.$$

Ce qui termine la preuve ■

2.5.1 Qualification d'une fonction noyau

Dans cette partie, on présente les conditions de qualification d'une fonction noyau barrière $\psi(t)$ et quelques propriétés concernant ces conditions. On commence par le lemme technique suivant :

Lemme 45 ([4]) *Soit $\psi(t)$ une fonction noyau barrière donnée. Si elle vérifie les deux propriétés suivantes :*

$$(i) \quad t\psi''(t) - \psi'(t) > 0, \quad t > 1,$$

$$(ii) \quad \psi'''(t) < 0, \quad t > 0.$$

Alors on a :

$$\psi''(t) \psi'(\beta t) - \beta \psi'(t) \psi''(\beta t) > 0, \quad t > 1, \quad \beta > 1.$$

Preuve. Supposons que vérifie (i) et (ii) du lemme. Soit $t > 1$, on considère

$$f(\beta) = \psi''(t) \psi'(\beta t) - \beta \psi'(t) \psi''(\beta t), \quad t > 1, \quad \beta > 1.$$

derivons par rapport β on a

$$f'(\beta) = \psi''(\beta t) [t\psi''(t) - \psi'(t) - \beta t\psi'(t) \psi'''(\beta t)] > 0, \quad \forall \beta > 1.$$

la dernière inégalité est satisfaite du fait que $\psi''(\beta t) > 0$, $t\psi''(t) - \psi'(t) > 0$, d'après (i) et $-\beta t\psi'(t) \psi'''(\beta t) > 0$, car $t > 1$, $\psi'(t) > 0$ et $\psi'''(\beta t) < 0$ d'après (ii). Ce qui implique que $f(\beta) > 0$, pour tout $\beta > 1$: Ce qui termine la preuve. ■

Commençons par une présentation de la définition d'une fonction noyau qualifiée.

Définition 46 Une fonction noyau f est dite qualifiée si $f \in C^3$ et f vérifie les propriétés suivantes

$$t\psi''(t) + \psi'(t) > 0, \quad t < 1, \quad (19)$$

$$t\psi''(t) - \psi'(t) > 0 \quad t > 1, \quad (20)$$

$$\psi'''(t) < 0, \quad t > 0, \quad (21)$$

$$2[\psi''(t)]^2 - \psi'(t) \psi'''(t) > 0 \quad t < 1, \quad (22)$$

$$\psi''(t) \psi'(\beta t) - \beta \psi'(t) \psi''(\beta t) > 0, \quad t > 1, \quad \beta > 1. \quad (23)$$

Remarque 47 D'après le Lemme 45, il est montré que (20) et (21) impliquent (23) Donc, pour montrer la qualification de $\psi(t)$, on vérifie seulement les conditions de (19) à (22).

2.5.2 Propriétés et relation entre les conditions de qualification

Le lemme suivant montre que toute fonction noyau éligible $\psi(t)$ est exponentiellement convexe, dont l'importance de ce lemme est justifié par la suite (voir [4], [38]).

Lemme 48 ([38]) *Soit ψ une fonction deux fois différentiable, alors les propriétés suivantes sont équivalentes*

- (a) $\psi(\sqrt{t_1 t_2}) \leq \frac{\psi(t_1) + \psi(t_2)}{2}$ pour tous $t_1, t_2 > 0$.
- (b) la fonction φ définie par $\varphi(\xi) = \psi(e^\xi)$ est convexe.
- (c) $t\psi''(t) + \psi'(t) \geq 0, t > 0$

Preuve. Signalons que la fonction φ est continue donc pour qu'elle soit convexe, il suffit qu'elle soit mid-convexe. Montrons que (a) \Leftrightarrow (b)

$$\begin{aligned}
 \psi(\sqrt{t_1 t_2}) &\leq \frac{\psi(t_1) + \psi(t_2)}{2} \\
 \Leftrightarrow \varphi\left(\frac{\xi_1 + \xi_2}{2}\right) &= \psi(\sqrt{e^{\xi_1} e^{\xi_2}}) \leq \frac{(\psi(e^{\xi_1}) + \psi(e^{\xi_2}))}{2}, \forall t_1, t_2 > 0 \\
 &= \frac{\varphi(\xi_1) + \varphi(\xi_2)}{2}, \forall \xi_1, \xi_2 \in \mathbb{R}, \text{ telque } t_1 = e^{\xi_1}, t_2 = e^{\xi_2} \\
 \Leftrightarrow \varphi &\text{ est mid-convexe}
 \end{aligned}$$

Concernant l'équivalence (b) \Leftrightarrow (c)

$$\begin{aligned}
 \varphi \text{ convexe} &\Leftrightarrow \varphi''(\xi) = [\psi(e^\xi)]'' = [\psi'(e^\xi) e^\xi]' = [\psi''(e^\xi) e^{2\xi} + e^\xi \psi'(e^\xi)] \\
 &= [\psi''(e^\xi) e^\xi + \psi'(e^\xi)] e^\xi \geq 0, \forall \xi \in \mathbb{R} \\
 \Leftrightarrow t\psi''(t) + \psi'(t) &\geq 0, \forall t > 0, \text{ telque } t = e^\xi
 \end{aligned}$$

■

Lemme 49 *Soit une fonction deux fois différentiable, alors les propriétés suivantes*

sont équivalentes :

$$\left\{ \begin{array}{l} 1) \psi \left(\sqrt{\frac{\xi_1^2 + \xi_2^2}{2}} \right) \leq \frac{\psi(\xi_1) + \psi(\xi_2)}{2}, \xi_1, \xi_2 > 0. \\ 2) \text{ la fonction } \phi \text{ définie par } \phi(t) = \psi(\sqrt{t}) \text{ est convexe, } t > 0. \\ 3) t\psi''(t) - \psi'(t) \geq 0, t > 0. \end{array} \right.$$

Preuve. Signalons que la fonction ϕ est continue donc pour qu'elle soit convexe, il suffit qu'elle soit mid-convexe.

Montrons que 1) \Leftrightarrow 2)

$$\begin{aligned} \psi \left(\sqrt{\frac{\xi_1^2 + \xi_2^2}{2}} \right) &\leq \frac{\psi(\xi_1) + \psi(\xi_2)}{2}, \xi_1, \xi_2 > 0 \\ \phi \left(\frac{t_1 + t_2}{2} \right) &= \psi \left(\sqrt{\frac{t_1 + t_2}{2}} \right) \leq \frac{\psi(\sqrt{t_1}) + \psi(\sqrt{t_2})}{2} \\ &= \frac{\phi(t_1) + \phi(t_2)}{2}, \forall (t_1, t_2) \in (\mathbb{R}_+^*)^2 \text{ telque } t_1 = \xi_1^2, t_2 = \xi_2^2 \\ &\Leftrightarrow \phi \text{ mid-convexe} \end{aligned}$$

Concernant l'équivalence 2) \Leftrightarrow 3)

$$\phi \text{ convexe} \Leftrightarrow \phi''(\xi) = \left[\psi(\sqrt{t}) \right]'' = \frac{1}{4t^{\frac{3}{2}}} (t\psi''(t) - \psi'(t)) \geq 0, t > 0$$

Ce qui termine la preuve. ■

2.5.3 Une borne supérieure de $\Psi(v)$ après chaque itération externe

Notons qu'au début de chaque itération externe du l'Algorithme (42), juste avant la mise à jour du paramètre μ avec le facteur $(1 - \theta)$, on a $\Psi(v) < \tau$. Après la réduction de μ en le multipliant par le facteur $(1 - \theta)$: $\mu_+ = (1 - \theta)\mu$, avec $0 < \theta < 1$, le vecteur v est mis à jour par la relation suivante : $v_+ = \frac{v}{\sqrt{1-\theta}}$. Ceci conduit en général à une augmentation de la valeur de $\Psi(v)$ dans les itérations externes, puis, au cours des itérations internes $\Psi(v)$ diminue jusqu'à ce qu'elle atteigne la première valeur inférieure ou égale à τ . Au cours de l'algorithme, les plus grandes valeurs de $\Psi(v)$ se produisent juste après les mises à jours de μ .

Pour cela, nous avons besoin d'une bonne estimation de la borne supérieure de $\Psi(v)$. En d'autre terme, on trouve une borne supérieure de $\Psi\left(\frac{v}{\sqrt{1-\theta}}\right)$ en fonction de $\Psi(v)$.

Il sera clair que certaines fonctions inverses liées à la fonction noyau et sa première dérivée jouent un rôle important pour analyser l'Algorithme 42. On introduit ces fonctions inverses ici. On note par

$$\varrho : [0, +\infty[\rightarrow [1, +\infty[, \text{ et } \rho : [0, +\infty[\rightarrow]0, 1]$$

les fonctions inverses de ψ pour $t \geq 1$ et $-\frac{1}{2}\psi'$ pour $t \in]0, 1]$, respectivement. Ceci equivalent à

$$\begin{aligned} s &= \psi(t) \Leftrightarrow t = \varrho(s), \text{ pour } t \geq 1 \\ s &= -\frac{\psi'(t)}{2} \Leftrightarrow t = \rho(s), \text{ pour } t \geq 1 \end{aligned}$$

Supposons dans toute la suite que est une fonction noyau qualifiée. On a le théorème suivant qui est représenté en [17] et dont la démonstration dépend de la condition (23).

Théorème 50 *Pour tout vecteur positif v et tout $\beta > 1$, on a :*

$$\Psi(\beta v) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(v)}{n}\right)\right).$$

La démonstration du théorème énoncé à la page 13 ref [4]. En conséquence du Théorème 50, on a si $\Psi(v) \leq \tau$ et $\beta = \frac{1}{\sqrt{1-\theta}}$ alors, après la mise à jour de μ , $\Psi(v)$ est égale à $\Psi\left(\frac{v}{\sqrt{1-\theta}}\right)$ qui est bornée supérieurement par $n\psi\left(\beta\varrho\left(\frac{\Psi(v)}{n}\right)\right)$. Comme ϱ est croissante, alors $\Psi(v) \leq \tau \Leftrightarrow \varrho\left(\frac{\Psi(v)}{n}\right) \leq \varrho\left(\frac{\tau}{n}\right)$.

Et donc pour tout vecteur positif v , si $\Psi(v) \leq \tau$ et $\frac{1}{\sqrt{1-\theta}} > 1$, alors :

$$\Psi_0 \leq L(n, \theta, \tau) = n\psi\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right),$$

où Ψ_0 désigne la valeur de $\Psi(v)$ après la mise à jour de μ .

2.5.4 Analyse de la décroissance de la fonction barrière de proximité Ψ

Dans cette sous-section, on va calculer le pas de déplacement α qui conserve la stricte faisabilité de nouveau itéré et réduire la valeur de proximité durant les itérations internes et on donne les résultats de complexité de l'algorithme.

Pour μ fixé, si on prend α le pas de déplacement, alors le nouveau itéré est donné par :

$$x_+ = x + \alpha\Delta x, \quad y^+ = y + \alpha\Delta y, \quad s^+ = s + \alpha\Delta s.$$

D'après, (12) on peut écrire

$$x_+ = x \left(e + \alpha \frac{\Delta x}{x} \right) = x \left(e + \alpha \frac{d_x}{v} \right) = \frac{x}{v} (v + \alpha d_x)$$

et

$$s_+ = s \left(e + \frac{\alpha}{v} \Delta s \right) = s \left(e + \frac{\alpha}{v} d_s \right) = \frac{s}{v} (v + \alpha d_s)$$

Donc pour μ fixé, on a :

$$v_+ = \sqrt{\frac{x_+ s_+}{\mu}} = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}, \quad (24)$$

Pour tout $\alpha > 0$, on considère la fonction suivante :

$$f(\alpha) = \Psi(v_+) - \Psi(v), \quad \text{pour tout } \alpha \geq 0. \quad (25)$$

Alors $f(\alpha)$ est la différence de la proximité entre le nouvel itéré et l'ancien itéré, pour μ fixé.

En utilisant le Lemme 48 (a) on obtient :

$$\begin{aligned} \Psi(v_+) &= \Psi\left(\sqrt{(v + \alpha d_x)(v + \alpha d_s)}\right) \\ &\leq \frac{1}{2} [\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)] \end{aligned}$$

Et par conséquent $f(\alpha) \leq f_1(\alpha)$, ou

$$f_1(\alpha) = \frac{1}{2} [\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)] - \Psi(v), \quad (26)$$

avec

$$f(0) = f_1(0) = 0,$$

Maintenant, pour estimer la décroissance de la fonction de proximité durant un pas, on utilise les deux dérivées successive de $f_1(\alpha)$ par rapport à α on obtient

$$f_1'(\alpha) = \frac{1}{2} \sum_{i=1}^n [\psi'(v_i + \alpha [d_x]_i) [d_x]_i + \psi'(v_i + \alpha [d_s]_i) [d_s]_i]$$

et

$$f_1''(\alpha) = \frac{1}{2} \sum_{i=1}^n [\psi''(v_i + \alpha [d_x]_i) [d_x]_i^2 + \psi''(v_i + \alpha [d_s]_i) [d_s]_i^2], \quad (27)$$

où les $[d_x]_i$ et $[d_s]_i$ désigne respectivement la $i^{\text{ème}}$ composante de des vecteurs d_x et d_s

D'après la définition de δ , et puisque $d_x + d_s = -\nabla\Psi(v)$, on trouve :

$$f_1'(0) = \frac{1}{2} \nabla\Psi(v)^T (d_x + d_s) = -\frac{1}{2} \nabla\Psi(v)^T \nabla\Psi(v) = -2(\delta(v))^2 \quad (28)$$

d'où $f_1''(\alpha) > 0$ si d_x ou $d_s \neq 0$, alors dans ce cas f_1 est strictement convexe.

Le lemme suivant présente une borne supérieure de $f_1''(\alpha)$ pour **LP**

Lemme 51 ([4]) *Soit $f_1(\alpha)$ défini dans (26) et δ défini dans (18), alors on a :*

$$f_1''(\alpha) \leq 2\delta^2 \psi''(v_{\min} - 2\alpha\delta) \quad (29)$$

Avec v_{\min} désigne le minimum de l'ensemble : $\{v_i \in \mathbb{R}_{++}, 1 \leq i \leq n\}$.

Preuve. On sait que d_x et d_s vérifient (15), par conséquent $d_x \perp d_s$, implique que $d_x + d_s = -\nabla\Psi(v)$, et on a

$$\|d_x + d_s\| = \|(d_x, d_s)\| = 2\delta.$$

Par conséquent, nous avons : $\|d_x\| \leq 2\delta$ et $\|d_s\| \leq 2\delta$. Donc :

$$v_i + \alpha [d_x]_i \geq v_{\min} - 2\alpha\delta, \quad v_i + \alpha [d_s]_i \geq v_{\min} - 2\alpha\delta, \quad 1 \leq i \leq n. \quad (30)$$

D'après (21) (ψ'' strictement décroissante) et on utilisons (27), on obtient :

$$\begin{aligned}
f_1''(\alpha) &\leq \frac{1}{2}\psi''(v_{\min} - 2\alpha\delta) \sum_{i=1}^n ([d_x]_i^2 + [d_s]_i^2) \\
&= \frac{1}{2}\psi''(v_{\min} - 2\alpha\delta) \|(d_x + d_s)\|^2 \\
&= 2\delta^2\psi''(v_{\min} - 2\alpha\delta)
\end{aligned}$$

■

Lemme 52 *Si le pas de déplacement α satisfait*

$$-\psi'(v_{\min} - 2\alpha\delta) + \psi'(v_{\min}) \leq 2\delta. \quad (31)$$

Alors

$$f_1'(\alpha) \leq 0.$$

Preuve. En utilisant (28) et (29) nous écrivons :

$$\begin{aligned}
f_1'(\alpha) &= f_1'(0) + \int_0^\alpha f_1''(\varsigma) d\varsigma \\
&\leq -2\delta^2 + 2\delta^2 \int_0^\alpha \psi''(v_{\min} - 2\varsigma\delta) d\varsigma \\
&= -2\delta^2 - \delta \int_0^\alpha \psi''(v_{\min} - 2\varsigma\delta) d(v_{\min} - 2\varsigma\delta) \\
&= -2\delta^2 - \delta [\psi'(v_{\min} - 2\alpha\delta) - \psi'(v_{\min})] \\
&= -2\delta^2 + 2\delta^2 = 0,
\end{aligned}$$

car on à :

$$-\psi'(v_{\min} - 2\alpha\delta) + \psi'(v_{\min}) \leq 2\delta,$$

alors, $f_1'(\alpha) \leq 0$. D'ou le résultat. ■

Lemme 53 *L'estimation de la plus grande valeur de α vérifiant (31), est donnée par :*

$$\bar{\alpha} = \frac{1}{2\delta} [\rho(\delta) - \rho(2\delta)] \quad (32)$$

Preuve. Nous voulons déterminer la plus grande valeur de α vérifiant (31). On fixe δ , v_{\min} et α . ■

Comme ψ'' est strictement décroissante alors

$$-\psi''(v_{\min} - 2\alpha\delta) + \psi''(v_{\min}) \leq 0.$$

Donc la valeur maximale pour l'expression à gauche dans (31) par rapport à v_{\min} est atteinte si v_{\min} atteint sa valeur minimale, et d'autre part la dérivée de l'expression à gauche dans (31) par rapport à α est

$$2\delta\psi''(v_{\min} - 2\alpha\delta) \geq 0.$$

Donc la valeur maximale pour l'expression à gauche dans (31) par rapport à α est atteinte pour la plus grande valeur de α .

Pour $v_{\min} \in]0, 1[$ alors

$$\delta(v) = \frac{1}{2} \|\nabla\Psi(v)\| \geq \frac{1}{2} |\psi'(v_{\min})| \geq -\frac{1}{2}\psi'(v_{\min}),$$

comme ρ est strictement décroissante alors $\rho(\delta) \leq v_{\min}$. Nous prenons

$$v_{\min} = \rho(\delta). \quad (33)$$

Donc $\psi'(v_{\min}) = -2\delta$, en le remplaçant dans l'expression à gauche dans (31), on obtient

$$-\psi'(v_{\min} - 2\alpha\delta) \leq 4\delta,$$

nous prenons

$$-\frac{1}{2}\psi'(v_{\min} - 2\alpha\delta) = 2\delta,$$

donc

$$v_{\min} - 2\alpha\delta = \rho(2\delta). \quad (34)$$

D'après (33) et (34), on a :

$$\begin{aligned} v_{\min} &= \rho(\delta). \\ 2\alpha\delta &= v_{\min} - \rho(2\delta), \end{aligned}$$

donc

$$\alpha = \frac{1}{2\delta} [\rho(\delta) - \rho(2\delta)].$$

Ce qui termine la preuve

Lemme 54 Soient ρ et $\bar{\alpha}$ définis dans le lemme (53). Alors on a

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}.$$

Par la suite, on note par :

$$\tilde{\alpha} = \frac{1}{\psi''(\rho(2\delta))}. \quad (35)$$

Preuve. Par la définition de ρ on a

$$-\psi'(\rho(\delta)) = 2\delta$$

on dérive par rapport à δ , on trouve

$$-\psi''(\rho(\delta)) \rho'(\delta) = 2$$

alors

$$\rho'(\delta) = -\frac{2}{\psi''(\rho(\delta))} < 0$$

et comme la fonction définie par $\psi''(\rho(s))$ est croissante, donc

$$\begin{aligned}\bar{\alpha} &= \frac{1}{2\delta} \int_{\frac{\delta}{2}}^{\delta} \rho'(s) ds \\ &= \frac{1}{\delta} \int_{\frac{\delta}{2}}^{\delta} \frac{ds}{\psi''(\rho(s))} \\ &\geq \frac{1}{\delta} \frac{1}{\psi''(\rho(2\delta))} \int_{\frac{\delta}{2}}^{\delta} ds,\end{aligned}$$

d'où

$$\psi''(\rho(2\delta)) = \max_{[\frac{\delta}{2}, 2\delta]} \psi''(\rho(s)),$$

ce qui donne

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}.$$

Ce qui termine la preuve ■

Le lemme suivant montre que la proximité Ψ est décroissante. Pour plus de détails voir [39]

Lemme 55 ([39]) *On suppose que $h(t)$ est une fonction convexe et deux fois différentiable, telle que :*

$$h(0) = 0, \quad h'(0) < 0,$$

et h atteint son minimum global à $t^ > 0$; et h'' est croissante pour tout $t \in [0, t^*]$, alors pour tout $t \in [0, t^*]$, on a :*

$$h(t) \leq \frac{th'(0)}{2}.$$

Preuve. On suppose que h'' est croissante pour tout $t \in [0, t^*]$, Posons $\gamma = h'$, alors γ' est croissante, c'est à dire

$$[\gamma'(t_2) - \gamma'(t_1)] [t_2 - t_1] \geq 0, \quad \forall t_1, t_2 \in [0, t^*]$$

L'opérateur γ' est donc monotone. Se rappeler que la condition

$$\langle \nabla f(x_2) - \nabla f(x_1), x_2 - x_1 \rangle \geq 0$$

est une condition nécessaire et suffisante de convexité de la fonction f . ■

Donc $\gamma = h'$ est une fonction convexe. La sécante passant par les points $(0, h'(0))$ et $(t^*, h'(t^*))$ est au dessus de la courbe $(t, h'(t))$ pour tout $t \in [0, t^*]$. L'équation de la sécante passant par les deux points est

$$sec(t) = h'(0) + t \frac{h'(t^*) - h'(0)}{t^*}.$$

Se rappeler que $h'(t^*) = 0$. On a donc

$$h'(t) \leq sec(t) = \left(1 - \frac{t}{t^*}\right) h'(0), \quad \forall t \in [0, t^*].$$

Alors

$$\int_0^t h'(\zeta) d\zeta \leq \int_0^t \left(1 - \frac{\zeta}{t^*}\right) h'(0) d\zeta, \quad \forall t \in [0, t^*],$$

donc

$$[h(t) - h(0)] \leq \frac{th'(0)}{2} + \frac{th'(0)}{2} \left(1 - \frac{t}{t^*}\right), \quad \forall t \in [0, t^*].$$

Se rappeler que $h(0) = 0$. On a donc

$$h(t) - \frac{th'(0)}{2} \leq \frac{th'(0)}{2} \left(1 - \frac{t}{t^*}\right), \quad \forall t \in [0, t^*].$$

Comme $h'(0) < 0$, alors

$$h(t) - \frac{th'(0)}{2} \leq 0, \quad \forall t \in [0, t^*]$$

Ce qui termine la preuve

Et par conséquent, on a le lemme suivant qui présente une borne supérieure pour la valeur de décroissance de la proximité à une itération interne.

Noter aussi que toute borne supérieure pour $f_1(\alpha)$ est aussi une borne supérieure pour $f(\alpha)$

Lemme 56 *Si le pas de déplacement α vérifie $\alpha \leq \bar{\alpha}$, alors :*

$$f(\alpha) \leq -\alpha\delta^2.$$

Preuve. Soit h une fonction définie par

$$h(\alpha) = -2\alpha\delta^2 + \alpha\delta\psi'(v_{\min}) - \frac{1}{2}\psi(v_{\min}) + \frac{1}{2}\psi(v_{\min} - 2\alpha\delta),$$

alors

$$\begin{aligned} h(0) &= f_1(0) = 0, \quad h'(0) = f_1'(0) = -2\delta^2, \\ h''(\alpha) &= 2\delta^2\psi''(v_{\min} - 2\alpha\delta) \end{aligned}$$

d'après le lemme (51) on a

$$f_1''(\alpha) \leq h''(\alpha), \tag{36}$$

et par conséquent on obtient

$$f_1'(\alpha) \leq h'(\alpha) \text{ et } f_1(\alpha) \leq h(\alpha),$$

donc si $\alpha \leq \bar{\alpha}$ on a

$$\begin{aligned} h'(\alpha) &= h'(0) + \int_0^\alpha h''(\xi) d\xi \\ &= -2\delta^2 + \int_0^\alpha 2\delta^2\psi''(v_{\min} - 2\xi\delta) d\xi \\ &= -2\delta^2 + \delta(-\psi'(v_{\min} - 2\alpha\delta) + \psi'(v_{\min})) \leq -2\delta^2 + 2\delta^2 = 0, \end{aligned}$$

avec h'' est croissante, donc d'après le lemme (55) on a

$$f(\alpha) \leq f_1(\alpha) \leq h(\alpha) \leq \alpha \frac{h'(0)}{2} = -\alpha\delta^2.$$

Ce qui termine la preuve. ■

En combinant les résultats des lemmes (54) et (56) on obtient

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}. \quad (37)$$

avec $\tilde{\alpha}$ étant la taille de pas par défaut, comme déjà indiqué en (35). Et on utilise $\tilde{\alpha}$ comme taille de pas par défaut.

Corollaire 57 *Le membre droit de l'expression (37) est strictement décroissante par rapport à δ :*

2.5.5 Borne d'itérations et analyse de la complexité

On doit compter le nombre d'itérations internes nécessaire pour revenir à la situation où $\Psi(v) \leq \tau$. On note que Ψ_0 désigne la valeur de $\Psi(v)$ après la mise à jour de μ , les valeurs suivantes dans la même itération externe sont notées par Ψ_k , $k = 1, 2, \dots, K$, avec K désigne le nombre total d'itérations internes produites durant une itération externe.

Nous avons besoin des lemmes ci-dessous, utilisés par J.Peng et al.[38], pour faciliter au lecteur la compréhension des propos établis le long des différentes preuves.

Lemme 58 ([38]) *Si $\alpha \in [0, 1]$ alors*

$$(1+t)^\alpha \leq 1 + \alpha t, \quad \forall t \geq -1.$$

Preuve. On définit la fonction

$$f(t) = (1+t)^\alpha - \alpha t - 1, \quad \text{pour } t \geq -1$$

Alors

$$f'(t) = \alpha(1+t)^{\alpha-1} - \alpha.$$

$$f''(t) = \alpha(\alpha-1)(1+t)^{\alpha-2}.$$

$$f''(t) \leq 0,$$

donc f est concave et $f'(0) = 0$, alors

$$f(t) \leq f(0) = 0.$$

C'est à-dire que

$$(1 + t)^\alpha \leq \alpha t + 1, \forall t \geq -1$$

Ce qui termine la preuve. ■

Lemme 59 ([38]) *Soit t_0, t_1, \dots, t_K une suite des nombres positifs qui vérifie :*

$$t_{k+1} \leq t_k - \beta t_k^{1-\gamma}, k = 0, 1, \dots, K - 1,$$

tels que $\beta > 0$ et $0 < \gamma < 1$, alors :

$$K \leq \left\lceil \frac{t_0^\gamma}{\beta\gamma} \right\rceil.$$

Preuve. En utilisant le lemme (58), nous pouvons écrire ■

$$\begin{aligned} t_{k+1}^\gamma &\leq (t_k - \beta t_k^{1-\gamma})^\gamma = t_k^\gamma (1 - \beta t_k^{-\gamma})^\gamma \\ &\leq t_k^\gamma (1 - \gamma\beta t_k^{-\gamma}) = t_k^\gamma - \gamma\beta, \end{aligned}$$

donc $t_k^\gamma \leq t_0^\gamma - \gamma\beta k$, en remplaçant k par K , on obtient $t_0^\gamma - \gamma\beta K > 0$, alors

$$K \leq \left\lceil \frac{t_0^\gamma}{\gamma\beta} \right\rceil.$$

Ce qui termine la preuve.

Lemme 60 *Soit K le nombre d'itérations internes, alors on a :*

$$K \leq \frac{(\Psi_0)^\gamma}{\beta\gamma}.$$

Preuve. La définition de K implique que $\Psi_{K-1} > \tau$ et $\Psi_K < \tau$ et (d'après la décroissance de $f(\tilde{\alpha})$ on obtient la valeur de γ et β)

$$\Psi_{k+1} < \Psi_k - \beta (\Psi_k)^{1-\gamma}, \quad k = 0, 1, \dots, K-1,$$

En utilisant le Lemme (59) avec $t_k = \Psi_k$, on obtient :

$$K \leq \frac{(\Psi_0)^\gamma}{\beta\gamma}.$$

Ce qui achève la preuve. ■

La détermination du nombre total d'itérations nécessaires pour trouver une solution optimale primale-duale nécessite le calcul du nombre d'itérations externes.

Le lemme suivant présente le nombre d'itérations externes à partir duquel l'Algorithme 42 converge

Théorème 61 *Soit k le nombre d'itérations externes nécessaires pour trouver une solution optimale primale-duale approchée à une précision $\varepsilon > 0$, alors on a :*

$$k \geq \frac{1}{\theta} \log \frac{n}{\varepsilon} \quad (38)$$

Preuve. D'après l'Algorithme (42), l'obtention d'une solution optimale primale-duale nécessite que

$$n\mu^k \leq \varepsilon$$

Notons que à chaque itération externe k se réduit par un facteur de $(1 - \theta)$ c'est à dire :

Si θ le paramètre de la trajectoire centrale avec $\mu^0 = 1$ et $\mu^k = (1 - \theta)^k \mu^0$, on a :

$$(1 - \theta)^k \mu^0 n \leq \varepsilon \Rightarrow (1 - \theta)^k \leq \frac{\varepsilon}{n},$$

En utilisant la fonction croissante \log on obtient :

$$k \log(1 - \theta) \leq \log \varepsilon - \log n,$$

puisque $\log(1 - \theta) \leq -\theta$ pour $0 < \theta < 1$,

$$\begin{aligned} k\theta &\geq \log\left(\frac{n}{\varepsilon}\right) \\ k &\geq \frac{1}{\theta} \log\left(\frac{n}{\varepsilon}\right) \end{aligned}$$

Ce qui termine la preuve. ■

Le nombre total d'itérations nécessaires pour trouver une solution optimale primale duale approchée avec des précisions $\varepsilon > 0$ et $\tau > 0$ n'est autre que la multiplication de nombre d'itérations externes par le nombre d'itérations internes. On obtient le nombre total d'itérations produites par l'Algorithme (42) qui est donnée par :

$$\frac{[\Psi_0]^\gamma}{\beta\gamma\theta} \log\left(\frac{n}{\varepsilon}\right) \quad (39)$$

Si on prend $\tau = \mathcal{O}(n)$ et $\theta = \Theta(1)$, dans (39), on trouve le nombre d'itérations pour **IPMs** à long-pas.

Par contre, si $\tau = \mathcal{O}(1)$, et $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right) = \frac{1}{2\sqrt{n}}$ dans (39), on trouve le nombre d'itérations pour **IPMs** à court-pas.

Chapitre III

MÉTHODES DE LA TRAJECTOIRE CENTRALE BASÉE SUR UNE NOUVELLE FONCTION NOYAU POUR PL

3.1 *Fonction noyau paramétrique avec double barrière*

3.1.1 Introduction

Plusieurs travaux montrent que l'itération théorique liée court-pas dans les **IPMs** est meilleure que celle liée long-pas. Cependant, il est devenu clair que la théorie de la complexité algorithmique des méthodes à long-pas, qui sont les méthodes les plus efficaces dans la pratique, peut être nettement améliorée en utilisant d'autres fonctions barrière.

La plupart des algorithmes **IPMs** pour l'optimisation linéaire sont basés sur la fonction barrière logarithmique [19, 42]. En 2001, Peng et al. dans [37] ont proposé des nouvelles variantes de **IPMs** basées sur une nouvelle fonction noyau non logarithmique.

Une telle fonction est fortement convexe et coercive, lisse sur son domaine qui n'est autre que l'ortan positif. Ils ont obtenu les résultats de la complexité algorithmique $\mathcal{O}\left(qn^{\frac{q+1}{2q}} \log \frac{n}{\varepsilon}\right)$ est la complexité des méthodes à long-pas et $\mathcal{O}\left(q^2 \sqrt{n} \log \frac{n}{\varepsilon}\right)$ est la complexité des méthodes à court-pas. Pour $q = \frac{\log n}{2}$, on obtient la meilleure borne pour les méthodes à long-pas, soit $\mathcal{O}\left(\sqrt{n} (\log n) \log \frac{n}{\varepsilon}\right)$. Cette borne est la meilleure amélioration

En 2002, Y.Q. Bai et al. dans [5] ont introduit une nouvelle fonction noyau non

logarithmique. Ils ont obtenu une complexité de $\mathcal{O}\left(qn \log \frac{n}{\varepsilon}\right)$ itérations pour les méthodes à long-pas et $\mathcal{O}\left(q^2\sqrt{n} \log \frac{n}{\varepsilon}\right)$ pour les méthodes à court-pas. En 2004, Y.Q. Bai et al. dans [4] ont présenté des nouvelles fonctions noyau non logarithmiques. Ils ont obtenu $\mathcal{O}\left(\sqrt{n}(\log n)^2 \log \frac{n}{\varepsilon}\right)$ itérations comme borne pour les méthodes à long-pas et $\mathcal{O}\left(\sqrt{n} \log \frac{n}{\varepsilon}\right)$ itérations pour les méthodes à court-pas. Dans le même article, Bai et al. ont introduit pour la première fois une nouvelle fonction noyau avec un terme barrière trigonométrique.

En 2012, El Ghami et al. dans [18] ont évalué la fonction noyau avec un terme barrière trigonométrique introduite précédemment par Bai et al. dans [4]. Ils ont obtenu une borne de $\mathcal{O}\left(n^{\frac{3}{4}} \log \frac{n}{\varepsilon}\right)$ itérations pour les méthodes à long-pas. Cette découverte a encouragé les chercheurs à améliorer cette borne pour les fonctions noyaux avec un terme barrière trigonométrique (voir [18, 4, 38, 40, 26, 29]).

En 2014, Peyghami et al. [40, 41] ont proposé deux fonctions noyau différentes avec des termes barrière trigonométriques–logarithmiques et exponentielles logarithmiques pour les *IPMs* dans *LO*. La même année, Cai et al. [12] a introduit une nouvelle fonction noyau paramétrique, combinant la fonction noyau classique avec un terme barrière trigonométrique, et obtenu les même bornes d’itération pour les méthodes à long-pas que [40].

En 2015, Kheirfam et Moslem dans [26] ont proposé une nouvelle fonction barrière trigonométrique, qui présente une complexité de $\mathcal{O}\left(n^{\frac{3}{4}} \log \frac{n}{\varepsilon}\right)$ pour les méthodes à long-pas. Li et Zhang [29], ont présenté un algorithme de points intérieurs de type primal-dual pour résoudre un problème linéaire, basé sur une fonction noyau trigonométrique il ont montré que la complexité correspondant à cet algorithme à grand pas est d’ordre $\mathcal{O}\left(n^{\frac{2}{3}} \log \frac{n}{\varepsilon}\right)$.

Récemment, en 2016, Bouafia et al. [10] ont proposé une fonctions de noyau paramétrées pour les *IPMs*. Ils ont montré que les résultats de la complexité polynomiale obtenu est d’ordre $\mathcal{O}\left(pn^{\frac{p+2}{2(p+1)}} \log \frac{n}{\varepsilon}\right)$ pour les méthodes à long-pas et $\mathcal{O}\left(p^2\sqrt{n} \log \frac{n}{\varepsilon}\right)$

pour les méthodes à petit pas .

Dans ce chapitre, on développe un algorithme d'une méthode de points intérieurs de type primal-dual basé sur une nouvelle fonction noyau pour résoudre le problème **PL**.

3.1.2 Nouvelles directions recherchées

L'idée principale dans **IPMs** est de trouver une nouvelle direction dans la méthode de trajectoire centrale a travers une fonction noyau basée sur le remplacement de la fonction barrière Ψ_l dans par la nouvelle fonction barrière Ψ . Par conséquence, la fonction noyau ψ_c , définie en (17), est remplacée par la nouvelle fonction noyau qualifiée ψ comme suit :

$$\begin{aligned} d_x + d_s &= -\nabla\Psi(v), \text{ où } \Psi(v) = \Psi(x, y, s) \\ &= \sum_{i=1}^n \psi(v_i) = \sum_{i=1}^n \left(v_i^2 - 1 - \log(v_i) + \frac{v_i^{-p} - 1}{p} \right), \end{aligned}$$

et le système de Newton modifié est le suivant

$$\begin{cases} \bar{A}d_x = 0, \\ \bar{A}^T \Delta y + d_s = 0, \\ d_x + d_s = -\nabla\Psi(v). \end{cases} \quad (40)$$

Les nouvelles directions recherchées $(d_x, \Delta y, d_s)$ sont obtenues par la résolution de système (40). Puis, on utilise (12) pour obtenir $(\Delta x, \Delta y, \Delta s)$. Durant la recherche d'une direction, en prenant une taille de pas convenable, on obtient le nouvel itéré donné par :

$$x_+ = x + \alpha\Delta x, \quad y_+ = y + \alpha\Delta y, \quad s_+ = s + \alpha\Delta s, \quad \text{où } \alpha \in]0, 1]. \quad (41)$$

On a d_x et d_s sont orthogonaux car

$$d_x^T d_s = d_x^T (-\bar{A}^T \Delta y) = -d_x^T \bar{A}^T \Delta y = -(\bar{A}d_x)^T \Delta y = 0 \times \Delta y = 0.$$

Pour analyser un algorithme de points intérieurs, on définit la mesure de proximité $\delta : \mathbb{R}_{++} \longrightarrow \mathbb{R}_+$ par :

$$\delta = \delta(v) \longrightarrow \frac{1}{2} \|\nabla \Psi(v)\| = \frac{1}{2} \|d_x + d_s\|$$

Comme d_x et d_s sont orthogonaux, on a

$$\begin{aligned} d_x &= d_s = 0 \\ \Leftrightarrow \nabla \Psi(v) &= 0 \\ \Leftrightarrow v &= e \\ \Leftrightarrow \Psi(v) &= 0 \\ \Leftrightarrow x &= x(\mu), s = s(\mu) \end{aligned}$$

Par conséquent, la valeur de $\Psi(v)$ peut être considérée comme une mesure de la distance entre un itéré donné (x, y, s) et μ -centre $(x(\mu), y(\mu), s(\mu))$.

Pour analyser un algorithme générique basé sur une nouvelle fonction noyau éligible. Nous allons expliciter l'algorithme de point intérieur en suivant l'interprétation de schéma décrit par Bai et al en [4] de la manière suivante :

Etape 1 : Résoudre l'équation $\frac{-1}{2}\psi'(t) = s$ pour obtenir $\rho(s)$, la fonction inverse de $\frac{-1}{2}\psi'(t)$, $t \in]0, 1]$ pour $t \leq 1$. Si l'équation est difficile à résoudre, trouver une borne inférieure pour $\rho(s)$.

Etape 2 : Calculer la décroissance de $\psi(v)$ en terme de $\delta = \delta(v)$ pour la taille de pas par défaut $\tilde{\alpha}$; tel que : $f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}$.

Etape 3 : Résoudre l'équation $\psi(t) = s$ pour obtenir $\varrho(s)$; la fonction inverse de $\psi(t)$ pour $t \geq 1$: Si l'équation est difficile à résoudre, trouver une borne supérieure pour $\varrho(s)$.

Etape 4 : Dédire une borne inférieure pour $\delta(v)$ en terme de $\psi(v)$ de la forme : $\delta(v) \geq \frac{1}{2}\psi'(\varrho(\psi(v)))$.

Etape 5 : Utiliser les résultats des étapes **3** et **4** pour obtenir les constantes positives γ et η , avec $\eta \in (0, 1]$ et tel que $f(\tilde{\alpha}) \leq -\eta\Psi(v)^{1-\gamma}$.

Etape 6 : Calculer une borne supérieure $\Psi_0(v)$ pour $\Psi(v)$ sachant que : $\Psi_0 \leq n\psi\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}\right)$.

Etape 7 : Dédire une borne supérieure pour le nombre d'itérations totales sous la forme $\frac{\Psi_0^\gamma}{\theta^\gamma \eta} \log \frac{n}{\varepsilon}$.

Etape 8 : Si $\tau = \mathcal{O}(n)$ et $\theta = \Theta(1)$ on obtient une borne d'itérations pour les algorithmes à grand pas. Si $\tau = \mathcal{O}(1)$ et $\theta = \Theta(n)$ on obtient une borne d'itérations pour les algorithmes à petit pas.

3.1.3 Propriétés de fonction noyau

Dans cette partie, nous présentons une nouvelle fonction noyau paramétrée et nous donnons ses propriétés qui sont essentielles à notre analyse de la complexité.

Nous rappelons que notre fonction, double barrière est définie par :

$$\psi(t) = t^2 - 1 - \log(t) + \frac{t^{-p} - 1}{p}, \quad p > 0 \quad (42)$$

Les trois premières dérivées de $\psi(t)$ par rapport à t sont :

$$\psi'(t) = 2t - \frac{1}{t} - \frac{1}{t^{p+1}} \quad (43)$$

$$\psi''(t) = 2 + \frac{1}{t^2} + \frac{p+1}{t^{p+2}} \quad (44)$$

$$\psi'''(t) = -\left(\frac{2}{t^3} + \frac{(p+1)(p+2)}{t^{p+3}}\right) \quad (45)$$

D'après (44)

$$\psi''(t) > 1, \quad p > 0, \quad t > 0. \quad (46)$$

En notant également que

$$\begin{cases} \text{i) } \psi(1) = \psi'(1) = 0, \\ \text{ii) } \lim_{t \rightarrow +\infty} \psi(t) = \lim_{t \rightarrow 0^+} \psi(t) = +\infty \end{cases} \quad (47)$$

(46) et (47) montrent que $\psi(t)$ est strictement convexe et minimale en 1. On définit $\psi(t)$ à partir de sa dérivée seconde comme suit :

$$\psi(t) = \int_0^t \int_0^x \psi''(y) dy dx. \quad (48)$$

De plus, les deux conditions **i)** et **ii)** indiquent donc que notre fonction $\psi(t)$ est une fonction barrière noyau

3.1.4 Eligibilité (Qualification) de la nouvelle fonction noyau

Le lemme suivant sert à prouver que la nouvelle fonction noyau définie en (42) est éligible.

Lemme 62 *Soit $\psi(t)$, définie dans (42), et $t > 0$. Alors,*

$$t\psi''(t) + \psi'(t) > 0, \quad t < 1 \quad (49)$$

$$t\psi''(t) - \psi'(t) > 0, \quad t > 1 \quad (50)$$

$$\psi'''(t) < 0, \quad t > 0 \quad (51)$$

$$2[\psi''(t)]^2 - \psi'(t)\psi'''(t) > 0, \quad t < 1 \quad (52)$$

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, \quad t > 1, \quad \beta > 1 \quad (53)$$

Preuve. Pour montrer (49), on utilise (43) et (44) et on obtient :

$$t\psi''(t) + \psi'(t) = 4t + \frac{p}{t^{p+1}} > 0, \text{ pour toute } t > 0 \text{ et } p > 0.$$

donc la condition (49) est satisfaite.

Pour montrer (50), on utilise encore (43) et (44), on a :

$$t\psi''(t) - \psi'(t) = \frac{p+2}{t^{p+1}} + \frac{2}{t} > 0 \text{ pour toute } t > 1,$$

le côté droit de la dernière égalité est positif, ce qui prouve (50). La condition (51) est vérifiée à partir de l'expression du $\psi'''(t)$. Finalement pour montrer, (52) on a

$$\begin{aligned} 2[\psi''(t)]^2 - \psi'(t)\psi'''(t) &= 2 \left[\left(2 + \frac{1}{t^2}\right)^2 + \frac{(p+1)^2}{t^{2p+4}} + 2(p+1)\frac{1}{t^{p+2}} \left(2 + \frac{1}{t^2}\right) \right] \\ &\quad + \left[\frac{2(p+1)(p+2)}{t^{p+2}} + \frac{4}{t^2} - \frac{(p+1)(p+2)}{t^{p+4}} - \frac{2}{t^4} \right] \\ &= 8 + \frac{12}{t^2} + \frac{p(p+1)}{t^{2p+4}} + \frac{2p^2 + 14p + 12}{t^{p+2}} - \frac{p(p-1)}{t^{p+4}} \\ &\quad + \frac{1}{t^{2p+4}} [8t^{2p+4} + 12t^{2p+2} + (2p^2 + 14p + 12)t^{p+2} + pg_p(t)], \end{aligned}$$

où $g_p(t) = p + 1 - (p-1)t^p$. On a deux cas à discuter :

Si $p \in]0, 1]$ la condition (52) est satisfaite.

Si $p > 1$. Alors

$$2[\psi''(t)]^2 - \psi'(t)\psi'''(t) > 0 \Leftrightarrow g_p(t) \geq 0.$$

Cela est certainement vrai si

$$t^p \leq \frac{p+1}{p-1} = 1 + \frac{2}{p-1} \Leftrightarrow t \leq \left(1 + \frac{2}{p-1}\right)^{\frac{1}{p}} \text{ et } t \in]0, 1].$$

Ceci est une inégalité valide, lorsque $0 < t \leq 1$, et par conséquent

$$2[\psi''(t)]^2 - \psi'(t)\psi'''(t) > 0$$

Finalement on obtient la démonstration du (52). Pour (53), d'après le Lemme 2.4 (Bai et al., 2004) il suffit que $\psi(t)$ satisfait (50) et (51). ■

Finalement, $\psi(t)$ est une fonction noyau éligible (qualifiée). Ce qui termine la preuve.

Le lemme suivant montre que toute fonction noyau éligible $\psi(t)$ est exponentielement convexe, dont l'importance de ce lemme est justifié par la suite (voir [4], [38])

Lemme 63 [38 Lemme 2.12] *Les propriétés suivantes sont équivalentes :*

- 1– $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2))$ pour tout $t_1, t_2 > 0$;
- 2– $t\psi''(t) + \psi'(t) > 0$, $t > 0$
- 3– la fonction $t \mapsto \psi(e^t)$ est convexe

Cette propriété est démontré par plusieurs chercheurs voir El Ghami et al. (2012) and Peng et al. (2002a).

Maintenant, on présente les lemmes techniques les plus utilisés dans ce chapitre concernant la nouvelle fonction noyau.

Lemme 64 *Pour $\psi(t)$, nous avons :*

$$\frac{1}{2}(t-1)^2 \leq \psi(t) \leq \frac{1}{2}[\psi'(t)]^2, \quad t > 0 \quad (54)$$

$$\frac{1}{2}\psi'(t) \leq \psi(t) \leq \frac{p+4}{2}(t-1)^2, \quad t \geq 1 \quad (55)$$

$$\|v\| \leq \sqrt{n} + \sqrt{2\Psi(v)}, \quad v \in \mathbb{R}_{++}^n \quad (56)$$

Preuve. Pour montrer (54), en utilisant (46), et (48) on a

$$\psi(t) = \int_0^t \int_0^x \psi''(y) dy dx \geq \int_0^t \int_0^x dy dx = \frac{1}{2}(t-1)^2,$$

ce qui prouve la première inégalité.

Nous allons maintenant passer à la démonstration de la deuxième inégalité :

$$\begin{aligned} \psi(t) &= \int_0^t \int_0^x \psi''(y) dy dx \leq \int_0^t \int_0^x \psi''(x) \psi''(y) dy dx \\ &= \int_0^t \psi''(x) \psi'(x) dx = \int_0^t \psi'(x) d[\psi'(x)] dx = \frac{1}{2}[\psi'(x)]^2. \end{aligned}$$

Pour (55) si $f(x) = 2\psi(t) - (t-1)\psi'(t)$, alors $f'(x) = \psi'(t) - \psi''(t)(t-1)$, $f''(x) = -(t-1)\psi'''(t)$ et $f'(1) = f(1) = 0$. Et comme $\psi'''(t) < 0$ on a $f''(x) \geq 0$ pour $t \geq 1$ on peut résumer ce qui est acquis de la manière suivante $f'(x)$ est une fonction croissante positive il en est de même de $f(x)$. Ce qui prouve la première inégalité

Nous arrivons maintenant à démontrer l'inégalité droite de (55), comme $\psi(1) = \psi'(1) = 0$, et en utilisant le développement de Taylor, on obtient

$$\begin{aligned}\psi(t) &= \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{6}\psi'''(\xi)(\xi-1)^3, \\ &= \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{6}\psi'''(\xi)(\xi-1)^3 \\ &\leq \frac{1}{2}\psi''(1)(t-1)^2 = \frac{p+4}{2}(t-1)^2\end{aligned}$$

pour certains ξ vérifiant $1 \leq \xi \leq t$. Ce qui termine la preuve.

Pour (56) alors en utilisant la première inégalité de (54), et l'inégalité de Cauchy-Schwartz peut obtenir

$$\begin{aligned}\Psi(v) &= \sum_{i=1}^n \psi(v_i) \geq \frac{1}{2} \sum_{i=1}^n (v_i - 1)^2 \\ &= \frac{1}{2} \left[\sum_{i=1}^n v_i^2 - 2 \sum_{i=1}^n v_i + n \right] \\ &= \frac{1}{2} [\|v\|^2 - 2e^T v + \|1\|^2] \\ &\geq \frac{1}{2} [\|v\|^2 - 2\|v\| \|e\| + n] \\ &= \frac{1}{2} [\|v\| - \sqrt{n}]^2.\end{aligned}$$

Cela implique que $\|v\| \leq \sqrt{n} + \sqrt{2\Psi(v)}$, $v \in \mathbb{R}_{++}^n$.

Ce qui achève la démonstration. ■

Remarque 65 Soit $g(t) = -\log(t) + \frac{t^{-p}-1}{p}$, $t > 0$. Alors $\psi(t) = t^2 - 1 + g(t)$, puisque $g'(t) = -\left(\frac{1}{t} + t^{-p-1}\right) < 0$, $g(t)$ est décroissante par rapport à $t \geq 1$

Soit $\varrho : [0, +\infty[\longrightarrow [1, +\infty[$ la fonction inverse de $\psi(t)$ pour $t \geq 1$ et soit $\rho : [0, +\infty[\longrightarrow]0, 1]$ la fonction inverse de $-\frac{1}{2}\psi'(t)$ pour tout $t \in]0, 1]$. Alors nous avons le lemme suivant :

Lemme 66 Pour $\psi(t)$ on a :

$$\sqrt{u+1} \leq \varrho(u) \leq \sqrt{2u+1}, \quad u \geq 0 \quad (57)$$

$$\rho(z) \geq \left(\frac{1}{2z+1} \right)^{\frac{1}{p+1}}. \quad (58)$$

Preuve. Pour montrer (57) soit $u = \psi(t)$ $t \geq 1$. Alors $\varrho(u) = t$, on utilisant (54) on a $u = \psi(t) \geq \frac{1}{2}(t-1)^2$, $t \geq 1$, ce qui implique que

$$t = \varrho(u) \leq \sqrt{2u+1}.$$

Par la définition de $\psi(t)$ on a $u = \psi(t) = g(t) + t^2 - 1$. On utilisant la remarque 65 et $g(1) = 0$, on a $g(t) \in]-\infty, 0]$, puisque $t^2 - 1 = u - g(t) \geq u$. Ce qui implique

$$t = \varrho(u) \geq \sqrt{u+1}.$$

Pour montrer (58), soit

$$z = \frac{-1}{2}\psi'(t), \quad t \in]0, 1] \Leftrightarrow 2z = -\psi'(t)$$

En utilisant la définition de

$$\rho : \rho(z) = t, \quad t \in]0, 1].$$

et en utilisant la définition de $\psi'(t)$, on a

$$-2t + \frac{1}{t} + \frac{1}{t^{p+1}} = 2z \Leftrightarrow \frac{1}{t^{p+1}} = 2t - \frac{1}{t} + 2z$$

puisque la fonction $h : t \longrightarrow 2t - \frac{1}{t}$ monotone strictement croissante $t \in]0, 1]$ et $h(1) = 0$. Donc

$$\frac{1}{t^{p+1}} \leq 2z + 1 \Leftrightarrow t = \rho(z) \geq \left(\frac{1}{2z+1} \right)^{\frac{1}{p+1}}, \quad p > 0 \text{ et } t \in]0, 1].$$

D'ou le résultat. ■

Lemme 67 Soit $\beta \geq 1$. Alors

$$\psi(\beta t) \leq \psi(t) + (\beta^2 - 1)t^2.$$

Preuve. Utilisant la remarque 65 on a $g(\beta t) - g(t) \leq 0$ for $\beta \leq 1$. Puisque

$$\begin{aligned}\psi(\beta t) &= \beta^2 t^2 + g(\beta t) + t^2 - 1 + g(t) - t^2 - g(t) \\ &= (\beta^2 - 1) t^2 + g(\beta t) + \psi(t) - g(t) \\ &= (\beta^2 - 1) t^2 + \psi(t) + (g(\beta t) - g(t)) \leq \psi(t) + (\beta^2 - 1) t^2\end{aligned}$$

Ce qui achève la démonstration. ■

Lemme 68 Soit $0 \leq \theta < 1$, $v_+ = \frac{v}{\sqrt{1-\theta}}$. Si $\Psi(v) \leq \tau$ alors pour $p > 0$ on a

$$\Psi(v_+) \leq \frac{p+4}{2(1-\theta)} \left[\sqrt{2\tau} + \sqrt{n\theta} \right]^2 \quad (59)$$

et

$$\Psi(v_+) \leq \frac{\tau(1+\theta) + n\theta + 2\theta\sqrt{2\tau n}}{1-\theta} \quad (60)$$

Preuve. Montrons (59), Comme $\frac{1}{\sqrt{1-\theta}} \geq 1$ et $\varrho\left(\frac{\Psi(v)}{n}\right) \geq 1$, nous avons $\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}} \geq 1$. Utilisant le théorème 50 avec $\beta = \frac{1}{\sqrt{1-\theta}}$, les inégalités (55), (57), et $\Psi(v) \leq \tau$, on a

$$\begin{aligned}\Psi(v_+) &\leq n\psi\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}}\right) \leq \frac{n(p+4)}{2} \left[\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}} - 1 \right]^2 \\ &= \frac{n}{2} \left(\frac{p+4}{1-\theta} \right) \left[\varrho\left(\frac{\Psi(v)}{n}\right) - \sqrt{1-\theta} \right]^2 \\ &\leq \frac{n}{2} \left(\frac{p+4}{1-\theta} \right) \left[\sqrt{\frac{2\Psi(v)}{n}} + 1 - \sqrt{1-\theta} \right]^2 \\ &= \frac{n}{2} \left(\frac{p+4}{1-\theta} \right) \left[\sqrt{\frac{2\Psi(v)}{n}} + \frac{\theta}{1+\sqrt{1-\theta}} \right]^2 \\ &\leq \frac{p+4}{2(1-\theta)} \left[\sqrt{2\tau} + n\theta \right]^2\end{aligned}$$

où la dernière inégalité découle de $\frac{\theta}{1+\sqrt{1-\theta}} = 1 - \sqrt{1-\theta} \leq \theta$.

Pour montrer (60), utilisant le théorème (50) avec $\beta = \frac{1}{\sqrt{1-\theta}}$, $\Psi(v) \leq \tau$, inégalité (57) et lemme (67) nous obtenon autre borne de $\Psi(v)$ comme suit

$$\begin{aligned} \Psi(v_+) &\leq n\psi\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}}\right) \leq n\left[\psi\left(\varrho\left(\frac{\Psi(v)}{n}\right)\right) + \frac{\theta}{1-\theta}\varrho^2\left(\frac{\Psi(v)}{n}\right)\right] \\ &= \Psi(v) + \frac{n\theta}{1-\theta}\varrho^2\left(\frac{\Psi(v)}{n}\right) \leq \Psi(v) + \frac{n\theta}{1-\theta}\left[\sqrt{\frac{2\Psi(v)}{n}} + 1\right]^2 \\ &\leq \tau + \frac{n\theta}{1-\theta}\left(\sqrt{\frac{2\tau}{n}} + 1\right)^2 = \frac{\tau + \tau\theta + n\theta + 2\theta\sqrt{2\tau n}}{1-\theta} \end{aligned}$$

D'où le résultat. ■

Notation 69 *On note par*

$$\bar{\Psi}_0 = \frac{\tau(1+\theta) + n\theta + 2\theta\sqrt{2\tau n}}{1-\theta} \quad (61)$$

et

$$\tilde{\Psi}_0 = \frac{p+4}{2(1-\theta)}\left[\sqrt{2\tau} + \sqrt{n\theta}\right]^2 \quad (62)$$

On désigne par $\bar{\Psi}_0$ la borne supérieure de $\Psi(v_+)$, pour les méthodes à long-pas (respectivement) $\tilde{\Psi}_0$ dans le cas des méthodes à court-pas, durant la procédure de Newton dans l'algorithme.

Remarque 70 *Pour les méthodes à long-pas avec $\tau = \mathcal{O}(n)$, $\theta = \Theta(1)$, on a $\bar{\Psi}_0 = \mathcal{O}(n)$, et dans le cas des méthodes à court-pas, avec $\tau = \mathcal{O}(1)$, $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$, on a $\tilde{\Psi}_0 = \mathcal{O}(1)$.*

3.1.5 Analyse de la complexité

Lemme 71 *Soit $\delta(v)$ défini dans (18). Alors, on a*

$$\delta(v) \geq \sqrt{\frac{\Psi(v)}{2}}, \quad v \in \mathbb{R}_{++}^n$$

En utilisant la seconde, inégalité (54) du lemme 64 on a

$$\begin{aligned}\Psi(v) &= \sum_{i=1}^n \psi(v_i) \leq \frac{1}{2} \sum_{i=1}^n [\psi'(v_i)]^2 \\ &= \frac{1}{2} \|\nabla \Psi\|^2 = 2\delta^2(v).\end{aligned}$$

Alors on a bien que

$$\delta(v) \geq \sqrt{\frac{\Psi(v)}{2}}, v \in \mathbb{R}_{+++}^n.$$

Ce qui termine la preuve.

► Tout au long de cette section, nous supposons que $\tau \geq 1$. En utilisant le lemme 71 et l'hypothèse que $\Psi(v) \geq \tau$ on a

$$\delta(v) \geq \frac{1}{\sqrt{2}}$$

Lemme 72 Soit $\bar{\alpha}$, la valeur définie dans le lemme 53 vérifiant $\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}$. Alors si $\Psi(t) \geq \tau \geq 1$, on a

$$\bar{\alpha} \geq \frac{1}{2 + (p+2)(4\delta+1)^{\frac{p+2}{p+1}}}.$$

Preuve. En utilisant les lemmes 45, 57 inégalité (58) et la définition de $\psi''(t)$, on a $\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}$ et pour $\rho(2\delta) = t \in (0, 1]$, cela implique que

$$\psi''(\rho(2\delta)) \geq 2 + \frac{p+1}{[\rho(2\delta)]^{p+2}} + \frac{1}{[\rho(2\delta)]^2},$$

on a

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))} \geq \frac{1}{2 + (p+1)(4\delta+1)^{\frac{p+2}{p+1}} + (4\delta+1)^{\frac{2}{p+1}}}.$$

Puisque $(4\delta+1)^{\frac{2}{p+1}} < (4\delta+1)^{\frac{p+2}{p+1}}$ pour $p \in]0, +\infty[$ ce qui implique que

$$\bar{\alpha} \geq \frac{1}{2 + (p+2)(4\delta+1)^{\frac{p+2}{p+1}}}.$$

Ce qui termine la preuve. ■

On pose

$$\tilde{\alpha} = \frac{1}{2 + (p+2)(4\delta+1)^{\frac{p+2}{p+1}}}. \quad (63)$$

$\tilde{\alpha}$ est le pas de déplacement et $\bar{\alpha} \geq \tilde{\alpha}$

En combinant les deux résultats, du lemme (56) et le pas de déplacement, défini dans (63) on obtient le lemme suivant :

Lemme 73 *On considère $\tilde{\alpha}$ le pas de déplacement, défini dans (63), et $\delta \geq 1$. Alors :*

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{2 + (p+2)(4\delta+1)^{\frac{p+2}{p+1}}}$$

Maintenant, nous exprimons la décroissance d'une itération interne en terme $\Psi(v)$ ont utilisant lemme (71)

Lemme 74 *Supposons $\delta \geq 1$ On a*

$$f(\tilde{\alpha}) \leq -\frac{\Psi^{\frac{p}{2(p+1)}}}{128(p+2)}$$

Preuve. Utilisant lemme (71), l'inégalité $\sqrt{2}\delta(v) \geq 1$, et on substitution la valeur de $\tilde{\alpha}$, on obtient

$$\begin{aligned} f(\tilde{\alpha}) &\leq -\frac{\delta^2}{2 + (p+2)(4\delta+1)^{\frac{p+2}{p+1}}} \leq -\frac{\delta^2}{3(p+2)(4\delta+\sqrt{2}\delta)^{\frac{p+2}{p+1}}} \\ &\leq -\frac{\delta^2}{90(p+2)\delta^{\frac{p+2}{p+1}}} \leq -\frac{\delta^{\frac{p}{p+1}}}{90(p+2)} \leq -\frac{\Psi^{\frac{p}{2(p+1)}}}{90\sqrt{2}(p+2)} \\ &\leq -\frac{\Psi^{\frac{p}{2(p+1)}}}{128(p+2)}, \quad 90\sqrt{2} \simeq 128. \end{aligned}$$

Ce qui termine la preuve. ■

Après la mise à jour de μ à $(1-\theta)\mu$, on a

$$\Psi(v_+) \leq \frac{\tau(1+\theta) + n\theta + 2\theta\sqrt{2\tau n}}{1-\theta} = \bar{\Psi}(v_+)$$

D'après la définition de $f(\alpha)$ on a

$$f(\tilde{\alpha}) = \Psi(v_+) - \Psi(v) \leq -\frac{\Psi^{\frac{p}{2(p+1)}}}{128(p+2)} \text{ pour tout } p > 0$$

Donc on comparant avec le lemme (60)

$$\eta = \frac{1}{128(p+2)} \text{ et } \gamma = 1 - \frac{p}{2(p+1)} = \frac{p+2}{2(p+1)}$$

Soient K_1 le nombre total d'itérations internes dans l'itération externe pour les algorithmes à petit pas, et K_2 le nombre total d'itérations internes dans l'itération externe pour les algorithmes à grand pas.

Alors pour $p > 0$, en utilisant le Lemme (75), on a

$$\begin{aligned} K_1 &\leq \frac{(\bar{\Psi}_0)^\gamma}{\eta\gamma} = 156p\bar{\Psi}_0^{\frac{p+2}{2(p+1)}} \\ K_2 &\leq \frac{(\tilde{\Psi}_0)^\gamma}{\eta\gamma} = 156p\tilde{\Psi}_0^{\frac{p+2}{2(p+1)}} \end{aligned}$$

Soit $\bar{\Psi}_0$ définit (61) alors, le nombre total d'itérations pour avoir une solution approchée avec $n\mu < \varepsilon$ est majoré par

$$156p\bar{\Psi}_0^{\frac{p+2}{2(p+1)}} \frac{1}{\theta} \log \frac{n}{\varepsilon} \quad (64)$$

Pour les algorithmes à grand pas, on considère $\tau = \mathcal{O}(n)$, $\theta = \Theta(1)$, par conséquent,

$$\Psi_0 \leq \bar{\Psi}_0 = \mathcal{O}(n)$$

et on a

$$\mathcal{O}\left((p+1)n^{\frac{p+2}{2(p+1)}} \log \frac{n}{\varepsilon}\right) \text{ itérations.}$$

Si on prend $p = \frac{\log n}{2} - 1$, la borne d'itérations devienne

$$\mathcal{O}\left(\sqrt{n} \log(n) \log \frac{n}{\varepsilon}\right)$$

qui est la meilleure complexité pour les algorithmes à grand pas jusqu'à maintenant.

Pour les algorithmes à petit pas, on considère $\tau = \mathcal{O}(1)$, $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$. En utilisant la borne supérieure $\tilde{\Psi}_0$, nous obtenons un majorant de nombre total d'itérations suivant :

$$156p\tilde{\Psi}_0^{\frac{p+2}{2(p+1)}}\frac{1}{\theta}\log\frac{n}{\epsilon}, \quad (65)$$

par conséquent,

$$\Psi_0 \leq \tilde{\Psi}_0 = \mathcal{O}(p),$$

et la complexité algorithmique devienne

$$\mathcal{O}\left(p\sqrt{n}\log\frac{n}{\epsilon}\right) \text{ itérations.}$$

Si on prend $p = 1$, alors la borne d'itérations devienne

$$\mathcal{O}\left(\sqrt{n}\log\frac{n}{\epsilon}\right)$$

qui est la meilleure complexité pour les algorithmes à pas court.

Nous résumons dans le tableau ci-après décrivant quelques résultats de la complexité algorithmique pour l'optimisation linéaire

Fonction noyau ψ_j	Long-pas	Court-pas	Réf
$\frac{t^2 - 1}{2} - \frac{e^{b(1-t)} - 1}{b}, b > 0$	$\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon}), b = \log(n)$	--	[8]
$\frac{t^2 - 1}{2} + \frac{6}{\pi} \tan\left(\pi \frac{1-t}{4t+2}\right)$?	?	[4]
$\frac{t^2 - 1}{2} + \frac{6}{\pi} \tan\left(\pi \frac{1-t}{4t+2}\right)$	$\mathcal{O}\left(n^{\frac{3}{4}} \log \frac{n}{\epsilon}\right)$	$\mathcal{O}(\sqrt{n} \log \frac{n}{\epsilon})$	[18]
$\frac{t^2 - 1}{2} - \log t + \frac{1}{8} \tan^2\left(\frac{1-t}{4t+2}\pi\right)$	$\mathcal{O}\left(n^{\frac{2}{3}} \log \frac{n}{\epsilon}\right)$	--	[41]
$\frac{t^2 - 1}{2} - \int_1^t e^{3\left(\left(\tan \frac{\pi}{2x+2}\right)^{-1}\right)} d\xi$	$\mathcal{O}(\sqrt{n} (\log n)^2 \log \frac{n}{\epsilon})$	--	[40]
$\frac{t^2-1}{2} - \log t + \lambda \tan^2\left(\frac{\pi(1-t)}{3t+2}\right),$ $0 < \lambda \leq \frac{8}{25}$	$\mathcal{O}\left(n^{\frac{2}{3}} \log \frac{n}{\epsilon}\right)$	--	[12]
$t^2 - 2t + \frac{1}{\sin\left(\frac{\pi t}{t+1}\right)}$	$\mathcal{O}\left(n^{\frac{3}{4}} \log \frac{n}{\epsilon}\right)$	--	[26]
$\frac{(t-1)^2}{2} + \frac{(t-1)^2}{2t} + \frac{1}{8} \tan^2\left(\frac{1-t}{4t+2}\pi\right)$	$\mathcal{O}\left(n^{\frac{2}{3}} \log \frac{n}{\epsilon}\right)$	--	[29]
$\frac{t^2-1}{2} + \frac{4}{\pi p} \left[\tan^p\left(\frac{\pi}{2t+2}\right) - 1\right], p \geq 2$	$\mathcal{O}\left(p n^{\frac{p+2}{2(p+1)}} \log \frac{n}{\epsilon}\right)$	$\mathcal{O}(p^2 \sqrt{n} \log \frac{n}{\epsilon})$	[10]
$t^2 - 1 - \log(t) + \frac{t^{-p}-1}{p}, p > 0$	$\mathcal{O}\left(p n^{\frac{p+1}{2p}} \log \frac{n}{\epsilon}\right)$	$\mathcal{O}(p \sqrt{n} \log \frac{n}{\epsilon})$	[11]

Conclusions et perspectives

Dans cette thèse nous nous sommes intéressés à la résolution d'un programme linéaire. Nous avons présenté dans le chapitre 3 une approche générale pour la résolution des problèmes linéaires à l'aide de la méthode des points intérieurs qui utilise une fonction noyau.

Notre objectif dans ce travail est de proposer une nouvelle fonction noyau paramétrée (qualifiée) pour améliorer la complexité algorithmique de l'algorithme de points intérieurs. Sous un choix spécial du paramètre p , tel que p dépend de n , c'est la valeur minimisant la complexité de l'itération. En particulier, en prenant $p = \left(\frac{\log n}{2} - 1\right)$, et on obtient la meilleure complexité connue jusqu'à présent pour ces algorithmes à grand pas qui est de l'ordre $\mathcal{O}\left(\sqrt{n}(\log n) \log \frac{n}{\varepsilon}\right)$. Pour les méthodes à court-pas, nous obtenons la meilleure borne de complexité, à savoir $\mathcal{O}\left(\sqrt{n} \log \frac{n}{\varepsilon}\right)$, en prenant simplement $p = \mathcal{O}(1)$.

Notre prochain travail est consacré au développement d'un algorithme primal-dual de points intérieurs basé sur la même fonction et une nouvelle fonction noyau avec d'autres techniques dont le but est d'améliorer la complexité et de généraliser nos travaux à plusieurs problèmes d'optimisation tels que le problème semi-défini (*SDP*), les problèmes de complémentarité linéaire.

De plus l'étude numérique concernant l'algorithme proposé de la résolution d'un problème (*PL*) par la méthode de trajectoire centrale basée sur les fonctions noyaux serait sans doute un objet de recherche.

L'ensemble de ces résultats sont publiés dans "Int. J. Mathematical Modelling and Numerical Optimisation" [11]

RÉFÉRENCES

- [1] F. Alizadeh, Combinatorial Optimization with Interior-Point Methods and Semidefinite Matrices, Ph.D. Thesis, Computer Science Department, University of Minnesota, Minneapolis, MN, 1991.
- [2] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, *SIAM J. Optim.* 5 (1995) 13-51.
- [3] K.M. Anstreicher, J.P. Vial, On the convergence of an infeasible primal-dual interior point method for convex programming. *Optimization methods and software*, 3 : 273 – 283, (1994)
- [4] Y. Q. Bai, M. El Ghami, C. Roos, A comparative study of kernel functions for primaldual interior point algorithms in linear optimization, *SIAM Journal on Optimization*, 15, 101 – 128, (2004).
- [5] Y. Q. Bai, C. Roos, A primal-dual interior point method based on a new kernel function with linear growth rate, in : *Proceedings of the 9th Australian Optimization Day, Perth, Australia*, (2002).
- [6] Y.Q. Bai, J.L. Guo, C. Roos. A new kernel function yielding the best known iteration bounds for primal-dual interior-point algorithms. *Acta Mathematica Sinica. English Series*, 25 (12) : 2169-2178, (2009).
- [7] Y.Q. Bai, G. Lesaja, C. Roos, G.Q. Wang, and M. El Ghami. A class of large- and small-update primal-dual interior-point algorithm for linear optimization. *Journal of optimization theory and application*, 138 (3) : 341-359, (2008).
- [8] Y. Q. Bai, M. El Ghami and C. Roos, A new efficient large-update primal-dual interior-point method based on a finite barrier, *SIAM J. Optim.* 13 (2003) 766-782 (electronic).

- [9] R. Bellman and K. Fan, On systems of linear inequalities in hermitian matrix variables, In V.L. Klee, editor, Convexity, Vol. 7 Proc. Symposia in Pure Mathematics, pages 1–11. Amer. Math. Soc., Providence, RI, 1963
- [10] M. Bouafia, D. Benterki. A ,Yassine : An efficient primal-dual interior point method for linear programming problems based on a new kernel function with a trigonometric barrier term. *J. Optim. Theory Appl.* **170**, 528–545 (2016)
- [11] B. Bounibane, E.a. Djefal. (2019). ‘Kernel function-based interior-point algorithms for linear optimisation’, *Int. J. Mathematical Modelling and Numerical Optimisation*, Vol. **9**, No. 2, pp 158-177
- [12] X. Z. Cai, G. Q. Wang, M. El Ghami, and Y. J. Yue, Complexity analysis of primaldual interior-point methods for linear optimization based on a new parametric kernel function with a trigonometric barrier term, *Abstract and Applied Analysis*, 11 pages, Art. ID 710158, (2014)
- [13] G. B. DANTZIG, *Linear Programming and Extensions*, Princeton University Press, Princeton, New Jersey, 1963.
- [14] A.S. El Bakry, R.A. Tapia, T. Tsuchiya, Y. Zang, On the formulation and theory of the newton interior point method for nonlinear programming. *JOTA*, 89(3) :507-541, 1996.
- [15] A. V. FIACCO ET G. P. MCCORMICK, *Nonlinear programming : Sequential unconstrained minimization techniques*, John Wiley and Sons, Inc., New York London-Sydney, 1968.
- [16] K. R. FRISCH, The logarithmic potential method of convex programming, Memorandum of May 13 (1955), University Institute of Economics, Oslo, Norway.
- [17] M. El Ghami. New primal-dual interior-point methods based on kernel functions. Phd thesis. Delft University, Netherland, (2005)

- [18] M. El Ghami, Z.A. Guennoun, S. Bouali, T. Steihaug, , Interior point methods for linear optimization based on a kernel function with a trigonometric barrier term. *J. Comput. Appl. Math.* 236, 3613–3623 (2012)
- [19] M. El Ghami, I. D. Ivanov, C. Roos, T. Steihaug, A polynomial-time algorithm for LO based on generalized logarithmic barrier functions, *Int. J. Appl. Math.* 21 99 –115 (2008).
- [20] C. C. Gonzaga, Path following methods for linear programming, *SIAM Review.* 34 (1992) 167–227
- [21] M. Halicka, E. De Klerk and C. Roos, On the convergence of the central path in semidefinite optimization, *SIAM J. Optim.* 12 (2002) 1090–1099.
- [22] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz, An interior-point method for semidefinite programming, *SIAM J. Optim.* 6 (1996) 342–361
- [23] J. Ji, F. A. Potra and R. Sheng, On the local convergence of a predictor-corrector method for semide. . . nite programming, *SIAM J. Optim.* 10 (1999) 195–210.
- [24] N. K. Karmarkar, A new polynomial-time algorithm for linear programming, in : *Proceedings of the 16th Annual ACM Symposium on Theory of Computing*, 4, 373–395, (1984).
- [25] L. G. Khachiyan, A polynomial algorithm in linear programming, *Soviet Mathematics Doklady*, 20 191–194, (1979).
- [26] B. Kheirfam, M. Moslem. :Apolynomial-time algorithm for linear optimization based on a new kernel function with trigonometric barrier term. *Yujor* 25 (2), 233-250 (2015)
- [27] V. KLEE ET G. J. MINTY, How good is the simplex algorithm ?, *Inequalities*, O. Shisha, ed, Academic Press, New York (1972), pp. 159-175.

- [28] M. Kojima, S. Shindoh, and S. Hara, Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices, *SIAM J. Optim.* 7 (1997) 86–125.
- [29] Xin, Li , Mingwang Zhang, Interior-point algorithm for linear optimization based on a new trigonometric kernel function, *Operations Research Letters*, 43(5), 471–475, (2015).
- [30] N. Megiddo, Pathways to the optimal set in linear programming, in : N. Megiddo (Ed.), *Progress in Mathematical Programming : Interior Point and Related Methods*, Springer-Verlag, New York, 131–158, (1989).
- [31] G.P. McCormick, J.E. Falck, The superlinear convergence of a nonlinear primal-dual algorithm. Technical Report GWU/OR/Serial T-550/91, Departement of Operation Research, The George Washinton University, Washington DC, 1991.
- [32] R. D. C. Monteiro and T. Tsuchiya, Polynomial convergence of a new family of primal-dual algorithms for semidefnite programming, *SIAM J. Optim.* 9 (1999) 551–577.
- [33] R. Monteiro, I. Adler, Interior path following algorithm primal-dual algorithms-part 2 : Convex quadratic programming.. *Mathematical Programmming*, 44 :43-66, 1989.
- [34] R. D. C. Monteiro, Primal-dual path-following algorithms for semide-finite programming, *SIAM J. Optim.* 7 (1997) 663–678.
- [35] R. D. C. Monteiro, Polynomial convergence of primal-dual algorithms for semide-finite programming based on Monteiro and Zhang family of directions, *SIAM J. Optim.* 8 (1998) 797–812.
- [36] Y. E. Nesterov, A. S. Nemirovski, Interior-point polynomial methods in convex programming, *SIAM Studies in Applied Mathematics*. SIAM Publications, Philadelphia, (1994).

- [37] J. Peng, C. Roos, and T. Terlaky, A new and efficient large-update interior-point method for linear optimization, *Journal of Computational Technologies*, 6, 61–80, (2001)
- [38] J. Peng, C. Roos, T. Terlaky. *Self-regularity. A new paradigm for Primal-Dual Interior Point Algorithm*. Princeton University Press. Princeton, (2002).
- [39] J. Peng, C. Roos, T. Terlaky. Self-regular functions and new search directions for linear and semide-finite optimization. *Mathematical Programming*, 93 : 129-171, (2002).
- [40] M. R. Peyghami, S. Fathi Hafshejani, Complexity analysis of an interior-point algorithm for linear optimization based on a new proximity function, *Numerical Algorithms*, 67, 33–48, (2014).
- [41] M. R. Peyghami, S. Fathi Hafshejani, L. Shirvani, Complexity of interior-point methods for linear optimization based on a new trigonometric kernel function, *Journal of Computational and Applied Mathematics*, 255, 74–85, (2014).
- [42] C. Roos, T. Terlaky, J. Ph. Vial, *Theory and algorithms for linear optimization, An Interior-Point Approach*, John Wiley & Sons, Chichester, UK, (1997).
- [43] J. Renegar, A polynomial-time algorithm, based on Newton’s method, for linear programming, *Mathematical Programming*, 40 59–93, (1988).
- [44] R. T. Rockafellar, *Convex analysis*, published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540, (1988).
- [45] N. Z. Shor, Utilization of the operation of space dilatation in the minimization of convex functions, *Kibernetika*, 1, 6–12, (1970).
- [46] G. Sonnevend, An analytic center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, in : A. Prkopa, J. Szelezsn, B. Strazicky (Eds.), *Lecture Notes in Control and Information Sciences*, vol. 84, Springer Verlag, Berlin, 1986, 866-876.

- [47] A.L. Tits, A. Wachter, S. Bakhtiari, T.J. Urban, and C.T. Lawrence A PrimalDual Interior-Point Method for Nonlinear Programming with Strong Global and Local Convergence Properties. *SIAM Journal on Optimization* 14(1), pp. 173-199, 2003
- [48] R.J. Vanderbei, An interior-point algorithm for quadratic programming *Statistics and Operations Research*. Princeton SOR-94-15.
- [49] J.P. Vial, Computational experience with a primal-dual interior point method for smooth convex programming. *Optimization methods and software*, 3 :285-310, 1994.
- [50] A. Wachter , An Interior Point Algorithm for Large-Scale Nonlinear Optimization with Applications in Process Engineering. Phd Thesis, Carnegie Mellon University, 2002.
- [51] G.Q. Wang, Y.Q. Bai. A new primal-dual path-following interior-point algorithm for semidefinite optimization. *Journal of Mathematical Analysis and Applications*, 353 (1) : 339-349, (2009)
- [52] G.Q. Wang, D.T. Zhu. A unified kernel function approach to primal-dual interiorpoint algorithms for convex quadratic SDO. *Numerical Algorithms*, 57 (4) : 537-558,(2011).
- [53] G.Q. Wang, Y.Q. Bai. Primal-dual interior-point algorithms for convex quadratic semidefinite optimization. *Nonlinear analysis : theory, methods & applications*, 71 (7-8) : 3389-3402, (2009).
- [54] G.Q. Wang, Y.Q. Bai, C. Roos. Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function. *Journal of mathematical modelling and algorithms*, 4 (4) : 409-433, (2005).
- [55] G.Q. Wang, Z.H. Zhang, D.T. Zhu. On extending primal-dual interior-point method for linear optimization to convex quadratic cone optimization. *Numerical*

- functional analysis and optimization, 34 (5) : 576-603, (2013).
- [56] S. J. Wright, primal-dual interior point Methods, SIAM, Philadelphia, USA, (1997)
- [57] S.J. Wright, An interior point algorithm for linearly constrained optimization. SIAM J. Optimization, 2 :450-473, 1992.
- [58] H. Yamashita, A globally convergent primal-dual interior point method for constrained optimization. Tech. Rep. Mathematical Systems Institute Inc, 2-5-3 Shinjuku, Shinjukuku, Tokyo, Japan, 1995.
- [59] Y. Ye, Interior point algorithms. Theory and Analysis, John-Wiley. Sons, Chichester, UK, (1997).
- [60] M.W. Zhang. A large-update interior-point algorithm for convex quadratic semi-definite optimization based on a new kernel function. Acta Mathematica Sinica. English Series, 28 (11) : 2313-2328, (2012)
- [61] Y. Zhang, On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming, SIAM J. Optim. 8 (1998) 365–386.
- [62] D. Zhao, M.W. Zhang. A primal-dual large-update interior point algorithm for semidefinite optimization based on a new parametric kernel function. Statistics, optimization and information computing, (1) : 41-61, (2013).