

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Résolution d'un problème d'admissibilité convexe par une méthode de réduction de potentiel utilisant le centre analytique et des coupes multiples.

Hespel, Anne

Award date:
1998

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Facultés Universitaires Notre-Dame de la Paix
Namur
Facultés des Sciences - Département de Mathématiques

RÉSOLUTION D'UN PROBLÈME D'ADMISSIBILITÉ CONVEXE
PAR UNE MÉTHODE DE RÉDUCTION DE POTENTIEL
UTILISANT LE CENTRE ANALYTIQUE
ET DES COUPES MULTIPLES.

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

Promoteur : J.J. Strodiot

Anne HESPEL

Année académique 1997-1998

Je tiens tout particulièrement à remercier
le Professeur J.-J. Strodiot, promoteur
de ce mémoire, pour ses conseils précieux
et sa grande disponibilité tout au long
de la rédaction.

Je ne voudrais pas non plus oublier ma famille,
qui a su si bien me soutenir et m'encourager
durant ces quatre années d'études.

Résumé

Ce mémoire étudie une méthode de réduction de potentiel destinée à résoudre le problème d'admissibilité convexe suivant : étant donnée une partie convexe \mathcal{C} de \mathcal{R}^n , trouver un point $y \in \mathcal{C}$, ou conclure que \mathcal{C} est vide.

Partant d'un polytope contenant \mathcal{C} , nous effectuons des coupes de façon à construire des approximations polyédriques de \mathcal{C} de plus en plus affinées. Pour chacune de celles-ci, nous calculons le centre analytique, et cela jusqu'à ce que l'un d'entre eux appartienne à \mathcal{C} .

Nous appliquons d'abord cette méthode au cas particulier du problème d'admissibilité linéaire, pour la généraliser ensuite au cas convexe. Enfin, nous proposons une stratégie permettant d'accélérer la convergence de cette méthode. Elle consiste à passer d'une approximation polyédrique à une autre en effectuant plusieurs coupes au lieu d'une seule comme précédemment.

Abstract

This thesis considers a potential reduction method intended to resolve the next convex admissibility problem : given a convex set \mathcal{C} of \mathcal{R}^n , find a point y in \mathcal{C} , or conclude that \mathcal{C} is empty.

Starting with a convex polytope containing \mathcal{C} , we perform cuts in order to build increasingly refined polyhedral approximations of \mathcal{C} . For each of them we compute the analytic center and this until one of them is included in \mathcal{C} .

First we apply this method in the particular case of linear admissibility problem. Then we generalize it to the convex case. Last we propose a strategy allowing to accelerate the convergence of the method in the convex case. It consists of reducing the polyhedral approximation of \mathcal{C} by performing several cuts instead of only one as previously.

Table des matières

Introduction	3
1 Formalisation de la méthode ACCPM.	7
1.1 Plans coupants	7
1.1.1 Position du problème	7
1.1.2 Principe de l'algorithme ACCPM.	7
1.2 Centre analytique	10
1.2.1 Définition	10
2 Calcul du centre analytique par une méthode de réduction de potentiel.	15
2.1 Introduction	15
2.2 Proximité du centre analytique	16
2.3 Algorithmes de calcul du centre analytique	30
2.3.1 Algorithme dual	30
2.3.2 Algorithme primal	40
2.3.3 Algorithme primal-dual	49
3 Algorithme ACCPM basé sur le principe de réduction de potentiel pour un problème d'admissibilité linéaire.	56
3.1 Introduction	56
3.2 Principe de l'algorithme ACCPM avec réduction de potentiel	56
3.3 Propriété combinatoire du centre analytique	57
3.4 Algorithme ACCPM pour la résolution du problème d'admissibilité linéaire	64
3.4.1 Algorithme ACCPM utilisant des centres analytiques exacts	66
3.4.2 Algorithme ACCPM utilisant des centres analytiques approximatifs	75
4 Algorithme ACCPM basé sur le principe de réduction de potentiel pour un problème d'admissibilité convexe.	87
4.1 Introduction	87
4.2 Préliminaires	88
4.3 Algorithme ACCPM pour la résolution du problème d'admissibilité convexe	89
4.4 Réduction de potentiel	89
4.5 Mise à jour du centre approximatif	91
4.6 Complexité algorithmique	93

5	Algorithme ACCPM générant plusieurs coupes pour le problème d'admissibilité convexe.	101
5.1	Introduction	101
5.2	Algorithme ACCPM générant plusieurs coupes	101
5.3	Réduction de potentiel	103
5.4	Complexité algorithmique	107
5.4.1	Borne sur le nombre total de coupes	107
5.4.2	Borne sur le nombre total de pas de Newton	116
	Conclusion	125
	Bibliographie	126

Introduction

Le but de ce mémoire est d'étudier des méthodes permettant de résoudre le problème d'admissibilité convexe suivant :

Etant donné un ensemble convexe \mathcal{C} de \mathcal{R}^n , trouver un point dans \mathcal{C} ou conclure que \mathcal{C} est vide.

Ce problème recouvre notamment celui de la recherche de la solution optimale d'un problème de programmation convexe lorsque la valeur optimale z^* de celui-ci est connue. En effet, si z^* désigne la valeur optimale du problème

$$(P) \quad \begin{cases} \min f(y) \\ \text{s.c. } y \in Y \\ \quad \quad l \leq y \leq h \end{cases}$$

où $f : \mathcal{R}^n \rightarrow \mathcal{R}$ est convexe, $Y \subseteq \mathcal{R}^n$ est convexe, et où l et h sont des bornes sur y , alors résoudre ce problème revient à trouver un point y_0 appartenant à l'ensemble convexe

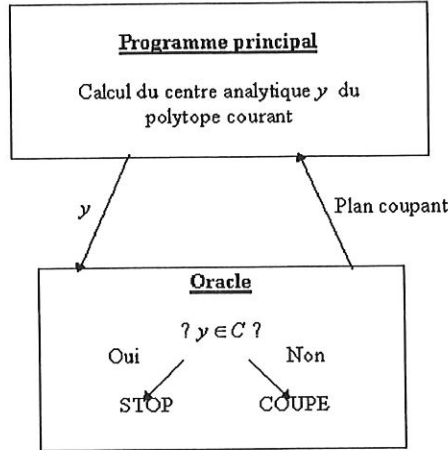
$$\mathcal{C} = \{y : f(y) \leq z^*, y \in Y, l \leq y \leq h\}.$$

La stratégie mise en oeuvre consiste à engendrer une suite de points à partir d'une suite d'approximations polyédriques de l'ensemble \mathcal{C} . De façon plus précise, on se donne x^k , un 'bon' point intérieur dans un ensemble contenant \mathcal{C} . Si $x^k \in \mathcal{C}$, alors le problème est résolu et l'algorithme se termine; sinon un hyperplan est engendré de sorte que l'ensemble contenant \mathcal{C} soit séparé en deux parties.

Si l'on est sûr que l'une des deux parties contient l'ensemble \mathcal{C} , l'autre peut être éliminée. Cela conduit à un nouvel ensemble contenant \mathcal{C} , plus petit que le précédent. On peut dès lors choisir un nouveau 'bon' point à l'intérieur du nouvel ensemble et de nouveau tester s'il appartient à \mathcal{C} .

La question qui se pose est comment sélectionner le 'bon' point intérieur de chaque ensemble contenant \mathcal{C} . Idéalement, il serait préférable de choisir un point 'centré', qui divise l'ensemble contenant \mathcal{C} en deux parties plus ou moins égales. Plusieurs propositions de 'centres' ont été faites, mais celle que nous allons retenir dans ce mémoire est celle du 'centre analytique'.

Ce choix du centre analytique comme 'bon' point intérieur, et ce principe de génération de coupes justifient le nom donné à la méthode, à savoir **la méthode ACCPM** (Analytic Center Cutting Plane Method).



La méthode ACCPM permet de résoudre d'autres problèmes que le problème d'admissibilité convexe.

Elle permet notamment de résoudre le problème de programmation convexe (P), même lorsque la valeur optimale z^* de celui-ci est inconnue [6]. Dans ce cas, le programme principal travaille sur une relaxation linéaire du problème original (P). Cette relaxation est affinée à chaque itération par adjonction d'une ou plusieurs coupes. A une itération donnée, le programme principal, au vu de la relaxation linéaire, envoie une proposition y à l'oracle. Celui-ci renvoie alors au programme principal une information sur le problème (P) en ce point, afin d'enrichir l'approximation linéaire de ce dernier.

Plus précisément, si y est admissible ($y \in Y, l \leq y \leq h$), l'oracle renvoie un plan de support de la fonction f en y . Sinon, il renvoie un plan séparant y de Y . Ces coupes sont appelées *coupes d'optimalité* dans le premier cas, et *coupes d'admissibilité* dans le second. La relaxation linéaire est donc définie par

$$\begin{cases} \min z \\ \text{s.c.} \begin{cases} z \geq f(y; y^i) & i \in I \\ g(y; y^j) \leq 0 & j \in J \\ l \leq y \leq h, \end{cases} \end{cases}$$

où $\{z \geq f(y; y^i) \ i \in I\}$ est l'ensemble des coupes d'optimalité définies en les points y^i envoyés à l'oracle, et $\{g(y; y^j) \leq 0 \ j \in J\}$ est l'ensemble des coupes d'admissibilité définies en les points y^j envoyés à l'oracle.

Nous minimisons en fait une approximation extérieure de la fonction f sur un domaine élargi de Y .

La procédure décrite plus haut prend fin lorsque le programme principal considère qu'il a atteint l'optimum.

Le point y dont il est question précédemment est en réalité le centre analytique de *l'ensemble de localisation*, défini par

$$\mathcal{L}_{\bar{z}} = \{(y, z) \in \mathcal{R}^{n+1} : z \leq \bar{z}, z \geq f(y; y^i), g(y; y^j) \leq 0, l \leq y \leq h, \forall i \in I, j \in J\},$$

où \bar{z} est une borne supérieure sur la solution optimale z^* du problème (P).

On montre que $\mathcal{L}_{\bar{z}}$ contient nécessairement la solution optimale z^* à chaque itération. La convergence de cet algorithme repose sur la mise à jour à chaque itération de $\mathcal{L}_{\bar{z}}$, d'une borne inférieure \underline{z} , et d'une borne supérieure \bar{z} . Cette démarche permet en effet d'atteindre une itération pour laquelle le saut de dualité $\bar{z} - \underline{z} < \varepsilon$, notre critère d'arrêt. (voir algorithme [6])

Citons encore le problème MinMax linéaire

$$\min_u \left\{ \max_{i \in I} \left\{ a_i - \sum_{j=1}^n g_{ij} u_j \right\} \mid l \leq u \leq h \right\},$$

qui lui aussi peut être résolu par une méthode ACCPM ([1]).

Beaucoup de questions peuvent se poser autour du fonctionnement de l'algorithme ACCPM pour la résolution du problème d'admissibilité convexe. Entre autres :

- Qu'est-ce que le centre analytique d'un polytope, et comment le générer ?
- Comment choisir les coupes de façon optimale ?
- Peut-on envisager l'ajout de plusieurs coupes simultanément ?
- Quel est l'ordre de convergence d'un tel algorithme ?

Toutes ces questions seront abordées dans ce mémoire selon le cheminement qui suit.

Dans le premier chapitre, nous formalisons la méthode ACCPM permettant de résoudre le problème d'admissibilité convexe, en introduisant les concepts de plan coupant et de centre analytique ([3],[4],[11]). Le deuxième chapitre se consacre au calcul du centre analytique d'un polytope borné à intérieur non vide. Nous envisageons ce calcul sous différents points de vue (dual, primal, et primal-dual), mais toujours suivant une méthode de réduction de potentiel [11].

Les concepts de base ayant été définis, nous abordons dans les chapitres suivants la résolution du problème d'admissibilité convexe au moyen des algorithmes ACCPM.

Le troisième chapitre traite d'un algorithme ACCPM destiné à résoudre le problème particulier d'admissibilité linéaire. L'ensemble \mathcal{C} est ici un polytope à intérieur non vide. Nous étudions à la fois la façon de concevoir l'algorithme, mais aussi sa complexité algorithmique [9]. Nous suivons le même cheminement dans les deux derniers chapitres. Le premier de ceux-ci propose un algorithme ACCPM pour le problème d'admissibilité convexe général, où une seule coupe est permise à chaque itération [4], et le second un algorithme ACCPM pour la résolution du même problème, mais où plusieurs coupes sont autorisées à chaque itération [10].

Tous ces algorithmes sont basés sur la réduction d'une fonction potentielle qui mesure en quelque sorte le volume des polyèdres contenant \mathcal{C} .

NOTATIONS :

Tout au long de ce mémoire, nous utiliserons les notations suivantes, où x et s sont deux vecteurs de \mathcal{R}^n .

- La matrice $diag(x)$ sera notée X .
- Le produit d'Hadamard (ou produit composante par composante) de x et s sera noté xs (de toute évidence, ce produit est commutatif).
- Le vecteur dont toutes les composantes sont égales à '1' sera noté e (sa dimension dépendra du contexte).

Avec ces notations, le produit scalaire de x et s pourra dès lors s'écrire :

$$x^T s = e^T X s = e^T S x = e^T (xs).$$

Nous utiliserons également les notations

- $\|\cdot\|$ pour désigner la norme Euclidienne $\|\cdot\|_2$.
- 'log' pour représenter le logarithme népérien.
- \underline{n} lorsque l'on prend tous les indices de 1 à n .

Chapitre 1

Formalisation de la méthode ACCPM.

Le but de ce chapitre est de définir de façon précise le fonctionnement de l'algorithme ACCPM en introduisant les notions de plans coupants et de centre analytique.

1.1 Plans coupants

1.1.1 Position du problème

Notre problème est de trouver un point intérieur à un ensemble convexe $\mathcal{C} \subseteq \mathcal{R}^m$, où \mathcal{C} satisfait aux hypothèses suivantes :

- 1) \mathcal{C} est contenu dans le cube $\Omega^0 = \{y \in \mathcal{R}^m : 0 \leq y \leq e\} = [0, 1]^m$
- 2) \mathcal{C} contient une boule de dimension m , et de rayon ε , où $\text{spdg } \varepsilon \leq 0.5$.

1.1.2 Principe de l'algorithme ACCPM.

Le principe de l'algorithme ACCPM est de générer des plans coupants qui permettront de limiter l'espace de recherche à un polytope contenant \mathcal{C} , et l'approximant le mieux possible.

Ainsi, partant du polytope Ω^0 défini plus haut, et de son centre analytique y^0 , nous faisons appel à un oracle (ou sous-routine). Celui-ci vérifie si y^0 appartient à $\text{int}(\mathcal{C})$ ou pas, et agit en conséquence :

-Si $y^0 \in \text{int}(\mathcal{C})$, l'algorithme se termine.

-Si $y^0 \notin \text{int}(\mathcal{C})$, l'oracle renvoie au programme principal une coupe, ou hyperplan, que nous pouvons exprimer sous la forme $\{y \in \mathcal{R}^m : a^T y \leq a^T y^0 - \gamma\} \supseteq \mathcal{C}$, où a est normalisé de telle sorte que $\|a\| = 1$, et où γ est choisi selon la profondeur que l'on souhaite donner à la coupe.

Les plans coupants sont donc générés par l'oracle selon certains critères que nous lui imposons de respecter.

Une fois cette nouvelle coupe obtenue, on recherche le centre analytique y^1 du nouveau

polytope Ω^1 ainsi défini, et on soumet ce point à l'oracle, qui de nouveau, agira selon deux possibilités.

L'algorithme prendra fin lorsque le centre analytique du polytope courant appartiendra à l'ensemble \mathcal{C} des solutions. Cela arrivera à une itération donnée, puisque l'on fait l'hypothèse que \mathcal{C} a un intérieur non vide.[cfr section (1.1.1)]

Notons que l'oracle renvoie également une coupe dans le cas où y^i (centre analytique de Ω^i) se trouve sur la frontière de \mathcal{C} . Cependant, dans un cas plus général où y^i est à l'extérieur de \mathcal{C} , on peut se demander comment choisir les paramètres a et γ qui définissent la coupe $\{y \in \mathcal{R}^m : a^T y \leq a^T y^i - \gamma\}$.

Choix du vecteur normal a :

Exprimons tout d'abord l'ensemble convexe \mathcal{C} par un système (fini ou infini) d'inégalités convexes :

$$\mathcal{C} = \{y \in \mathcal{R}^m : f_i(y) \leq 0 \quad i = 1, 2, \dots\}$$

où chaque $f_i : \mathcal{R}^m \rightarrow \mathcal{R}$ est convexe.

Il semble dès lors pertinent de choisir a comme étant $g_i / \|g_i\|$, où g_i est un sous-gradient arbitraire d'une des fonctions f_i en \bar{y} qui satisfait $f_i(\bar{y}) \geq 0$ (i.e. une des contraintes non strictement vérifiées par \bar{y} , le centre analytique du polytope courant).

g_i est donc un élément de $\partial f_i(\bar{y})$, où $\partial f_i(\bar{y})$ est l'ensemble des sous-gradients de f_i en \bar{y} , et est défini par :

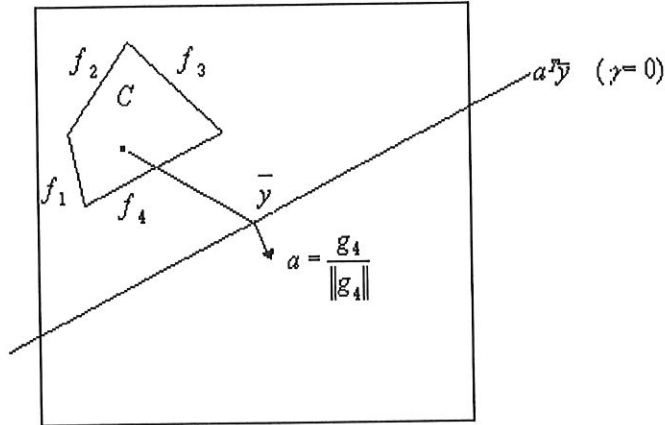
$$\partial f_i(\bar{y}) = \{g_i \in \mathcal{R}^m : f_i(y) \geq f_i(\bar{y}) + g_i^T (y - \bar{y}), \forall y \in \mathcal{R}^m\}.$$

Notons que

$$\forall g_i \in \partial f_i(\bar{y}), f_i(y) \leq f_i(\bar{y}) \Rightarrow g_i^T (y - \bar{y}) \leq 0$$

et donc,

$$\text{si } f_i(\bar{y}) \geq 0 \text{ et } f_i(y) \leq 0, \text{ alors } g_i^T (y - \bar{y}) \leq 0.$$



on choisit $a = g_4 / \|g_4\|$ puisque $f_4(\bar{y}) \geq 0$.
 Pour un y tel que $f_4(y) \leq 0$, on a bien $g_4^T(y - \bar{y}) \leq 0$.

Le choix d'une telle valeur pour a est motivé par le fait que cela assure que le nouveau polytope obtenu en ajoutant la dernière contrainte $a^T y \leq a^T \bar{y} - \gamma$ contienne toujours l'ensemble \mathcal{C} des solutions, si toutefois on choisit γ de façon à ce que la coupe n'intersecte pas \mathcal{C} !

En réalité, il n'est pas nécessaire d'imposer $g_i \in \partial f_i(\bar{y})$. Tout ce que nous voulons, c'est trouver un indice i pour lequel $f_i(\bar{y}) \geq 0$, et choisir a de telle sorte que $a^T(y - \bar{y}) \leq 0$ pour tout y dans $\{y \in \mathcal{R}^m : f_i(y) \leq 0\}$.

Ce même type d'approche s'applique également quand les fonctions f_i sont convexes et différentiables. Dans ce cas, on choisit $a = g_i / \|g_i\|$ où $g_i = \nabla f_i(\bar{y})$, le gradient d'une fonction f_i en \bar{y} telle que $f_i(\bar{y}) \geq 0$. Cela est de nouveau motivé par le fait que

$$f_i(y) \leq f_i(\bar{y}) \Rightarrow \nabla f_i(\bar{y})^T (y - \bar{y}) \leq 0,$$

puisque f_i est convexe.

Choix du terme indépendant γ :

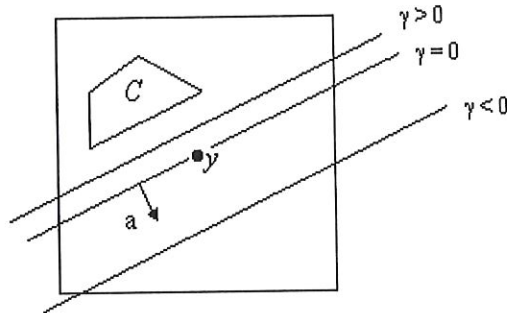
Goffin et Vial ([2]) ont proposé une terminologie pour les coupes, selon les valeurs prises par γ .

Si $\gamma > 0$, on dit que la coupe est **profonde**. (anglais : *deep*)

Si $\gamma < 0$, on dit que la coupe est **peu profonde**. (anglais : *shallow*)

Si $\gamma = 0$, la coupe passe par le centre analytique \bar{y} .

D'un point de vue graphique, cela signifie que la nouvelle coupe est plus ou moins éloignée de l'ensemble \mathcal{C} des solutions.



1.2 Centre analytique

1.2.1 Définition

Fonction potentielle duale

Soit Ω un polytope borné de \mathcal{R}^m défini par n ($> m$) inégalités linéaires, i.e.

$$\Omega = \{y \in \mathcal{R}^m : c - A^T y \geq 0\},$$

où $A \subseteq \mathcal{R}^{m \times n}$ est de rang plein, et où $c \in \mathcal{R}^n$.

Nous appellerons Ω le *polytope dual*.

Notons l'intérieur de Ω par $\text{int}(\Omega) = \{y \in \mathcal{R}^m : c - A^T y > 0\}$. Supposons que $\text{int}(\Omega) \neq \emptyset$.

Pour définir le centre analytique, nous choisissons la distance suivante :

$$d(y, \Omega) = \prod_{j=1}^n (c_j - a_j^T y), \quad y \in \Omega$$

où a_j est la j^{e} colonne de A .

La distance d'un point de Ω à Ω n'est autre que le produit des écarts aux contraintes formant le polytope Ω .

Cette distance est dès lors fonction aussi bien du point y choisi dans Ω , que des paramètres A et c .

Traditionnellement, on pose $s := c - A^T y$ et on l'appelle *variable d'écart*.

Remarquons que l'hypothèse selon laquelle A est de rang plein permet de définir une bijection entre les points y de Ω et les vecteurs d'écart s .

Le logarithme de $d(y, \Omega)$ est appelé *fonction barrière*, ou *fonction potentielle*, et s'écrit :

$$\mathcal{B}(y, \Omega) = \log d(y, \Omega) = \sum_{j=1}^n \log(c_j - a_j^T y) = \sum_{j=1}^n \log s_j.$$

Le point intérieur y^a de Ω qui maximise cette fonction est appelé le *centre analytique* de Ω . y^a est donc solution du problème

$$\begin{cases} \max & \sum_{j=1}^n \log s_j \\ \text{s.c.} & s = c - A^T y \geq 0. \end{cases} \quad (1.1)$$

Notons que ce problème est bien défini puisque l'on fait l'hypothèse que $\text{int}(\Omega) \neq \emptyset$, et que Ω est borné. La valeur optimale est donc finie, et atteinte en un point intérieur tel que $s > 0$ (pas sur la frontière car cela signifierait que le max est en un point tel que $\exists j \in \underline{n}$ tel que $s_j = 0$, et donc $\mathcal{B}(y, \Omega) = \log \prod_{j=1}^n s_j = -\infty!$)

On notera

$$\mathcal{B}(\Omega) := \mathcal{B}(y^a, \Omega) = \max_{y \in \Omega} \log d(y, \Omega).$$

y^a est donc aussi le point qui maximise le produit des écarts, puisqu'on a équivalence des problèmes

$$\max_{s=c-A^T y} \sum_{j=1}^n \log s_j \quad \text{et} \quad \max_{s=c-A^T y} \log \prod_{j=1}^n s_j \quad \text{et} \quad \max_{s=c-A^T y} \prod_{j=1}^n s_j.$$

Ce y^a est défini de façon unique puisque la fonction barrière $\mathcal{B}(y, \Omega)$ est strictement concave sur $\text{int}(\Omega)$. Par conséquent, y^a peut être obtenu en résolvant les conditions d'optimalité associées au problème(1.1).

Or, le Lagrangien du problème (1.1) est

$$\mathcal{L}(s, y; x) = - \sum_{j=1}^n \log s_j + x^T (A^T y + s - c)$$

où x est le multiplicateur de Lagrange.

Et

$$\begin{aligned} \nabla_s \mathcal{L}(s, y; x) &= -\frac{1}{s} + x & \text{où} & \quad \frac{1}{s} = \left(\frac{1}{s_1}, \frac{1}{s_2}, \dots, \frac{1}{s_n} \right) \\ \nabla_y \mathcal{L}(s, y; x) &= Ax \end{aligned}$$

Les conditions de Kuhn-Tucker sont donc :

$$\begin{aligned} xs &= e \\ Ax &= 0 & x > 0 \\ s &= c - A^T y & s > 0 \end{aligned} \tag{1.2}$$

où le produit xs signifie que l'on multiplie composante par composante. xs est donc le vecteur $(x_1s_1, x_2s_2, \dots, x_ns_n)^T$.

Remarque :

On peut définir un centre analytique même lorsque des égalités sont présentes dans le système des contraintes. Dans ce cas, Ω s'exprime :

$$\Omega = \{y \in \mathcal{R}^m : c - A^T y \geq 0, By = b\}.$$

Le centre analytique est choisi sur l'hyperplan $\{y : By = b\}$, et tel qu'il maximise le produit des variables d'écart $s = c - A^T y$.

Donc, l'intérieur de Ω ne l'est pas au sens topologique du terme.

Ici, Ω a un intérieur si :

$$\text{int}(\mathcal{R}_+^n) \cap \{s : s = c - A^T y \text{ pour } y \text{ tel que } By = b\} \neq \emptyset$$

où $\text{int}(\mathcal{R}_+^n) := \{s \in \mathcal{R}_+^n : s > 0\}$.

Ainsi, si Ω ne possède qu'un seul point intérieur y avec $s = c - A^T y$, on dira que $\text{int}(\Omega)$ est non vide.

Fonction potentielle primale

Le centre analytique peut être obtenu différemment, à partir d'une fonction potentielle primale.

Soit

$$\mathcal{X}_\Omega := \{x \in \mathcal{R}^n : Ax = 0, x \geq 0\}.$$

On définit la fonction potentielle primale sur $\text{int}(\mathcal{X}_\Omega)$ par :

$$\mathcal{P}(x, \Omega) = n \log(c^T x) - \sum_{j=1}^n \log x_j, \quad x \in \text{int}(\mathcal{X}_\Omega).$$

Cette fonction est aussi appelée *fonction potentielle de Karmarkar*.

Remarquons que $\mathcal{P}(x, \Omega)$ est homogène en x , i.e.

$$\mathcal{P}(x, \Omega) = \mathcal{P}(tx, \Omega) \quad \forall t > 0,$$

puisque

$$\begin{aligned} \mathcal{P}(tx, \Omega) &= n \log(tc^T x) - \sum_{j=1}^n \log tx_j \\ &= n(\log t + \log c^T x) - (n \log t + \sum_{j=1}^n \log x_j) \\ &= n \log c^T x - \sum_{j=1}^n \log x_j \\ &= \mathcal{P}(x, \Omega). \end{aligned}$$

Nous pouvons donc normaliser x de telle sorte que $c^T x = n$.

Considérons alors le problème de minimisation de $\mathcal{P}(x, \Omega)$ sur $\text{int}(\mathcal{X}_\Omega)$, i.e.

$$\begin{cases} \min & \mathcal{P}(x, \Omega) \\ \text{s.c.} & x \in \text{int}(\mathcal{X}_\Omega) \end{cases} \quad (1.3)$$

Comme $\mathcal{P}(x, \Omega)$ est homogène, le problème (1.3) revient donc à

$$\begin{cases} \min & \left(-\sum_{j=1}^n \log x_j\right) \\ \text{s.c.} & x \in \text{int}(\mathcal{X}_\Omega), \quad c^T x = n. \end{cases} \quad (1.4)$$

Ce problème étant convexe, sa solution peut être obtenue en résolvant les conditions de Kuhn-Tucker associées.

Or,

$$\mathcal{L}(x; y, y_0) = -\sum_{j=1}^n \log x_j - y^T(Ax) + y_0(c^T x - n)$$

où y et y_0 sont les multiplicateurs de Lagrange.

De plus,

$$\nabla_x \mathcal{L}(x; y, y_0) = -x^{-1} - A^T y + cy_0.$$

Les conditions d'optimalité s'écrivent alors

$$\begin{aligned} -x^{-1} - A^T y + cy_0 &= 0 \\ Ax &= 0 \\ c^T x &= n \quad x > 0. \end{aligned} \tag{1.5}$$

Si l'on multiplie la première équation de (1.5) par x à gauche, on a

$$\begin{aligned} 0 &= -x^T x^{-1} - x^T A^T y + c^T x y_0 \\ &= -n + n y_0 - 0 \\ &= n(y_0 - 1). \end{aligned}$$

Dès lors, $y_0 = 1$, et les conditions d'optimalité deviennent

$$\begin{aligned} -x^{-1} + s &= 0 \\ Ax &= 0 \\ A^T y + s &= c \quad s > 0 \\ c^T x &= n \quad x > 0 \end{aligned}$$

ou

$$\begin{aligned} xs &= e \\ Ax &= 0 \\ A^T y + s &= c \quad s > 0 \\ c^T x &= n \quad x > 0. \end{aligned}$$

Notons que les trois premières équations entraînent la quatrième car

$$n = e^T e = e^T(xs) = x^T(c - A^T y) = c^T x$$

puisque $Ax = 0$.

Donc, la dernière équation est redondante, et nous retrouvons les conditions d'optimalité (1.2) du problème dual (1.1), qui nous permettait de trouver le centre analytique de Ω .

Chapitre 2

Calcul du centre analytique par une méthode de réduction de potentiel.

2.1 Introduction

Ce chapitre a pour but d'exposer différentes manières de calculer le centre analytique d'un polytope borné à intérieur non vide, en tenant un raisonnement basé sur la réduction de potentiel.

La première approche du problème se fera dans l'espace dual $\Omega = \{y \in \mathcal{R}^m : A^T y \leq c\}$, où l'on envisagera le calcul du centre analytique dans trois cas : tout d'abord à partir d'un point qui est déjà centre analytique approximatif, ensuite à partir d'un point intérieur au polytope Ω , et enfin depuis un point extérieur à Ω .

L'approche suivante sera une approche *primale*, basée sur l'algorithme de Karmarkar, qui permet d'approcher le centre analytique de Ω en partant d'un point $x^0 \in \text{int}(\mathcal{K}_p)$, où $\mathcal{K}_p = \{x : Ax = 0, e^T x = n, x \geq 0\}$.

Nous envisagerons également un algorithme affiné.

Finalement, la dernière approche sera *primale-duale* et permettra d'évaluer le centre analytique de Ω depuis une paire (x, s) , où $x \in \text{int}(\mathcal{X}_\Omega)$ et $s \in \text{int}(\mathcal{S}_\Omega)$, avec

$$\begin{aligned} \text{int}(\mathcal{X}_\Omega) &= \{x : Ax = 0, x > 0\} \\ \text{int}(\mathcal{S}_\Omega) &= \{s : A^T y + s = c, s > 0\}. \end{aligned}$$

2.2 Proximité du centre analytique

Avant d'introduire des procédures numériques pour calculer le centre analytique, voyons comment exprimer la proximité d'un point par rapport au centre analytique.

A. Approche duale

Rappelons-nous que Ω est un polytope borné dans \mathcal{R}^m , défini par $n(> m)$ inégalités, i.e.

$$\Omega = \{y \in \mathcal{R}^m : c - A^T y \geq 0\}.$$

Pour un point $y \in \text{int}(\Omega)$, la *fonction potentielle duale* de Ω est

$$\mathcal{B}(y, \Omega) = \sum_{j=1}^n \log s_j, \quad s = c - A^T y$$

que l'on notera simplement $\mathcal{B}(y)$ par la suite.

Idéalement, on devrait mesurer la proximité de y ($\in \Omega$) à y^a (centre analytique de Ω) par la différence

$$\mathcal{B}(y^a) - \mathcal{B}(y) = \max_{y \in \Omega} \mathcal{B}(y) - \mathcal{B}(y).$$

Le problème est que l'on ne connaît pas y^a , ni $\mathcal{B}(y^a)$! Nous utiliserons donc une autre méthode.

On sait que y^a est l'unique maximum de $\mathcal{B}(y)$ sur Ω (car $\mathcal{B}(y)$ est strictement concave). Dès lors, $\nabla \mathcal{B}(y^a) = 0$.

Une mesure de proximité pourrait donc être le résidu $\nabla \mathcal{B}(y)$ en y .

Notons que

$$\nabla_y \mathcal{B}(y) = \nabla_y \left(\sum_{j=1}^n \log(c_j - a_j^T y) \right) = -AS^{-1}e$$

et

$$\nabla^2 \mathcal{B}(y) = \begin{pmatrix} (\mathcal{B}(y))''_{y_1^2} & \dots & (\mathcal{B}(y))''_{y_1 y_m} \\ & \ddots & \\ (\mathcal{B}(y))''_{y_m y_1} & \dots & (\mathcal{B}(y))''_{y_m^2} \end{pmatrix}$$

où par exemple,

$$(\mathcal{B}(y))''_{y_1^2} = -(-a_{11})(s_1)^{-2}(-a_{11}) - (-a_{12})(s_2)^{-2}(-a_{12}) - \dots - (-a_{1n})(s_n)^{-2}(-a_{1n}),$$

et donc

$$\begin{aligned} \nabla^2 \mathcal{B}(y) &= - \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} s_1^{-2} & \dots & 0 \\ & \ddots & \\ 0 & \dots & s_n^{-2} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \\ &= -AS^{-2}A^T \end{aligned}$$

où $s = c - A^T y > 0$.

Donc, la condition d'optimalité associée au problème $\max_{y \in \Omega} \mathcal{B}(y)$ est

$$\nabla \mathcal{B}(y) = -AS^{-1}e = 0.$$

Supposons maintenant que l'on veuille approcher la solution optimale de ce problème (i.e. le centre analytique) en utilisant la méthode de Newton .

La direction de Newton Δy dans l'espace dual, associée au problème $\max_{y \in \Omega} \mathcal{B}(y)$ est donnée par

$$\Delta y = -(\nabla^2 \mathcal{B}(y))^{-1} \nabla \mathcal{B}(y) = -(AS^{-2}A^T)^{-1} AS^{-1}e. \quad (2.1)$$

En termes d'écarts, on obtient donc

$$\Delta s = -A^T \Delta y = A^T (AS^{-2}A^T)^{-1} AS^{-1}e \quad (2.2)$$

puisque $(\Delta y, \Delta s)$ est admissible seulement si $\Delta s + A^T \Delta y = 0$ (i.e. $A^T(y + \Delta y) + (s + \Delta s) = c$).

Δs est en réalité la direction de Newton en termes d'écarts, projetée sur l'ensemble admissible $\{y : A^T y + s = c\}$. Elle est obtenue par la résolution du problème $\{\min \frac{1}{2} \Delta s^T S^{-2} \Delta s - \Delta s^T s^{-1} : \Delta s + A^T \Delta y = 0\}$ où la fonction objectif est l'approximation quadratique de $\mathcal{B}(y, \Omega) = \sum_{j=1}^n \log s_j$.

Définissons alors $p(s)$ tel que $\Delta s = -sp(s)$. On obtient

$$\begin{aligned} p(s) &= -S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}e \\ \delta(s)^2 &= \|p(s)\|^2 = e^T S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}e. \end{aligned} \quad (2.3)$$

On remarque que $\delta(s) = 0 \Rightarrow \nabla \mathcal{B}(y) = 0$, car la matrice $S^{-1}A^T(AS^{-2}A^T)^{-1}$ est de rang plein puisque A l'est par hypothèse, et S , matrice diagonale des écarts $s > 0$, l'est aussi. Ainsi, si $\delta(s) = 0$, alors $s = s^a$ et $y = y^a$, le centre analytique de Ω .

Remarquons que $\delta(s)$ peut s'exprimer de manière plus élégante si l'on définit

$$x(s) = S^{-1}(I - S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1})e. \quad (2.4)$$

Ce point est admissible puisque $Ax(s) = 0$.

Nous vérifions alors que

$$\begin{aligned} p(s) &= -S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}e = Sx(s) - e \\ \delta(s) &= \|p(s)\| = \|Sx(s) - e\|. \end{aligned} \quad (2.5)$$

Or, nous avons vu que si $\delta(s) = 0$, on se trouve au centre analytique de Ω . Ceci est bien confirmé par le fait que dans ce cas, $Sx(s) = e$, où $s = c - A^T y > 0$ est admissible, et $x(s) > 0$ également puisque $Ax(s) = 0$. Et donc, les conditions d'optimalité associées aux problèmes (1.1) et (1.4) sont vérifiées, ce qui signifie bien que $y = y^a$, $s = s^a$, et $x(s^a) = x^a$, solution du problème (1.4).

On appellera *paire centrale* de Ω la paire (x^a, s^a) .

B. Approche primale

Une autre mesure de proximité peut se faire via la direction de Newton associée au problème primal (1.3), plutôt que celle associée au problème dual (1.1).

Comme $\mathcal{P}(x, \Omega)$, notée $\mathcal{P}(x)$, est homogène, on peut fixer $c^T x = n$.

Dès lors,

$$\begin{aligned} \nabla \mathcal{P}(x) &= \nabla(n \log c^T x - \sum_{j=1}^n \log x_j) \\ &= \frac{n}{c^T x} c - X^{-1} e \\ &= c - X^{-1} e \end{aligned} \quad (2.6)$$

et

$$\nabla^2 \mathcal{P}(x) = X^{-2}.$$

La direction de Newton associée au problème primal (1.3) est donc obtenue par la résolution du problème

$$\begin{cases} \min (c - X^{-1}e)^T \Delta x + \frac{1}{2} \Delta x^T X^{-2} \Delta x \\ \text{s.c. } A \Delta x = 0 \end{cases}$$

où la fonction objectif n'est autre que l'approximation quadratique de $\mathcal{P}(x)$, et où la contrainte assure le caractère admissible de la direction Δx .

Ce problème étant convexe, sa solution est aussi celle des conditions de Kuhn-Tucker associées.

Sachant que le Lagrangien du problème est

$$\mathcal{L}(\Delta x; u) = (c - X^{-1}e)^T \Delta x + \frac{1}{2} \Delta x^T X^{-2} \Delta x + u^T (A \Delta x),$$

où u est le vecteur des multiplicateurs de Lagrange, ces conditions de Kuhn-Tucker s'écrivent

$$X^{-2} \Delta x + c - X^{-1}e + A^T u = 0 \quad (2.7)$$

$$A \Delta x = 0. \quad (2.8)$$

En multipliant (2.7) par AX^2 , nous obtenons

$$A \Delta x + AX^2(c - X^{-1}e) + AX^2 A^T u = 0, \quad \text{où } A \Delta x = 0.$$

$AX^2 A^T = (AX)(AX)^T$ étant inversible, nous pouvons extraire u :

$$u = -(AX^2 A^T)^{-1} AX^2 (c - X^{-1}e).$$

Si nous injectons cette valeur de u dans (2.7), nous avons

$$\Delta x = -X^2(c - X^{-1}e) + X^2 A^T (AX^2 A^T)^{-1} AX^2 (c - X^{-1}e).$$

En définissant $p(x) = -X^{-1} \Delta x$, on obtient

$$\begin{aligned} p(x) &= X(c - X^{-1}e) - X A^T (AX^2 A^T)^{-1} AX^2 (c - X^{-1}e) \\ &= (I - X A^T (AX^2 A^T)^{-1} AX)(Xc - e). \end{aligned} \quad (2.9)$$

On définit comme précédemment

$$\begin{aligned} \delta(x)^2 &:= \|p(x)\|^2 \\ &= (Xc - e)^T (I - X A^T (AX^2 A^T)^{-1} AX)(Xc - e). \end{aligned} \quad (2.10)$$

Posons ensuite

$$y(x) = (AX^2 A^T)^{-1} AX(Xc - e) \quad \text{et} \quad s(x) = c - A^T y(x). \quad (2.11)$$

Nous avons donc

$$\begin{aligned} p(x) &= X[c - A^T ((AX^2 A^T)^{-1} AX(Xc - e))] - e \\ &= Xs(x) - e \end{aligned} \quad (2.12)$$

et

$$\delta(x) = \|Xs(x) - e\|.$$

De nouveau, si $\|Xs(x) - e\| = 0$, alors $\delta(x) = 0$ et la direction de Newton projetée sur l'ensemble admissible Δx est nulle. Nous avons donc atteint le minimum de $\mathcal{P}(x)$ sur $\mathcal{X}_\Omega = \{x : Ax = 0, x > 0\}$.

Ainsi, $x = x^a$ et $y(x^a) = y^a$, la paire centrale de Ω .

C. Approche primale-duale

La troisième mesure de proximité est obtenue en utilisant à la fois le primal et le dual.

Pour un $x \in \mathcal{X}_\Omega = \{x : Ax = 0, x \geq 0\}$, et un $y \in \Omega$, ou $s = c - A^T y \in \mathcal{S}_\Omega = \{s : s = c - A^T y, s \geq 0\}$, cette troisième mesure peut être définie par

$$\eta(x, s) = \|X(c - A^T y) - e\| = \|Xs - e\|.$$

Grâce aux conditions d'optimalité (1.2) des problèmes (1.1) et (1.3) étudiés au chapitre 1, nous savons que si $\eta(x, s) = 0$, alors $x = x^a, y = y^a$.

Introduisons alors la *fonction potentielle primale-duale* homogène, définie par

$$\psi_n(x, s) = n \log(x^T s) - \sum_{j=1}^n \log(x_j s_j). \quad (2.13)$$

Notons que pour toute paire (x, s) admissible, $\psi_n(x, s) \geq n \log n$ et qu'au centre analytique, $\psi_n(x, s) = n \log n$.

En effet, nous savons que x^a est le point qui maximise $\sum_{j=1}^n \log x_j$ sur \mathcal{X}_Ω et que s^a est le point qui maximise $\sum_{j=1}^n \log s_j$ sur \mathcal{S}_Ω , donc la paire (x^a, s^a) est la paire qui maximise $\sum_{j=1}^n \log(x_j s_j)$ sur $(\mathcal{X}_\Omega \times \mathcal{S}_\Omega)$.

Or, au centre analytique, cette fonction s'annule car $x_j^a s_j^a = 1 \quad \forall j$.

Dès lors, en une paire admissible (x, s) tel que $x^T s = n$, $\sum_{j=1}^n \log(x_j s_j) \leq 0$ et $\psi_n(x, s) \geq n \log n$.

Comme $\psi_n(x, s)$ est homogène, cette dernière inégalité est satisfaite en toute paire (x, s) admissible.

Pour établir une relation entre ces trois mesures de proximité $\delta(x), \delta(s)$, et $\eta(x, s)$, nous aurons besoin du lemme suivant.

Lemme 2.2.1

Si $x \in (\mathcal{R}^n)^+$ est tel que $\|x\|_\infty < 1$, alors

$$e^T x \geq \sum_{i=1}^n \log(1 + x_i) \geq e^T x - \frac{\|x\|^2}{2(1 - \|x\|_\infty)}.$$

preuve : [8]

Comme, pour tout i , $\log(1 + x_i) \leq x_i$, la première inégalité s'obtient directement par sommation.

Pour montrer la deuxième inégalité, remarquons d'abord que pour chaque i ,

$$\begin{aligned} \log(1 + x_i) &= x_i - \frac{x_i^2}{2} + \frac{x_i^3}{3} - \frac{x_i^4}{4} + \dots \\ &= x_i - \frac{x_i^2}{2} \left(1 - \frac{2}{3}x_i + \frac{2}{4}x_i^2 - \frac{2}{5}x_i^3 + \dots\right) \end{aligned}$$

et donc

$$\begin{aligned} \log(1 + x_i) &\geq x_i - \frac{x_i^2}{2} \left(1 + \frac{2}{3}|x_i| + \frac{2}{4}|x_i|^2 + \frac{2}{5}|x_i|^3 + \dots\right) \\ &\geq x_i - \frac{x_i^2}{2} (1 + |x_i| + |x_i|^2 + |x_i|^3 + \dots) \\ &= x_i - \frac{x_i^2}{2(1 - |x_i|)} \\ &\geq x_i - \frac{x_i^2}{2(1 - \|x\|_\infty)}. \end{aligned}$$

En sommant sur tous les i , on trouve

$$\sum_{i=1}^n \log(1 + x_i) \geq e^T x - \frac{\|x\|^2}{2(1 - \|x\|_\infty)}.$$

CQFD.

Définissons encore la propriété de convergence quadratique.

Définition 2.2.1

- En un point admissible intérieur, un pas primal Δx satisfait la propriété de convergence quadratique si $x^+ = x + \Delta x$ est admissible intérieur et si

$$\delta(x^+) \leq \delta(x)^2, \quad \text{i.e.} \quad \|e - X^+ s(x^+)\| \leq \|e - Xs(x)\|^2. \quad (2.14)$$

- En un point admissible intérieur, un pas dual Δs satisfait la propriété de convergence quadratique si $s^+ = s + \Delta s$ est admissible intérieur et si

$$\delta(s^+) \leq \delta(s)^2, \text{ i.e. } \|e - S^+x(s^+)\| \leq \|e - Sx(s)\|^2. \quad (2.15)$$

- En un point admissible intérieur, un pas primal-dual $(\Delta x, \Delta s)$ satisfait la propriété de convergence quadratique si $(x^+, s^+) = (x + \Delta x, s + \Delta s)$ est admissible intérieure et si

$$\eta(x^+, s^+) \leq \eta(x, s)^2, \text{ i.e. } \|e - X^+s^+\| \leq \|e - Xs\|^2. \quad (2.16)$$

Le théorème suivant établit une relation entre les trois mesures de proximité. De plus, il montre que les pas de Newton primaux et duaux définis respectivement en des points x tels que $\delta(x) \leq \theta < 1$ et s tels que $\delta(s) \leq \theta < 1$ vérifient la propriété de convergence quadratique.

Théorème 2.1

Soit (y, s) un point intérieur, y^a le centre analytique de Ω et s^a l'écart associé; et soit x un point intérieur de \mathcal{X}_Ω et x^a le minimiseur du potentiel primal, avec $c^T x^a = n$.

i) Pour toute paire admissible

$$\delta(s) \leq \eta(x, s) \quad \text{et} \quad \delta(x) \leq \eta(x, s).$$

Si $\delta(s) < 1$, alors $x(s) \in \text{int}(\mathcal{X}_\Omega)$ et

$$\eta(x(s), s) = \delta(s).$$

Si $\delta(x) < 1$, alors $y(x) \in \text{int}(\Omega)$ et

$$\eta(x, s(x)) = \delta(x).$$

ii) Si $\eta(x, s) < 1$, alors il existe un $\hat{x} \geq 0$ avec $A\hat{x} = 0$, et $c^T \hat{x} = s^T \hat{x} = n$ tel que

$$\eta(\hat{x}, s) \leq \frac{\eta(x, s)}{\sqrt{1 - \frac{\eta(x, s)^2}{n}}}. \quad (2.17)$$

iii) Si $\eta(x, s) < 1$, avec $c^T x = s^T x = n$, alors

$$\psi_n(x, s) - \psi_n(x^a, s^a) \leq \frac{\eta(x, s)^2}{2(1 - \eta(x, s))} \quad (2.18)$$

$$\mathcal{P}(x) - \mathcal{P}(x^a) \leq \frac{\eta(x, s)^2}{2(1 - \eta(x, s))} \quad (2.19)$$

$$\mathcal{B}(y^a) - \mathcal{B}(y) \leq \frac{\eta(x, s)^2}{2(1 - \eta(x, s))}. \quad (2.20)$$

iv) Soit

$$\Delta y = -(AS^{-2}A^T)^{-1}AS^{-1}e, \quad y^+ = y + \Delta y, \quad s^+ = c - A^T y^+.$$

Alors, si $\delta(s) < 1$, $s^+ = s + \Delta s$ est admissible intérieur et

$$\delta(s^+) \leq \delta(s)^2. \quad (2.21)$$

Soit

$$\Delta x = -X(I - XA^T(AX^2A^T)^{-1}AX)(Xc - e) \quad \text{et} \quad x^+ = x + \Delta x.$$

Alors, si $\delta(x) < 1$, x^+ est admissible intérieur et

$$\delta(x^+) \leq \delta(x)^2. \quad (2.22)$$

v) Si $\eta(x, s) < 1$, alors

$$\|S^{-1}s^a - e\| \leq \frac{\eta(x, s)}{1 - \eta(x, s)} \quad (2.23)$$

$$\|X^{-1}x^a - e\| \leq \frac{\eta(x, s)}{1 - \eta(x, s)}. \quad (2.24)$$

preuve :

i) Soit $s > 0$. On vérifie que $\delta(s)$ (cfr 2.3) est la norme minimale et que $x(s)$ (cfr 2.4) est le minimum du problème aux moindres carrés

$$\begin{cases} \min \|Sx - e\|^2 \\ \text{s.c. } Ax = 0. \end{cases} \quad (2.25)$$

En effet, étant en présence d'un problème convexe, sa solution peut être obtenue grâce aux conditions nécessaires et suffisantes de Kuhn-Tucker.

Sachant que le Lagrangien vaut

$$\begin{aligned} \mathcal{L}(x; \lambda) &= (Sx - e)^T(Sx - e) - (Ax)^T \lambda \\ &= x^T S^2 x - 2s^T x + n - x^T A^T \lambda, \end{aligned}$$

où λ est le vecteur des multiplicateurs de Lagrange, les conditions de Kuhn-Tucker s'écrivent

$$\begin{cases} 2S^2 x - 2s - A^T \lambda = 0 \\ Ax = 0 \end{cases}$$

ou

$$\begin{cases} x = S^{-1}e + S^{-1} \frac{S^{-1}A^T \lambda}{2} \\ Ax = 0. \end{cases}$$

La solution de ce système est

$$\begin{cases} \lambda = -2(AS^{-2}A^T)^{-1}AS^{-1}e. \\ x = x(s) = S^{-1}(I - S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1})e. \end{cases}$$

La norme minimale est donc effectivement $\delta(s) = \|Sx(s) - e\|$.

Cette propriété de $x(s)$ nous permet de dire que, puisque le x dans $\eta(x, s)$ est admissible pour le problème (2.25), nous devons avoir

$$\delta(s) = \|Sx(s) - e\| = \eta(x(s), s) \leq \eta(x, s).$$

Supposons ensuite que $\delta(s) = \|Sx(s) - e\| < 1$. Alors, le point $x(s) \in \text{int}(\mathcal{X}_\Omega)$, car $s > 0$ par hypothèse. De plus, par définition, on a $\delta(s) = \eta(x(s), s)$.

Démontrons maintenant la seconde partie de i).

Soit $x > 0$. On vérifie que $\delta(x)$ (cfr 2.10) est la valeur minimale et $(y(x), s(x))$ (cfr 2.11) la solution du problème aux moindres carrés

$$\begin{cases} \min \|Xs - e\| \\ \text{s.c. } s = c - A^T y. \end{cases} \quad (2.26)$$

En effet, ce problème peut se mettre sous la forme d'un problème sans contrainte

$$\begin{aligned} & \min_y \|X(c - A^T y) - e\| \\ \Leftrightarrow & \min_y \|X(c - A^T y) - e\|^2 \\ \Leftrightarrow & \min_y \|(Xc - e) - XA^T y\|^2 \end{aligned} \quad (2.27)$$

où

$$\begin{aligned} \|(Xc - e) - XA^T y\|^2 &= \|Xc - e\|^2 + y^T AX^2 A^T y - 2(Xc - e)^T (XA^T y) \\ &= \|Xc - e\|^2 + y^T AX^2 A^T y - 2y^T AX(Xc - e). \end{aligned} \quad (2.28)$$

La condition d'optimalité associée à ce problème est

$$\nabla_y (\|Xc - e\|^2 + y^T AX^2 A^T y - 2y^T AX(Xc - e)) = 0$$

i.e.

$$\begin{aligned} & 2AX^2 A^T y - 2AX(Xc - e) = 0 \\ \Leftrightarrow & AX^2 A^T y = AX(Xc - e) \\ \Leftrightarrow & y = (AX^2 A^T)^{-1} AX(Xc - e) \end{aligned}$$

où $(AX^2 A^T)$ est inversible car définie positive.

Ainsi, $y = y(x)$ (cfr 2.11), et donc $s = c - A^T y = c - A^T y(x) = s(x)$.

$\delta(x) = \|Xs(x) - e\|$ est donc effectivement la norme minimale du problème (2.26).

Cette propriété de $(y(x), s(x))$ nous permet de dire que, puisque la paire (x, s) dans $\eta(x, s)$ est un point admissible pour le problème (2.26), nous devons avoir

$$\delta(x) = \eta(x, s(x)) \leq \eta(x, s).$$

Aussi, si $\delta(x) = \|Xs(x) - e\| < 1$, le point $y(x) \in \text{int}(\Omega)$, i.e. $s(x) > 0$, car $x > 0$ par hypothèse. De plus, par définition, on a $\delta(x) = \eta(x, s(x))$.

ii) Soit $\hat{x} = (\frac{n}{x(s)^T s})x(s)$ et $\eta(\hat{x}, s) = \|S\hat{x} - e\|$. Alors

$$\begin{aligned} A\hat{x} &= \frac{n}{x(s)^T s} Ax(s) = 0 \quad \text{car } Ax(s) = 0 \\ c^T \hat{x} &= \frac{n}{x(s)^T s} c^T x(s) = n \end{aligned}$$

puisque $x(s)^T s = x(s)^T (c - A^T y) = c^T x(s) - (Ax(s))^T y = c^T x(s)$. Par conséquent,

$$\begin{aligned} \delta(s)^2 &= \|Sx(s) - e\|^2 \\ &= \|Sx(s) - \frac{x(s)^T s}{n} e + \frac{x(s)^T s}{n} e - e\|^2 \\ &= \|Sx(s) - \frac{x(s)^T s}{n} e\|^2 + \|\frac{x(s)^T s}{n} e - e\|^2 \\ &= \|S\hat{x} - e\|^2 (\frac{x(s)^T s}{n})^2 + (1 - \frac{x(s)^T s}{n})^2 \|e\|^2 \\ &= \eta(\hat{x}, s)^2 (\frac{x(s)^T s}{n})^2 + (1 - \frac{x(s)^T s}{n})^2 n \end{aligned} \quad (2.29)$$

où le double produit n'apparaît pas dans la troisième égalité car on vérifie que

$$(Sx(s) - \frac{x(s)^T s}{n} e)^T (\frac{x(s)^T s}{n} e - e) = \frac{x(s)^T s}{n} x(s)^T s - x(s)^T s - (\frac{x(s)^T s}{n})^2 n + \frac{x(s)^T s}{n} n = 0.$$

Considérons l'équation (2.29) comme une équation quadratique en la variable $\frac{x(s)^T s}{n}$. Comme elle a une racine réelle (puisque $x(s), s \in \mathcal{R}^n$ et $\frac{x(s)^T s}{n} \in \mathcal{R}$), le discriminant est positif, i.e.

$$4n^2 - 4(\eta(\hat{x}, s)^2 + n)(n - \delta(s)^2) \geq 0$$

ou

$$\eta(\hat{x}, s)^2 \leq \frac{n^2}{n - \delta(s)^2} - n = \frac{n\delta(s)^2}{n - \delta(s)^2} \leq \frac{n\eta(x, s)^2}{n - \eta(x, s)^2} = \frac{\eta(x, s)^2}{1 - \eta(x, s)^2/n} \quad (2.30)$$

où la deuxième inégalité est obtenue grâce au point i) du théorème, applicable puisque $\eta(x, s) < 1$ par hypothèse.

L'inégalité (2.17) se déduit facilement de (2.30).

iii) Soit $\eta = \eta(x, s) = \|xs - e\| < 1$.
 Par le lemme (2.2.1), et comme $c^T x = s^T x = n$,

$$\begin{aligned}
 \sum_{j=1}^n \log(x_j s_j) &= \sum_{j=1}^n \log(1 + (x_j s_j - 1)) \\
 &\geq e^T (Xs - e) - \frac{\|xs - e\|^2}{2(1 - \|xs - e\|_\infty)} \\
 &\geq e^T Xs - n - \frac{\eta^2}{2(1 - \eta)} \quad \text{puisque } \|xs - e\|_\infty < \eta \\
 &= -\frac{\eta^2}{2(1 - \eta)} \quad \text{puisque } x^T s = n.
 \end{aligned} \tag{2.31}$$

Notons x^a et y^a ($s^a = c - A^T y^a$) la paire centrale de Ω .
 On a donc $X^a s^a - e = 0$, et par la définition de la fonction potentielle primale-duale ψ_n (cfr 2.13)

$$\psi_n(x, s) - \psi_n(x^a, s^a) = \sum_{j=1}^n \log(x_j^a s_j^a) - \sum_{j=1}^n \log(x_j s_j) \tag{2.32}$$

$$= -\sum_{j=1}^n \log(x_j s_j) \quad \text{puisque } x^a s^a - e = 0 \tag{2.33}$$

$$\leq \frac{\eta^2}{2(1 - \eta)} \quad \text{par (2.31)}. \tag{2.34}$$

Pour obtenir les résultats équivalents (2.19) et (2.20) associés aux fonctions potentielles primale et duale, décomposons la partie droite de (2.32)

$$\begin{aligned}
 \psi_n(x, s) - \psi_n(x^a, s^a) &= \sum_{j=1}^n \log x_j^a - \sum_{j=1}^n \log x_j + \sum_{j=1}^n \log s_j^a - \sum_{j=1}^n \log s_j \\
 &= \mathcal{P}(x) - \mathcal{P}(x^a) + \mathcal{B}(y^a) - \mathcal{B}(y)
 \end{aligned} \tag{2.35}$$

puisque $\mathcal{P}(x) = n \log(c^T x) - \sum_{j=1}^n \log(x_j)$, où $c^T x = n$ par hypothèse.

Comme y^a maximise $\mathcal{B}(y)$ sur $\text{int}(\Omega)$ et que x^a minimise $\mathcal{P}(x)$ sur $\text{int}(\mathcal{X}_\Omega)$, nous avons

$$\mathcal{B}(y^a) - \mathcal{B}(y) \geq 0 \quad \text{et} \quad \mathcal{P}(x) - \mathcal{P}(x^a) \geq 0.$$

Dès lors, par (2.34) et (2.35), nous avons aussi

$$\begin{aligned}
 \mathcal{B}(y^a) - \mathcal{B}(y) &\leq \frac{\eta^2}{2(1 - \eta)} \\
 \text{et } \mathcal{P}(x) - \mathcal{P}(x^a) &\leq \frac{\eta^2}{2(1 - \eta)}
 \end{aligned}$$

ce qui démontre iii).

iv) Démontrons tout d'abord (2.21).

Nous savons, grâce à i), que $x(s)$ est solution du problème $\{\min \|Sx - e\| \text{ s.c. } Ax = 0\}$.

Dès lors, nous avons aussi

$$\delta(s^+) = \|S^+x(s^+) - e\| \leq \|S^+x(s) - e\|. \quad (2.36)$$

Mais

$$s^+ = c - A^T y^+ = c - A^T(y + \Delta y) = s - A^T \Delta y = s + \Delta s = s - Sp(s),$$

où

$$\begin{aligned} p(s) &= -S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1} = Sx(s) - e \quad (\text{par (2.3) et (2.5)}) \\ \Delta s &= -Sp(s) = A^T(AS^{-2}A^T)^{-1}AS^{-1}e = -A^T \Delta y. \end{aligned} \quad (2.37)$$

Ainsi,

$$s^+ = s - Sp(s) = s - S(Sx(s) - e) = 2s - S^2x(s). \quad (2.38)$$

Dès lors,

$$\begin{aligned} \|S^+x(s) - e\|^2 &= \|(2S - S^2X(s))x(s) - e\|^2 \\ &= \|(Sx(s) - e)\|^2 \\ &= \sum_{j=1}^n (s_j x_j(s) - 1)^4 \\ &\leq \left(\sum_{j=1}^n (s_j x_j(s) - 1)^2 \right)^2 \\ &= \|Sx(s) - e\|^4 \\ &= \delta(s)^4. \end{aligned} \quad (2.39)$$

Et donc, par (2.36) et (2.39), on a $\delta(s^+) \leq \delta(s)^2$.

Reste à prouver que $s^+ = s + \Delta s$ est admissible intérieur. Or, par hypothèse, $\delta(s) = \|p(s)\| = \|Sx(s) - e\| < 1$, et donc on vérifie également que $\|p(s)\|_\infty < 1$.

Dès lors, $s^+ = s - sp(s) = s(e - p(s)) > 0$, et s^+ est un point intérieur de \mathcal{S}_Ω .

Pour ce qui est de l'admissibilité de s^+ , elle est immédiate puisque

$$A^T y^+ + s^+ = A^T(y + \Delta y) + (s + \Delta s) = c + (A^T \Delta y + \Delta s) = c \quad \text{par (2.37)}. \quad (2.40)$$

Démontrons maintenant (2.22).

Grâce à la preuve de i), nous savons que $s(x)$ est la solution du problème

$\{\min \|Xs - e\| \text{ s.c. } s = c - A^T y\}$.

Dès lors, nous avons aussi

$$\delta(x^+) = \|X^+s(x^+) - e\| \leq \|X^+s(x) - e\|. \quad (2.41)$$

Or, $x^+ = x + \Delta x = x - Xp(x)$, où

$$p(x) = (I - XA^T(AX^2A^T)^{-1}AX)(Xc - e) = Xs(x) - e \quad (\text{par (2.9) et (2.12)}).$$

Dès lors,

$$x^+ = x - xp(x) = x - X(Xs(x) - e) = 2x - X^2s(x) \quad (2.42)$$

et donc, de la même manière que dans la preuve de (2.21), on obtient

$$\|X^+s(x) - e\|^2 \leq \delta(x)^4. \quad (2.43)$$

Par (2.41) et (2.43), on obtient $\delta(x^+) \leq \delta(x)^2$.

Reste à prouver que $x^+ = x - xp(x)$ est admissible intérieur. Or, on sait par hypothèse que $\delta(x) = \|p(x)\| = \|Xs(x) - e\| < 1$, et donc on vérifie également que $\|p(x)\|_\infty < 1$. D'où $x^+ = x - xp(x) = x(e - p(x)) > 0$, et x^+ est intérieur à \mathcal{X}_Ω .

Quant à l'admissibilité de x^+ , elle est immédiate puisque $Ax^+ = A(x + \Delta x) = Ax = 0$ (par (2.8)).

v) L'hypothèse $\eta(x, s) < 1$ du point i) étant vérifiée, on a

$$\delta(s) \leq \eta(x, s) < 1. \quad (2.44)$$

Posons $y^0 = y$, $s^0 = c - A^T y^0$, et la séquence $\{y^k, s^k\}$ générée par la méthode de Newton duale.

On montre par induction que

$$\|(S^0)^{-1}s^k - e\| \leq \sum_{j=1}^{2^k-1} \delta(s)^j. \quad (2.45)$$

De toute évidence, cela est vrai pour $k = 0$.

Supposons que ce le soit pour k , et montrons que cela l'est aussi pour $k + 1$.

Comme nous savons que

$$s^{k+1} = s^k - S^k p(s^k) = s^k - S^k(S^k x(s^k) - e),$$

nous avons

$$\begin{aligned} \|(S^0)^{-1}(s^{k+1} - s^k)\| &= \|(S^0)^{-1}S^k(S^k x(s^k) - e)\| \\ &\leq \|(S^0)^{-1}S^k\| \|S^k x(s^k) - e\| \\ &\leq \left(\sum_{j=0}^{2^k-1} \delta(s)^j \right) \delta(s^k) \\ &\leq \left(\sum_{j=0}^{2^k-1} \delta(s)^j \right) \delta(s)^{2^k} \quad (\text{par (2.44) et (2.21)}) \end{aligned} \quad (2.46)$$

où (2.46) est obtenue par hypothèse de récurrence et car, $(S^0)^{-1}S^k$ étant diagonale,

$$\|(S^0)^{-1}S^k\| = \|(s^0)^{-1}s^k\|_\infty = (s_i^0)^{-1}s_i^k \leq |(s_i^0)^{-1}s_i^k - 1| + 1 \leq \sum_{j=1}^{2^k-1} \delta(s)^j + 1 = \sum_{j=0}^{2^k-1} \delta(s)^j$$

où l'indice i correspond à la composante maximale de $(s^0)^{-1}s^k$.

Donc,

$$\begin{aligned} \|(S^0)^{-1}s^{k+1} - e\| &\leq \|(S^0)^{-1}s^k - e\| + \|(S^0)^{-1}(s^{k+1} - s^k)\| \\ (\text{par HR}) &\leq \left(\sum_{j=1}^{2^k-1} \delta(s)^j\right) + \left(\sum_{j=0}^{2^k-1} \delta(s)^j\right)\delta(s)^{2^k} \\ &= \sum_{j=1}^{2^k-1} \delta(s)^j + \delta(s)^{2^k} + \sum_{j=2^{k+1}-1}^{2^{k+1}-1} \delta(s)^j \\ &= \sum_{j=1}^{2^{k+1}-1} \delta(s)^j. \end{aligned}$$

La relation (2.45) est donc vérifiée pour $k+1$.

Grâce à cette relation, on obtient

$$\begin{aligned} \|(S^0)^{-1}s^{k+1} - e\| &\leq \frac{\delta(s) - \delta(s)^{2^{k+1}}}{1 - \delta(s)} \quad (\text{car progression géométrique}) \\ &\leq \frac{\delta(s)}{1 - \delta(s)} \quad \text{car } \delta(s) < 1 \quad \text{par hyp.} \\ &\leq \frac{\eta(x, s)}{1 - \eta(x, s)} \quad \text{par (2.44).} \end{aligned}$$

Comme cette propriété est vérifiée à chaque itération, elle le sera également en $s^a = \lim_{k \rightarrow \infty} s^k$, le centre analytique de \mathcal{S}_Ω . On a donc prouvé (2.23).

Pour prouver (2.24), on utilise une démarche tout à fait similaire, qui consiste à démontrer par induction, et sous l'hypothèse $\delta(x) \leq \eta(x, s) < 1$ (par i)), que

$$\|(X^0)^{-1}x^k - e\| \leq \sum_{j=1}^{2^k-1} \delta(x)^j \quad (2.47)$$

où $x^0 = x$ et $\{x^k\}$ est la séquence engendrée par la méthode de Newton primale.

CQFD.

2.3 Algorithmes de calcul du centre analytique

Avant toute chose, il est important d'introduire une définition primordiale pour la compréhension du raisonnement, celle de (η) -centre (analytique) approximatif.

Définition 2.3.1

Soit (\hat{x}, \hat{s}) une paire intérieure telle que $\hat{x} > 0$ et $\hat{s} = c - A^T \hat{y} > 0$ pour un certain \hat{y} , et soit $\eta > 0$. On dira que $\hat{y} \in \Omega$ est un η -centre si

$$\begin{cases} \|e - \hat{x}\hat{s}\| \leq \eta \\ A^T \hat{y} + \hat{s} = c, \hat{s} > 0 \\ A\hat{x} = 0, \hat{x} > 0. \end{cases}$$

N.B : Par abus de langage, on dit aussi que la paire (\hat{x}, \hat{s}) est centre analytique approximatif.

2.3.1 Algorithme dual

Soit donné $y \in \Omega$. Par *iv*) du théorème (2.1), on remarque qu'une fois que y satisfait $\delta(s) \leq \eta < 1$, c'est à dire une fois que y est un η -centre de Ω , la méthode de Newton duale génère une séquence $\{y^k\}$ qui converge vers y^a de manière quadratique.

Dans cette section, nous allons nous intéresser à la manière dont un centre analytique approximatif, ou η -centre de Ω , peut être généré. Nous envisageons pour cela deux hypothèses. Dans le premier cas, nous supposons la connaissance d'un point intérieur de Ω . Et dans le second, les éléments de $\text{int}(\Omega)$ seront inconnus.

Générer un centre analytique approximatif à partir d'un point intérieur de Ω

Soit $y \in \text{int}(\Omega)$.

Appliquons la méthode de Newton avec recherche linéaire. Ceci revient à résoudre le problème linéaire avec contrainte suivant

$$\begin{cases} \max \nabla \mathcal{B}(y)^T d \\ \text{s.c. } \|d\|_{-\nabla^2 \mathcal{B}(y)} \leq \alpha \quad (\alpha < 1), \end{cases} \quad (2.48)$$

qui admet pour solution la direction de Newton amortie

$$\Delta y = -\frac{\alpha}{\sqrt{-\nabla \mathcal{B}(y)^T (\nabla^2 \mathcal{B}(y))^{-1} \nabla \mathcal{B}(y)}} (\nabla^2 \mathcal{B}(y))^{-1} \nabla \mathcal{B}(y). \quad (2.49)$$

En effet, il est clair que Δy vérifie la contrainte de (2.48).

Soit ensuite d une direction admissible non nulle. Montrons que $\nabla \mathcal{B}(y)^T \Delta y \geq \nabla \mathcal{B}(y)^T d$. Si on définit $r := \sqrt{-\nabla \mathcal{B}(y)^T (\nabla^2 \mathcal{B}(y))^{-1} \nabla \mathcal{B}(y)}$, nous avons

$$\nabla \mathcal{B}(y)^T \Delta y = -\frac{\alpha}{r} \nabla \mathcal{B}(y)^T (\nabla^2 \mathcal{B}(y))^{-1} \nabla \mathcal{B}(y) = \alpha \frac{r^2}{r} = \alpha r.$$

Comme $\|d\|_{-\nabla^2 \mathcal{B}(y)} \leq \alpha$ et $r = \|\nabla \mathcal{B}(y)\|_{-[\nabla^2 \mathcal{B}(y)]^{-1}}$, nous en déduisons

$$\begin{aligned} \nabla \mathcal{B}(y)^T \Delta y &\geq \|\nabla \mathcal{B}(y)\|_{-[\nabla^2 \mathcal{B}(y)]^{-1}} \|d\|_{-[\nabla^2 \mathcal{B}(y)]} \\ &\geq \nabla \mathcal{B}(y)^T d \end{aligned}$$

où la dernière inégalité provient d'une propriété liant produit scalaire et norme, à savoir

$$|\langle a, b \rangle| \leq \|a\|_{M^{-1}} \|b\|_M \quad \text{où } M \text{ est inversible et symétrique.}$$

Cette propriété est bien applicable dans notre cas puisque $\nabla^2 \mathcal{B}(y) = -AS^{-2}A^T = -(AS^{-1})(AS^{-1})^T$ est défini négatif et symétrique.

Le problème (2.48) admet donc comme solution

$$\Delta y = -\alpha \frac{(AS^{-2}A^T)^{-1}AS^{-1}e}{\sqrt{e^T S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}e}},$$

c'est à dire la direction de Newton amortie (2.49).

Par (2.3), Δy s'écrit encore

$$\Delta y = -\alpha \frac{(AS^{-2}A^T)^{-1}AS^{-1}e}{\|p(s)\|}.$$

Quant à Δs , il s'écrit

$$\begin{aligned} \Delta s = -A^T \Delta y &= \alpha \frac{A^T(AS^{-2}A^T)^{-1}AS^{-1}e}{\|p(s)\|} \\ &= -\alpha \frac{Sp(s)}{\|p(s)\|}. \end{aligned} \tag{2.50}$$

Notons que Δy est bien admissible puisque

$$s^+ = c - A^T y^+ = c - A^T(y + \Delta y) = s - A^T \Delta y = S(e - S^{-1}A^T \Delta y)$$

est positif. En effet, $s > 0$ et $e - S^{-1}A^T \Delta y > 0$, étant donné la contrainte du problème (2.48), i.e. $\|S^{-1}A^T \Delta y\|^2 \leq \alpha^2 < 1$. Cela assure bien que $\|S^{-1}A^T \Delta y\|_\infty \leq \alpha < 1$.

Par (2.50), nous avons

$$e^T S^{-1} \Delta s = -\alpha \frac{e^T p(s)}{\|p(s)\|} = \alpha \|p(s)\| \tag{2.51}$$

car $-e^T p(s) = \|p(s)\|^2$.

En effet, par définition de $p(s)$, il faut montrer que $-e^T(Sx(s) - e) = (Sx(s) - e)^T(Sx(s) - e)$ ou encore que $(Sx(s))^T(Sx(s) - e) = 0$.

Or, $Sx(s) - e = p(s) = -S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}e$, et donc le produit $(Sx(s))^T(Sx(s) - e)$

est bien nul puisque $Ax(s) = 0$ (cfr 2.4).

Par conséquent, $s^+ = s + \Delta s = s - \alpha \frac{Sp(s)}{\|p(s)\|} = S(e - \alpha \frac{p(s)}{\|p(s)\|})$, et

$$\|S^{-1}s^+ - e\| = \alpha < 1. \quad (2.52)$$

Nous pouvons dès lors appliquer le lemme (2.2.1) pour obtenir

$$\begin{aligned} \mathcal{B}(y^+) - \mathcal{B}(y) &= \sum_{j=1}^n \log s_j^+ - \sum_{j=1}^n \log s_j \\ &= \sum_{j=1}^n \log\left(\frac{s_j^+}{s_j}\right) = \sum_{j=1}^n \log(1 + s_j^{-1}(s_j^+ - s_j)) \\ &\geq e^T S^{-1}(s^+ - s) - \frac{\|S^{-1}(s^+ - s)\|^2}{2(1 - \|S^{-1}(s^+ - s)\|_\infty)} \\ &\geq \alpha \|p(s)\| - \frac{\alpha^2}{2(1 - \alpha)} \quad \text{par (2.51) et (2.52)}. \end{aligned}$$

Donc, si $\|p(s)\| \geq \alpha$, nous avons $\mathcal{B}(y^+) - \mathcal{B}(y) \geq \alpha^2 - \frac{\alpha^2}{2(1-\alpha)}$. Dès lors, tant que $\alpha < 1/2$, $\mathcal{B}(y^+) - \mathcal{B}(y) \geq \delta$, où $\delta = \alpha^2 - \frac{\alpha^2}{2(1-\alpha)} > 0$ est constant.

En d'autres mots, la fonction potentielle duale augmente d'au moins une constante à chaque itération. Notons également que $\mathcal{B}(y)$ est bornée supérieurement par $\mathcal{B}(y^a)$ puisque le centre analytique y^a de Ω est par définition le point qui maximise le potentiel dual sur Ω . Par conséquent, il faut qu'en un nombre fini de pas, on atteigne $\delta(s) = \|p(s)\| \leq \alpha$, sinon $\mathcal{B}(y)$ augmenterait indéfiniment. Ceci implique que nous allons atteindre un centre analytique approximatif (et donc la région de convergence quadratique) en un nombre fini de pas. Le nombre total d'itérations pour atteindre cette région est borné par $\mathcal{O}(\mathcal{B}(y^a) - \mathcal{B}(y))$, puisqu'il faut au plus $\frac{\mathcal{B}(y^a) - \mathcal{B}(y)}{\delta}$ itérations.

Générer un centre analytique approximatif lorsqu'on ne connaît pas de point intérieur à Ω .

Considérons l'ensemble $\Omega(\hat{c}) = \{y \in \mathcal{R}^m : A^T y \leq \hat{c}\}$.

On a donc $\Omega(c) = \Omega$, dont l'intérieur est non vide et borné par hypothèse. Dès lors, si $\hat{c} \geq c$, $\text{int}(\Omega(\hat{c}))$ est non vide et borné, puisque $\text{int}(\Omega(\hat{c})) = \text{int}(\Omega(c)) \cup \{y : c \leq A^T y < \hat{c}\}$, où \hat{c} est fini. Remarquons encore que $y = 0 \in \text{int}(\Omega(\hat{c}))$ si $\hat{c} > 0$.

Choisissons c^0 tel que $c^0 \geq e$ et $c^0 \geq c$.

Par les résultats de la section précédente, nous pouvons générer un centre approximatif de $\Omega^0 := \Omega(c^0)$ en $\mathcal{O}(\mathcal{B}(\Omega^0) - \mathcal{B}(0, \Omega^0))$ itérations puisque $y = 0 \in \text{int}(\Omega^0)$.

Posons y^0 un centre approximatif de Ω^0 .

A partir de ce point y^0 , nous allons générer une séquence de $\{c^k\}$ et $\{y^k\}$, où $c \leq c^{k+1} \leq c^k$ et où y^k est un centre approximatif de $\Omega^k := \Omega(c^k)$. Nous vérifierons que $\mathcal{B}(\Omega) \leq \mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) - \delta$, où δ est une constante positive.

Nous poursuivrons ce processus jusqu'à ce que $c^{k+1} = c$. Dans ce cas, y^{k+1} sera centre approximatif de $\Omega = \Omega(c)$.

Décrivons tout d'abord un algorithme conceptuel, utilisant des centres analytiques exacts.

Algorithme 2.3.1

BEGIN

Soit (y^0, s^0) le centre analytique exact de $\Omega^0 = \Omega(c^0)$, et β une constante dans $(0, 1)$.

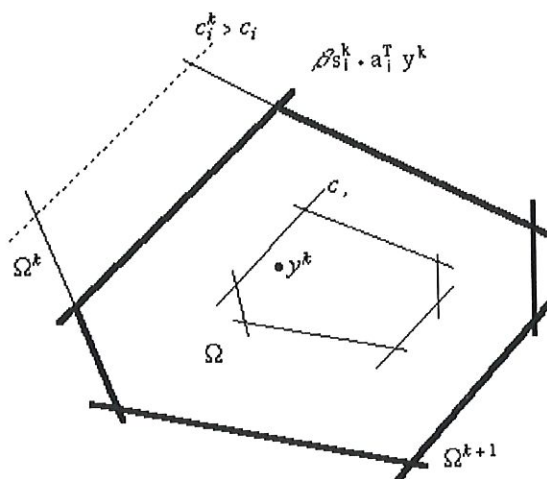
Poser $k := 0$.

WHILE $c^k \neq c$ **DO**

1. Translater une contrainte

Trouver i tq $c_i^k > c_i$ et mettre à jour

$$\begin{aligned} c_i^{k+1} &= \max\{c_i, \beta(c_i^k - a_i^T y^k) + a_i^T y^k\} \\ &= \max\{c_i, \beta s_i^k + a_i^T y^k\} \\ c_j^{k+1} &= c_j^k \quad \text{pour } j \neq i. \end{aligned}$$



NB : Le fait de prendre le maximum permet de ne pas translater la coupe jusqu'à l'intérieur de Ω .

2. Mise à jour du centre

Calculer le centre analytique y^{k+1} de Ω^{k+1} en utilisant la méthode de Newton pure à partir de y^k , qui est déjà centre approximatif de Ω^{k+1} .

Démontrer que y^k est effectivement un centre approximatif de Ω^{k+1} peut se faire par l'intermédiaire du lemme suivant

Lemme 2.3.1

$\exists \bar{x} > 0$ tq $A\bar{x} = 0$ et $\eta(\bar{x}, \bar{s}) = \|\bar{X}\bar{s} - e\| < 1$, où $\bar{s} = c^{k+1} - A^T y^k \in \mathcal{R}^n$.

En effet, grâce au théorème (2.1 i)), cela prouvera que $\delta(\bar{s}) \leq \eta(\bar{x}, \bar{s}) < 1$, et donc, par définition, y^k sera centre approximatif de Ω^{k+1} .

La preuve de ce lemme suivra de celle du lemme (2.3.3), que nous rencontrerons plus loin, et où il faudra poser $\eta = 0$.

3. Poser $k := k + 1$ et retourner au pas 1

ENDDO

END.

Lemme 2.3.2

La fonction potentielle duale $\mathcal{B}(\Omega^k) := \mathcal{B}(y^k, \Omega^k), k \geq 0$ vérifie

$$\mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) - (1 - \beta) \quad \text{si } c_i^{k+1} > c_i \quad (2.53)$$

$$\mathcal{B}(\Omega^{k+1}) < \mathcal{B}(\Omega^k) \quad \text{si } c_i^{k+1} = c_i \quad (2.54)$$

preuve :

- Si $c_i^{k+1} > c_i$, alors $c_i^{k+1} = \beta(c_i^k - a_i^T y^k) + a_i^T y^k$, et le nouveau polytope est

$$\Omega^{k+1} = \{y \in \mathcal{R}^m : c_j^k - a_j^T y \geq 0 \quad j \neq i, \beta s_i^k + a_i^T y^k - a_i^T y \geq 0\}.$$

Considérons alors y^{k+1} le centre analytique de Ω^{k+1} . Comme y^k est le centre analytique de Ω^k , $\exists x^k > 0$ tel que $X^k s^k = X^k(c - A^T y^k) = e$ et $Ax^k = 0$.

Comme $c_j^{k+1} = c_j^k$ pour $j \neq i$, $c_i^{k+1} = \beta s_i^k + a_i^T y^k$ et $s^{k+1} = c^{k+1} - A^T y^{k+1}$, nous avons

$$e^T X^k s^{k+1} = e^T X^k (c^{k+1} - A^T y^{k+1}) = e^T X^k c^{k+1} = e^T X^k c^k + e^T X^k (c^{k+1} - c^k)$$

où

$$(c^{k+1} - c^k)_j = \begin{cases} 0 & \text{si } j \neq i \\ \beta s_i^k + a_i^T y^k - c_i^k & \text{si } j = i. \end{cases}$$

Dès lors,

$$e^T X^k s^{k+1} = e^T X^k c^k + \beta x_i^k s_i^k - x_i^k (c_i^k - a_i^T y^k) = e^T X^k c^k - (1 - \beta) x_i^k s_i^k = n - 1 + \beta,$$

car $x^k s^k = e$ et $(x^k)^T c^k = n$.

Et donc,

$$\begin{aligned} \frac{\exp \mathcal{B}(\Omega^{k+1})}{\exp \mathcal{B}(\Omega^k)} &= \prod_{j=1}^n \frac{s_j^{k+1}}{s_j^k} = \prod_{j=1}^n s_j^{k+1} x_j^k \\ &\leq \left(\frac{1}{n} \sum_{j=1}^n s_j^{k+1} x_j^k \right)^n \\ &= \left(\frac{n-1+\beta}{n} \right)^n = \left(1 + \frac{\beta-1}{n} \right)^n \\ &\leq \exp(\beta-1) \end{aligned}$$

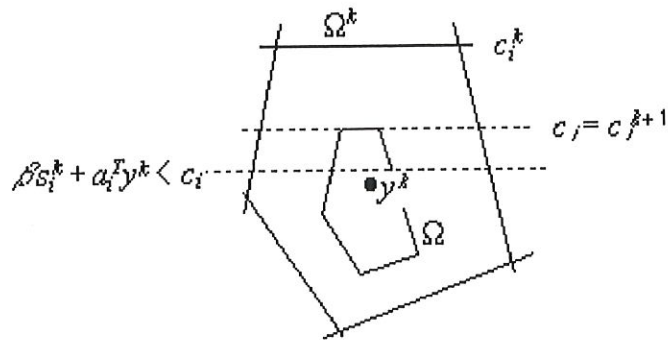
où la première inégalité vient de la relation entre moyenne géométrique et moyenne arithmétique, à savoir

$$\left(\prod x_i \right)^{1/n} \leq \frac{1}{n} \left(\sum x_i \right) \quad \forall x \in \mathcal{R}_+^n. \quad (2.55)$$

Ainsi, $\mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) \leq \beta - 1$, et (2.53) est vérifié.

- Si $c_i^{k+1} = c_i$, le nouveau polytope est

$$\Omega^{k+1} = \{y \in \mathcal{R}^m : c_j^k - a_j^T y \geq 0 \quad j \neq i, c_i - a_i^T y \geq 0\}.$$



Nous pouvons exprimer la nouvelle coupe sous la forme

$$\beta'(c_i^k - a_i^T y^k) + a_i^T y^k - a_i^T y \geq 0, \quad \beta' \in (0, 1)$$

avec $\beta' > \beta$ car quand $\beta = 0$, la coupe passe par y^k , et quand $\beta = 1$, il n'y a pas de translation.

Ne connaissant pas la valeur exacte de β' , on se contente de dire que $\mathcal{B}(\Omega^{k+1}) < \mathcal{B}(\Omega^k)$, tout en sachant que la diminution est plus grande que celle occasionnée dans le cas où $c_i^{k+1} > c_i$ puisqu'on se rapproche de y^k .

Notons encore que ce dernier cas ne peut survenir que n fois puisque Ω est constitué de n contraintes, et que, à chaque fois que l'on rencontre ce cas, on retire un candidat dans la recherche d'un i tel que $c_i^k > c_i$.

CQFD.

L'algorithme (2.3.1) s'arrêtera donc toujours après au plus $\mathcal{O}(\mathcal{B}(\Omega^0) - \mathcal{B}(\Omega)) + n$ itérations. Le premier terme correspond à la relation (2.53), puisque dans ce cas, le potentiel diminue au moins d'une constante $(1 - \beta)$ à chaque itération, et l'on s'arrête lorsque $\Omega^{k+1} = \Omega$. Le second terme 'n' provient de la relation (2.54), vérifiée lorsque $c_i^{k+1} = c_i$, égalité ne pouvant survenir que n fois. La somme de ces deux termes constitue donc une borne supérieure pour le nombre d'itérations de l'algorithme conceptuel.

Numériquement, il n'est pas possible de calculer le centre analytique exact. Nous utiliserons donc de préférence un algorithme utilisant des centres analytiques approximatifs à chaque itération.

Algorithme 2.3.2

BEGIN

Soit (y^0, s^0) un centre analytique approximatif de $\Omega^0 = \Omega(c^0)$, avec $\delta(s^0) \leq \eta < 1$, et soit $\beta \in (0, 1)$ tels que $\eta + (1 - \beta)(1 + \eta) < 1$.

Poser $k := 0$.

WHILE $c^k \neq c$ **DO**

1. Translation d'une contrainte.

Trouver i tel que $c_i^k > c_i$ et mettre à jour

$$\begin{aligned} c_i^{k+1} &= \max\{c_i, \beta s_i^k + a_i^T y^k\} \\ c_j^{k+1} &= c_j^k \quad \text{pour } j \neq i. \end{aligned}$$

2. Mettre à jour le centre analytique approximatif.

Calculer y^{k+1} , η -centre de Ω^{k+1} tel que $\delta(s^{k+1}) \leq \eta$, en utilisant la méthode de Newton pure à partir de y^k , dont on montre plus loin qu'il est déjà centre analytique approximatif de Ω^{k+1} .

3. Poser $k := k + 1$ et retourner au pas 1.

ENDDO

END.

Le lemme suivant montre qu'effectivement, y^k est déjà centre approximatif de Ω^{k+1} .

Lemme 2.3.3

$\exists \bar{x} > 0$ tel que $A\bar{x} = 0$ et $\|\bar{X}\bar{s} - e\| \leq \eta + (1 - \beta)(1 + \eta) < 1$, où $\bar{s} = c^{k+1} - A^T y^k$.

preuve :

Soit $\bar{x} = x(s^k) > 0$ avec $s^k = c^k - a^T y^k$. Alors, $A\bar{x} = Ax(s^k) = 0$, car $(x(s^k), y^k)$ est une paire admissible.

De plus, $\|\bar{x}s^k - e\| = \|x(s^k)s^k - e\| = \delta(s^k) \leq \eta$ puisque y^k est η -centre de Ω^k par hypothèse.

Notons encore que

$$\begin{cases} \bar{s}_j = s_j^k & \text{pour } j \neq i \\ \bar{s}_i = c_i^{k+1} - a_i^T y^k & \text{où } c_i^{k+1} = \max\{c_i, \beta s_i^k + a_i^T y^k\}. \end{cases}$$

Dès lors, $\bar{s}_i \geq \beta s_i^k$, et

$$\begin{aligned} \|\bar{X}\bar{s} - e\| &= \|\bar{X}s^k - e + \bar{X}\bar{s} - \bar{X}s^k\| \\ &\leq \|\bar{X}s^k - e\| + \|\bar{X}(s^k - \bar{s})\| \\ &\leq \|\bar{X}s^k - e\| + |\bar{x}_i s_i^k (1 - \beta)| \\ &\leq \eta + (1 + \eta)(1 - \beta) < 1. \end{aligned}$$

CQFD.

Le lemme précédent montre que, après translation d'une inégalité, y^k est déjà dans la région de convergence quadratique de Ω^{k+1} puisque nous choisissons $\eta + (1 + \eta)(1 - \beta) < 1$. Donc, un centre analytique approximatif plus approché encore pour Ω^{k+1} , i.e. y^{k+1} tel que $\delta(s^{k+1}) \leq \eta$ pour Ω^{k+1} , peut être obtenu à partir de y^k en un nombre constant de pas de Newton, puisque pour chaque pas, nous avons $\delta(s^+) \leq \delta(s)^2$.

Vérifions maintenant que la fonction potentielle duale est de nouveau réduite d'une constante après chaque translation, bien que l'on travaille avec des centres analytiques approchés.

Lemme 2.3.4

Soit (y^k, s^k) un centre analytique approximatif de Ω^k , avec $\delta(s^k) \leq \eta < 1$, et Ω^{k+1} défini par l'algorithme. Alors

$$\begin{aligned} \mathcal{B}(\Omega^{k+1}) &\leq \mathcal{B}(\Omega^k) - \delta & \text{si } c_i^{k+1} > c_i, \text{ (où } \delta > 0). \\ \mathcal{B}(\Omega^{k+1}) &< \mathcal{B}(\Omega^k) & \text{si } c_i^{k+1} = c_i. \end{aligned}$$

preuve :

Prouvons tout d'abord la première inégalité.

Soit (y^a, s^a) et (y_+^a, s_+^a) les centres analytiques exacts de Ω^k et Ω^{k+1} respectivement. Nous vérifions que

$$A(S^a)^{-1}e = Ax^a = 0$$

et

$$(c^k)^T (S^a)^{-1}e = (c^k - A^T y^a)^T (S^a)^{-1}e = (s^a)^T (S^a)^{-1}e = n.$$

De plus, $s_+^a = c^{k+1} - A^T y_+^a$ où d'une part $c_j^{k+1} = c_j^k$ pour $j \neq i$, et d'autre part $c_i^{k+1} = \beta(c_i^k - a_i^T y^k) + a_i^T y^k = \beta s_i^k + (c_i^k - s_i^k) = c_i^k - (1 - \beta)s_i^k$.

Notons encore que

$$e^T (S^a)^{-1} s_+^a = e^T (S^a)^{-1} (c^{k+1} - A^T y_+^a) = e^T (S^a)^{-1} c^{k+1} = e^T (S^a)^{-1} c^k + e^T (S^a)^{-1} (c^{k+1} - c^k),$$

où

$$(c^{k+1} - c^k)_j = \begin{cases} 0 & \text{si } j \neq i \\ -(1 - \beta)s_i^k & \text{si } j = i. \end{cases}$$

Donc

$$\begin{aligned} e^T(S^a)^{-1}s_+^a &= e^T(S^a)^{-1}c^k - (1 - \beta)\frac{s_i^k}{s_i^a} \\ &\leq n - (1 - \beta)(1 - \eta), \end{aligned}$$

puisque nous savons, par le point v) du théorème (2.1), que $\|(S^k)^{-1}s^a - e\| \leq \frac{\eta}{1 - \eta}$.

Dès lors,

$$\begin{aligned} \frac{\exp \mathcal{B}(\Omega^{k+1})}{\exp \mathcal{B}(\Omega^k)} &= \prod_{j=1}^n \left(\frac{s_+^a}{s^a}\right)_j \\ &\leq \left(\frac{1}{n} \sum_{j=1}^n \left(\frac{s_+^a}{s^a}\right)_j\right)^n \\ &\leq \left(\frac{n - (1 - \beta)(1 - \eta)}{n}\right)^n \\ &\leq \exp(-(1 - \beta)(1 - \eta)), \end{aligned} \tag{2.56}$$

où l'on utilise la relation entre moyenne géométrique et arithmétique (cfr 2.55) et la relation $(1 + \frac{x}{n})^n \leq e^x$, qui est toujours vérifiée quand $x \geq 0$.

Puisque (2.56) signifie aussi que $\mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) - (1 - \beta)(1 - \eta)$, la première inégalité sera satisfaite avec $\delta = (1 - \beta)(1 - \eta) < 1$.

La deuxième inégalité se prouve de la même manière que (2.54). En effet, on ne peut trouver une relation précise entre $\mathcal{B}(\Omega^{k+1})$ et $\mathcal{B}(\Omega^k)$, mais l'on sait que le potentiel va diminuer.

De plus, comme précédemment, ce cas ne peut survenir que n fois.

CQFD.

Grâce à ce lemme, on peut conclure que le nombre d'itérations nécessaires à l'algorithme (2.3.2) pour générer un centre analytique approximatif de Ω^{k+1} est en $\mathcal{O}(\mathcal{B}(\Omega^0) - \mathcal{B}(\Omega)) + n$, de la même manière que nous l'avons conclu pour l'algorithme (2.3.1), utilisant des centres analytiques exacts.

2.3.2 Algorithme primal

Commençons par démontrer une propriété vérifiée par la fonction potentielle primale.

Lemme 2.3.5

Soit $x \in \text{int}(\mathcal{X}_\Omega) = \{x \in \mathcal{R}^n : Ax = 0, x > 0\}$ et $d_x \in \mathcal{R}^n$ admissible.
Posons $x^+ = x + d_x$ et $\|X^{-1}d_x\| < 1$. Alors

$$\mathcal{P}(x^+) - \mathcal{P}(x) \leq \nabla \mathcal{P}(x)^T d_x + \frac{\|X^{-1}d_x\|^2}{2(1 - \|X^{-1}d_x\|_\infty)}. \quad (2.57)$$

preuve :

Rappelons que $\mathcal{P}(x, \Omega) = n \log(c^T x) - \sum_{j=1}^n \log x_j$.

Or, par la concavité de la fonction logarithme, nous avons

$$n \log(c^T x^+) - n \log(c^T x) = n \log\left(\frac{c^T x^+}{c^T x}\right) = n \log\left(1 + \frac{1}{c^T x}(c^T x^+ - c^T x)\right) \leq \frac{n}{c^T x} c^T d_x.$$

Et par le lemme (2.2.1),

$$\begin{aligned} -\sum_{j=1}^n \log x_j^+ + \sum_{j=1}^n \log x_j &= -\sum_{j=1}^n \log\left(\frac{x_j^+}{x_j}\right) = -\sum_{j=1}^n \log(1 + x_j^{-1}(x_j^+ - x_j)) \\ &\leq -e^T X^{-1}d_x + \frac{\|X^{-1}d_x\|^2}{2(1 - \|X^{-1}d_x\|_\infty)}. \end{aligned}$$

Dès lors,

$$\mathcal{P}(x^+) - \mathcal{P}(x) \leq \frac{n}{c^T x} c^T d_x - e^T X^{-1}d_x + \frac{\|X^{-1}d_x\|^2}{2(1 - \|X^{-1}d_x\|_\infty)} = \nabla \mathcal{P}(x)^T d_x + \frac{\|X^{-1}d_x\|^2}{2(1 - \|X^{-1}d_x\|_\infty)}.$$

CQFD.

Algorithme de Karmarkar.

Le principe de l'algorithme de Karmarkar est de réduire la fonction potentielle primale à partir d'un point intérieur de \mathcal{X}_Ω , et donc de s'approcher de x^a , la partie primale de la paire centrale (x^a, y^a) .

Comme la fonction potentielle primale est homogène, on peut normaliser x tel que $e^T x = n$ et travailler dans la région

$$\mathcal{K}_p = \{x \in \mathcal{R}^n : Ax = 0, e^T x = n, x \geq 0\}.$$

Le problème devient

$$\begin{cases} \min \mathcal{P}(x, \Omega) \\ \text{s.c. } x \in \mathcal{K}_p. \end{cases}$$

Remarquons que si $Ae = 0$, e est alors centre analytique de \mathcal{K}_p , puisqu'il devient solution du problème

$$\begin{aligned} & \max \sum_{j=1}^n \log x_j \\ & \text{s.c. } \begin{cases} Ax = 0 \\ e^T x = n \\ x \geq 0. \end{cases} \end{aligned} \quad (2.58)$$

En effet, il suffit de voir que e est déjà solution du problème obtenu à partir du problème (2.58) en oubliant la contrainte $Ax = 0$. Si l'on vérifie cela, e sera aussi solution du problème (2.58), puisque $Ae = 0$ et que les points vérifiant les trois contraintes de (2.58) sont aussi des points vérifiant les deux contraintes du problème simplifié.

Pour obtenir la solution du problème simplifié convexe, nous résolvons les conditions de Kuhn-Tucker associées. Or,

$$\mathcal{L}(x; \lambda, \mu) = -\sum_{j=1}^n \log x_j + \lambda(e^T x - n) + \mu^T x$$

où λ et μ sont les multiplicateurs de Lagrange, et

$$\frac{\partial \mathcal{L}}{\partial x_j} = -\frac{1}{x_j} + \lambda + \mu_j \quad \forall j.$$

Les conditions de Kuhn-Tucker sont donc

$$\begin{cases} -\frac{1}{x} + \lambda e + \mu = 0 \\ \mu_i x_i = 0 \quad \forall i \in \underline{n} \\ e^T x = n \\ \mu \geq 0 \\ x \geq 0. \end{cases}$$

Or, $x > 0$ car sinon, la troisième condition n'est pas satisfaite. Par la condition de complémentarité, $\mu = 0$.

La première condition assure alors que $x_i = \frac{1}{\lambda} \quad \forall i \in \underline{n}$. Comme $\sum_{j=1}^n x_j = n$, on obtient finalement que $\lambda = 1$ et donc $x_j = 1 \quad \forall j$, c'est à dire $x = e$.

Comme par hypothèse $Ae = 0$, e est bien solution du problème (2.58).

Cette condition selon laquelle $Ae = 0$ est cependant rarement vérifiée.

Par contre, si l'on choisit $x^k \in \text{int}(\mathcal{K}_p)$, e sera toujours centre analytique de

$$\mathcal{K}'_p = \{x' : A^k x' = 0, \quad e^T x' = n, \quad x' \geq 0\}$$

où $A^k = AX^k$, car $A^k e = AX^k e = Ax^k = 0$.

La transformation de Karmarkar, définie par

$$T(x) = \frac{n(X^k)^{-1}x}{e^T(X^k)^{-1}x} \quad \text{pour } x \in \mathcal{K}_p$$

permet de transformer un point x de \mathcal{K}_p en un point x' de \mathcal{K}'_p , et en particulier, le x^k choisi plus haut en centre analytique e de \mathcal{K}'_p .

En effet, $A^k T(x) = \frac{nAX^k(X^k)^{-1}x}{e^T(X^k)^{-1}x} = 0$ car $Ax = 0$, $e^T T(x) = \frac{ne^T(X^k)^{-1}x}{e^T(X^k)^{-1}x} = n$, et $T(x) \geq 0$. De plus $T(x^k) = e$.

Revenons maintenant au problème qui nous intéresse, c'est à dire la réduction de la fonction potentielle $\mathcal{P}(x, \Omega)$, notée $\mathcal{P}(x)$. Nous devons pour cela générer une suite $\{x^k\}$ telle que la différence $\mathcal{P}(x^{k+1}) - \mathcal{P}(x^k)$ entre deux itérés successifs diminue à chaque itération. Notre démarche consistera cependant à travailler avec la fonction potentielle

$$\mathcal{P}'(x') = n \log((c^k)^T x') - \sum_{j=1}^n \log x'_j$$

définie en $x' \in \mathcal{K}'_p$ et telle que $c^k = X^k c$, car nous vérifions facilement que la différence entre les valeurs de la fonction potentielle en deux points x^1 et x^2 de $\text{int}(\mathcal{K}_p)$ est invariante sous la transformation de Karmarkar, i.e.

$$\mathcal{P}'(T(x^2)) - \mathcal{P}'(T(x^1)) = \mathcal{P}(x^2) - \mathcal{P}(x^1). \quad (2.59)$$

Pour réduire la fonction potentielle $\mathcal{P}'(x')$, nous passons par la réduction de sa linéarisation autour de e , donnée par $\mathcal{P}'(e) + \nabla \mathcal{P}'(e)^T(x' - e)$. Cela revient à rechercher une direction de descente partant du point e .

Or,

$$\begin{aligned} \nabla \mathcal{P}'(e)^T x' &= \left(\frac{n}{(c^k)^T e} c^k - e \right)^T x' \\ &= \left(\frac{n}{c^T x^k} c^k - e \right)^T x' \\ &= \frac{n}{c^T x^k} (c^k)^T x' - n. \end{aligned}$$

Comme $c^T x^k$ et n sont fixés, nous résolvons simplement le problème sans contraintes

$$\begin{cases} \min (c^k)^T x' \\ \text{s.c.} \begin{cases} A^k x' = 0, \quad e^T x' = n \\ \|x' - e\| \leq \alpha \quad (\alpha > 0 \text{ fini}) \end{cases} \end{cases} \quad (2.60)$$

où la contrainte d'inégalité permet d'éviter à la direction de descente d'avoir une norme infinie.

Corollaire 2.3.1

La solution du problème (2.60) est telle que

$$x' - e = -\alpha \frac{p^k}{\|p^k\|} \quad (2.61)$$

$$\begin{aligned} \text{où } p^k &= \left(I - \begin{pmatrix} A^k \\ e \end{pmatrix}^T \begin{pmatrix} A^k(A^k)^T & 0 \\ 0 & n \end{pmatrix}^{-1} \begin{pmatrix} A^k \\ e \end{pmatrix} \right) c^k \\ &= c^k - (A^k)^T y^k - \lambda^k e \\ \text{et } \lambda^k &= \frac{c^k{}^T e}{n} = \frac{c^T x^k}{n}. \end{aligned}$$

preuve :

Pour résoudre le problème (2.60), nous passons par l'intermédiaire d'un problème semblable, à partir duquel nous effectuerons des analogies afin de se ramener au problème (2.60).

Soit le problème

$$\begin{cases} \min c^T x \\ \text{s.c. } \begin{cases} Ax = 0 \\ x^T x \leq 1 \end{cases} \quad (\text{i.e. } \|x\|^2 \leq 1). \end{cases} \quad (2.62)$$

Ce problème étant convexe, sa solution peut être obtenue par les équations de Kuhn-Tucker associées. Or

$$\mathcal{L}(x; y, z) = c^T x - y^T Ax + z(x^T x - 1)$$

où y et z sont les multiplicateurs de Lagrange.

Les conditions de Kuhn-Tucker sont donc

$$\begin{cases} c - A^T y + 2zx = 0 \\ Ax = 0 \\ x^T x \leq 1 \\ z(x^T x - 1) = 0, \quad z \geq 0. \end{cases}$$

La première équation peut encore s'écrire $x = -\frac{1}{2z}(c - A^T y)$ $z \neq 0$.

La seconde équation peut elle s'exprimer $Ac - AA^T y = 0$.

Si l'on réécrit la troisième équation en fonction de la première, on obtient alors

$$z \geq \frac{\|c - A^T y\|}{2}.$$

Comme x n'est défini que si $z \neq 0$, la quatrième équation impose que $x^T x = 1$, et $z = \frac{\|c - A^T y\|}{2}$.

On obtient donc la solution

$$\begin{cases} x^* = -\frac{c - A^T y}{\|c - A^T y\|} & \text{si } c - A^T y \neq 0 \\ \text{tq } Ac - AA^T y = 0. \end{cases}$$

Essayons maintenant d'établir une analogie entre les problèmes (2.60) et (2.62).
Le problème (2.60) peut encore s'exprimer

$$\begin{cases} \min (c^k)^T(x' - e) \\ \text{s.c.} \begin{cases} A^k(x' - e) + A^k e = 0 \\ e^T(x' - e) = 0 \\ (x' - e)^T(x' - e) \leq \alpha^2. \end{cases} \end{cases}$$

où $A^k e$ est facultatif car $A^k e = AX^k e = Ax^k = 0$.

Dès lors, le problème (2.60) prend une forme équivalente à celle du problème (2.62), à savoir

$$\begin{cases} \min (c^k)^T(x' - e) \\ \text{s.c.} \begin{cases} \begin{pmatrix} A^k \\ e^T \end{pmatrix} (x' - e) = 0 \\ (x' - e)^T(x' - e) \leq \alpha^2. \end{cases} \end{cases}$$

On identifie donc respectivement x , A , c , 1 du problème (2.62) à $x' - e$, $((A^k)^T \ e)^T$, c^k , α^2 du problème (2.60), et la solution de (2.60) est par conséquent telle que $x' - e = -\alpha \frac{p^k}{\|p^k\|}$, où

$$\begin{aligned} p^k &= c^k - ((A^k)^T \ e) \begin{pmatrix} y^k \\ \lambda^k \end{pmatrix} \\ &= c^k - (A^k)^T y^k - \lambda^k e \\ \text{avec } \begin{pmatrix} A^k \\ e^T \end{pmatrix} c^k - \begin{pmatrix} A^k \\ e^T \end{pmatrix} ((A^k)^T \ e) \begin{pmatrix} y^k \\ \lambda^k \end{pmatrix} &= 0. \end{aligned} \quad (2.63)$$

La deuxième équation de (2.63) donne la valeur de λ . En effet, $e^T c^k - n\lambda^k = 0$ signifie que $\lambda^k = \frac{(c^k)^T e}{n} = \frac{e^T x^k}{n}$. Et donc, $p^k = X^k(c - A^T y^k) - (\frac{e^T x^k}{n})e$.

CQFD.

Théorème 2.2

Soit $x' \in \mathcal{K}'_p$, la solution du problème (2.60) vérifiant (2.61). Alors

$$\mathcal{P}'(x') - \mathcal{P}'(e) \leq -\alpha \frac{n}{c^T x^k} \|p^k\| + \frac{\alpha^2}{2(1-\alpha)}, \quad (2.64)$$

où α est tel que $\|x' - e\| \leq \alpha$.

preuve :

Grâce au lemme (2.3.5), appliqué à \mathcal{P}' , nous avons que

$$\begin{aligned} \mathcal{P}'(x') - \mathcal{P}'(e) &\leq \nabla \mathcal{P}'(e)^T(x' - e) + \frac{\alpha^2}{2(1-\alpha)} \\ &= \left(\frac{n}{(c^k)^T e} c^k - e \right)^T \left(-\alpha \frac{p^k}{\|p^k\|} \right) + \frac{\alpha^2}{2(1-\alpha)} \end{aligned} \quad (2.65)$$

$$\begin{aligned}
&= -\alpha \frac{n}{c^T x^k} \frac{(c^k)^T p^k}{\|p^k\|} + \alpha \frac{e^T p^k}{\|p^k\|} + \frac{\alpha^2}{2(1-\alpha)} \\
&= -\alpha \frac{n}{c^T x^k} \frac{1}{\|p^k\|} [(c^k)^T p^k - (\frac{c^T x^k}{n}) e^T p^k] + \frac{\alpha^2}{2(1-\alpha)} \\
&= -\alpha \frac{n}{c^T x^k} \frac{1}{\|p^k\|} [(X^k c)^T p^k - (\frac{c^T x^k}{n}) e^T p^k] + \frac{\alpha^2}{2(1-\alpha)}. \quad (2.66)
\end{aligned}$$

Or,

$$(X^k c)^T p^k - (\frac{c^T x^k}{n}) e^T p^k = [X^k c - X^k A^T y^k - (\frac{c^T x^k}{n}) e]^T p^k = (p^k)^T p^k \quad (2.67)$$

car

$$\begin{aligned}
-(X^k A^T y^k)^T p^k &= -(X^k A^T y^k)^T [X^k c - X^k A^T y^k - (\frac{c^T x^k}{n}) e] \\
&= -(y^k)^T A (X^k)^2 c + (y^k)^T A (X^k)^2 A^T y^k + (\frac{c^T x^k}{n}) (y^k)^T A X^k e \\
&= 0
\end{aligned}$$

puisque d'une part, $A X^k e = A x^k = 0$, et d'autre part, la première équation de (2.63) assure que $A (X^k)^2 c - A (X^k)^2 A^T y^k = 0$. Dès lors, $-(y^k)^T [A (X^k)^2 c - A (X^k)^2 A^T y^k] = 0$. Par (2.66) et (2.67), on obtient le résultat souhaité.

CQFD.

Le théorème précédent montre que tant que $\|p^k\| \geq \eta \frac{c^T x^k}{n} > 0$, on peut choisir une valeur appropriée de α dans le problème (2.60), de telle sorte que

$$\mathcal{P}'(x') - \mathcal{P}'(e) \leq -\zeta$$

pour un constante positive $\zeta = \alpha \eta - \frac{\alpha^2}{2(1-\alpha)}$.

Notons que si $\alpha = 1/2$ et $\eta = 3/4$, alors $\zeta = 1/8$.

La relation(2.59) permet de traduire facilement cette dernière inégalité en termes de variables $x \in \mathcal{K}_p$, puisque si l'on pose $x^{k+1} = T^{-1}(x')$,

$$\mathcal{P}(x^{k+1}) - \mathcal{P}(x^k) \leq -\zeta$$

car $T^{-1}(e) = \frac{n X^k e}{e^T X^k e} = x^k$.

Notons encore que $\|p^k\|$ ne va pas rester indéfiniment supérieur à $\eta \frac{c^T x^k}{n}$. Le contraire signifierait en effet que $\mathcal{P}(x)$ tend vers $-\infty$, puisqu'à chaque itération $\mathcal{P}(x)$ diminue au moins d'une valeur ζ positive. Ceci est absurde puisque la valeur optimale de $\mathcal{P}(x)$ est atteinte au centre analytique de $\mathcal{X}_\Omega = \{x : Ax = 0, c^T x = n, x \geq 0\}$, et est finie.

Ainsi, en $\mathcal{O}(\mathcal{P}(x^0) - \mathcal{P}(x^a))$ itérations, nous générons une paire (x^k, y^k) telle que

$$\|p^k\| = \|X^k(c - A^T y^k) - (\frac{c^T x^k}{n})e\| \leq \eta \frac{c^T x^k}{n}$$

ou

$$\|\frac{n}{c^T x^k} X^k(c - A^T y^k) - e\| \leq \eta < 1.$$

Ceci traduit bien le fait que y^k est un centre analytique approximatif pour Ω , puisque le lemme (2.3.3) est vérifié, avec $\bar{x} = \frac{n}{c^T x^k} x^k$ et $\bar{s} = c - A^T y^k$.

Par la relation *iv*) du théorème (2.1), un centre analytique plus approché pour Ω peut être obtenu par les méthodes de Newton primale ou duale.

Algorithme de réduction de potentiel avec transformation affine

Dans cette section, nous montrons que la transformation affine $x' = (X^k)^{-1}x$ permet d'obtenir la même réduction de potentiel que l'algorithme de Karmarkar.

Prenons $x^k > 0$ tel que $Ax^k = 0$.

Le gradient de la fonction potentielle primale en x^k est $\nabla \mathcal{P}(x^k) = \frac{n}{c^T x^k} c - (X^k)^{-1}e$.

La réduction de $\mathcal{P}(x)$ peut se faire par l'intermédiaire de celle de sa linéarisation autour de x^k . On doit donc résoudre le problème

$$\begin{cases} \min \nabla \mathcal{P}(x^k)^T x \\ \text{s.c.} \begin{cases} Ax = 0 \\ \|(X^k)^{-1}(x - x^k)\| \leq \alpha \end{cases} \end{cases} \quad (2.68)$$

où la normalisation $\|(X^k)^{-1}(x - x^k)\| \leq \alpha$ est choisie de telle sorte que l'on puisse se ramener à un problème du même type que (2.62).

En effet, si l'on définit $x' = (X^k)^{-1}x$, le problème (2.68) devient

$$\begin{cases} \min (\frac{n}{c^T x^k} X^k c - e)^T x' \\ \text{s.c.} \begin{cases} A^k x' = 0 \\ \|x' - e\| \leq \alpha \end{cases} \end{cases}$$

ou encore

$$\begin{cases} \min (\frac{n}{c^T x^k} X^k c - e)^T (x' - e) \\ \text{s.c.} \begin{cases} A^k (x' - e) = 0 \\ \|x' - e\| \leq \alpha \end{cases} \end{cases} \quad (2.69)$$

puisque $A^k e = AX^k e = Ax^k = 0$.

Par analogie avec le problème (2.62), la solution x' de ce problème est telle que

$$x' - e = -\alpha \frac{p^k}{\|p^k\|} \quad (2.70)$$

où

$$\begin{aligned} p^k &= (\frac{n}{c^T x^k} X^k c - e) - (A^k)^T y^k \\ &= X^k (\frac{n}{c^T x^k} c - A^T y^k) - e \end{aligned}$$

avec

$$\begin{aligned}
& A^k \left(\frac{n}{c^T x^k} X^k c - e \right) - A^k (A^k)^T y^k = 0 \\
i.e. \quad & A(X^k)^2 \left(\frac{n}{c^T x^k} c - (X^k)^{-1} e \right) - A(X^k)^2 A^T y^k = 0 \quad (2.71) \\
i.e. \quad & y^k = (A(X^k)^2 A^T)^{-1} A(X^k)^2 \nabla \mathcal{P}(x^k)
\end{aligned}$$

Pour revenir à notre problème de départ en x , on pose $x^{k+1} = X^k x'$, où x' est la solution du problème (2.69), de telle sorte que, par (2.70)

$$x^{k+1} - x^k = -\alpha \frac{X^k p^k}{\|p^k\|}.$$

Théorème 2.3

$$\mathcal{P}(x^{k+1}) - \mathcal{P}(x^k) \leq -\alpha \|p^k\| + \frac{\alpha^2}{2(1-\alpha)} \quad (2.72)$$

où α est défini dans le problème (2.69).

preuve :

En utilisant le lemme (2.3.5), on peut établir que

$$\begin{aligned}
\mathcal{P}(x^{k+1}) - \mathcal{P}(x^k) & \leq \nabla \mathcal{P}(x^k)^T (x^{k+1} - x^k) + \frac{\alpha^2}{2(1-\alpha)} \\
& = \left(\frac{n}{c^T x^k} c - (X^k)^{-1} e \right)^T \left(-\alpha \frac{X^k p^k}{\|p^k\|} \right) + \frac{\alpha^2}{2(1-\alpha)} \\
& = -\frac{\alpha}{\|p^k\|} \left(\frac{n}{c^T x^k} X^k c - e \right)^T p^k + \frac{\alpha^2}{2(1-\alpha)}
\end{aligned}$$

Or,

$$\left(\frac{n}{c^T x^k} X^k c - e \right)^T p^k = \left(\frac{n}{c^T x^k} X^k c - e - X^k A^T y^k \right)^T p^k = (p^k)^T p^k$$

car, grâce à (2.71), on vérifie que

$$\begin{aligned}
-(X^k A^T y^k)^T p^k & = -(X^k A^T y^k)^T \left(\frac{n}{c^T x^k} X^k c - e - X^k A^T y^k \right) \\
& = -(y^k)^T \left[\frac{n}{c^T x^k} A(X^k)^2 c - A X^k e - A(X^k)^2 A^T y^k \right] \\
& = 0.
\end{aligned}$$

Par suite, (2.72) est vérifié.

CQFD.

Dès lors, aussi longtemps que $\|p^k\| \geq \eta > 0$, on peut choisir un α approprié tel que

$$\mathcal{P}(x^{k+1}) - \mathcal{P}(x^k) \leq -\zeta$$

où $\zeta = \alpha\eta - \frac{\alpha^2}{2(1-\alpha)}$ est une constante positive.

De nouveau, si $\alpha = 1/2$ et $\eta = 3/4$, alors $\zeta = 1/8$.

Comme dans le cas de l'algorithme de Karmarkar, $\|p^k\|$ ne peut pas rester indéfiniment supérieur à η , sinon $\mathcal{P}(x)$ tendrait vers $-\infty$. Donc, en $\mathcal{O}(\mathcal{P}(x^0) - \mathcal{P}(x^a))$ itérations, nous générons une paire (x^k, y^k) telle que, si x^k est normalisé de telle sorte que $c^T x^k = n$, on ait

$$\|p^k\| = \|X^k(c - A^T y^k) - e\| \leq \eta < 1,$$

ce qui signifie que y^k est centre analytique approximatif de Ω .

2.3.3 Algorithme primal-dual

Principe

Supposons que nous ayons $x \in \text{int}(\mathcal{X}_\Omega)$ et $s \in \text{int}(\mathcal{S}_\Omega)$ tels que $\eta(x, s) \leq \|Xs - e\| \leq \eta < 1$. Pour générer un nouvel itéré x^+ et (y^+, s^+) , la méthode de Newton primale-duale se base sur le raisonnement suivant.

Notre but étant de trouver le centre analytique de $\Omega = \{y \in \mathcal{R}^m : A^T y \leq c\}$, nous devons résoudre le système de Kuhn-Tucker (1.2). Une façon d'approcher la solution est d'effectuer des itérations de Newton appliquées au système

$$F(x, y, s) = \begin{pmatrix} A^T y + s - c \\ Ax \\ Xs - e \end{pmatrix} = 0$$

La direction de Newton primale-duale en (x, y, s) sera donc définie par le système

$$J(F(x, y, s))d = -F(x, y, s)$$

où J désigne le jacobien de F .

Ce système se réécrit

$$\begin{pmatrix} O & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} d_x \\ d_y \\ d_s \end{pmatrix} = - \begin{pmatrix} 0 \\ 0 \\ xs - e \end{pmatrix}$$

ou encore

$$\begin{cases} A^T d_y + d_s = 0 \\ Ad_x = 0 \\ Sd_x + Xd_s = e - Xs \end{cases} \quad (2.73)$$

On posera alors

$$x^+ = x + d_x, \quad y^+ = y + d_y, \quad s^+ = s + d_s.$$

Etablissons un lemme sur la centralité de (x^+, y^+, s^+) .

Lemme 2.3.6

Soit (x^+, y^+, s^+) , généré par le système (2.73). Alors (x^+, s^+) est admissible intérieure et

$$\|X^+ s^+ - e\| \leq \frac{\sqrt{2}\eta^2}{4(1-\eta)}.$$

preuve :

Remarquons que $\|X^+ s^+ - e\| = \|D_x d_s\|$ puisque, par (2.73), $X^+ s^+ = (x + d_x)(s + d_s) = xs + xd_s + sd_x + d_x d_s = e + d_x d_s$.

Si l'on multiplie chaque terme de la troisième équation de (2.73) par $(XS)^{-1/2}$, on obtient

$$Dd_x + D^{-1}d_s = r := (XS)^{-1/2}(e - Xs),$$

où $D = S^{1/2} X^{-1/2}$.

Si l'on pose $p = Dd_x$ et $q = D^{-1}d_s$, on remarque que $p^T q = d_x^T d_s = 0$ car $d_x \in \text{Ker} A$ et $d_s \in \text{Im} A^T$.

De plus,

$$\begin{aligned} \|D_x d_s\|^2 &= \|Pq\|^2 = \sum_{j=1}^n (p_j q_j)^2 \\ &= \sum_{p_j q_j < 0} (p_j q_j)^2 + \sum_{p_j q_j > 0} (p_j q_j)^2 \\ &\leq \left(\sum_{p_j q_j < 0} p_j q_j \right)^2 + \left(\sum_{p_j q_j > 0} p_j q_j \right)^2 \end{aligned}$$

Or, $\sum_{j=1}^n p_j q_j = \sum_{p_j q_j > 0} p_j q_j + \sum_{p_j q_j < 0} p_j q_j = 0$ car $p^T q = 0$.

D'où $\sum_{p_j q_j > 0} p_j q_j = -\sum_{p_j q_j < 0} p_j q_j$.

Comme ces deux sommes sont positives, nous avons aussi $(\sum_{p_j q_j > 0} p_j q_j)^2 = (\sum_{p_j q_j < 0} p_j q_j)^2$.

Dès lors,

$$\begin{aligned} \|D_x d_s\|^2 &\leq 2 \left(\sum_{p_j q_j > 0} p_j q_j \right)^2 \\ &\leq 2 \left(\sum_{p_j q_j > 0} \frac{(p_j + q_j)^2}{4} \right)^2 \\ &\leq 2 \left(\frac{\|r\|^2}{4} \right)^2 \quad \text{où } \|r\|^2 = \sum_{j=1}^n (p_j + q_j)^2. \end{aligned}$$

La deuxième inégalité provient de la relation entre moyenne géométrique et arithmétique (2.55).

Par conséquent, comme $\|Xs - e\| \leq \eta$, et que donc $\|r\|^2 \leq \|(xs)^{-1/2}\|_\infty^2 \|e - xs\|^2 \leq \frac{\eta^2}{1-\eta}$, nous obtenons le résultat attendu, à savoir

$$\|X^+ s^+ - e\| = \|D_x d_s\| \leq \frac{\sqrt{2}}{4} \frac{\eta^2}{1-\eta}.$$

Il reste à montrer que (x^+, s^+) est admissible intérieure.

L'admissibilité est immédiate compte tenu des conditions (2.73). Pour prouver le caractère intérieur, prenons $0 \leq \alpha \leq 1$, et définissons $x(\alpha) = x + \alpha d_x$ et $s(\alpha) = s + \alpha d_s$.

Par définition de d_s et d_x , on a

$$\begin{aligned} x(\alpha)s(\alpha) &= xs + \alpha(xd_s + sd_x) + \alpha^2 d_x d_s \\ &= xs + \alpha(e - xs) + \alpha^2 d_x d_s \\ &= xs(1 - \alpha) + \alpha(e + \alpha d_x d_s). \end{aligned}$$

Si $\alpha = 1$, alors $x(1)s(1) = e + d_x d_s$. Or, $\|d_x d_s\|_\infty \leq \|d_x d_s\| = \|e - x^+ s^+\| < 1$ (par la première partie de la preuve). Donc, chaque composante de $d_x d_s \in]-1, 1[$, et $x_i(1)s_i(1) > 0 \forall i$.

Dès lors, comme $x = x(0) > 0$ et $s = s(0) > 0$, on a aussi $x_i(\alpha)s_i(\alpha) > 0 \forall i$ et $\forall \alpha \in [0, 1]$. Cela n'assure pas pour autant que $x_i(\alpha) > 0$ et $s_i(\alpha) > 0 \forall i$ et $\forall \alpha \in [0, 1]$. Cependant, si l'on suppose qu'il existe un indice i et un α tels que $x_i(\alpha) < 0$, alors $x_i(\alpha) = x_i(0) + \alpha(d_x)_i < 0$. Par conséquent, $\exists \bar{\alpha} \in]0, \alpha]$ tel que $x_i(\bar{\alpha}) = 0$. Ceci est impossible puisque l'on a vu que $x_i(\alpha)s_i(\alpha) > 0 \forall i$.

On conclut donc que $x_i(\alpha) > 0, s_i(\alpha) > 0 \forall i, \forall \alpha \in [0, 1]$.

CQFD.

N.B. : Remarquons que nous obtiendrions la convergence quadratique, i.e. $\|X^+ s^+ - e\|^2 \leq \eta^2$, si $\eta \leq 1 - \frac{1}{2\sqrt{2}}$.

Algorithme utilisant le potentiel primal-dual.

Nous allons construire un algorithme primal-dual permettant de calculer le centre analytique de Ω , qui convient pour toute paire de départ (x, s) intérieure telle que $s^T x = (c - A^T y)^T x = c^T x = n$.

Il a pour but de minimiser la fonction potentielle primale-duale (2.13) sur $\text{int}(\mathcal{X}_\Omega \times \mathcal{S}_\Omega)$, et donc de se rapprocher de la paire centrale (x^a, s^a) . Nous nous arrêterons lorsque nous identifierons une paire centrale approximative telle que y^k est un η -centre de Ω (où $\eta \leq 1 - \frac{1}{2\sqrt{2}}$). Partant d'une telle paire, le lemme précédent nous assure que la méthode de Newton primale-duale converge quadratiquement vers la paire centrale.

Lemme 2.3.7

Soient les directions d_x, d_y , et d_s générées par le système (2.73). Posons

$$\theta = \frac{\alpha \sqrt{\min(Xs)}}{\|(XS)^{-1/2}(e - Xs)\|} \quad (2.74)$$

où $\alpha \in]0, 1]$ et $\min(v \in \mathcal{R}^n) := \min_{j \in \underline{n}}(v_j)$, alors nous avons

$$\psi_n(x + \theta d_x, s + \theta d_s) - \psi_n(x, s) \leq -\alpha \sqrt{\min(Xs)} \|(XS)^{-1/2}(e - Xs)\| + \frac{\alpha^2}{2(1 - \alpha)}. \quad (2.75)$$

preuve :

Remarquons tout d'abord que pour le choix de θ en (2.74), nous avons

$$\|\theta S^{-1} d_s\|^2 + \|\theta X^{-1} d_x\|^2 \leq \alpha^2. \quad (2.76)$$

En effet, par (2.73), nous avons $Sd_x + Xd_s = e - Xs$, et donc

$$X^{-1}d_x + S^{-1}d_s = X^{-1}S^{-1}(e - Xs).$$

Si nous définissons, comme dans la preuve du lemme (2.3.6), $r := (XS)^{-1/2}(e - Xs)$, nous vérifions également que

$$\theta X^{-1}d_x + \theta S^{-1}d_s = \theta(XS)^{-1/2}r. \quad (2.77)$$

Ainsi,

$$\|\theta X^{-1}d_x + \theta S^{-1}d_s\|^2 = \|\theta X^{-1}d_x\|^2 + \|\theta S^{-1}d_s\|^2 + 2\theta^2 X^{-1}S^{-1}d_x^T d_s \quad (2.78)$$

où $d_x^T d_s = 0$ puisque $d_x \in \text{Ker}A$, et $d_s \in \text{Im}A^T$.

De plus,

$$\|\theta X^{-1}d_x + \theta S^{-1}d_s\|^2 \leq \theta^2 \|(XS)^{-1/2}\|^2 \|r\|^2 \quad \text{par (2.77),} \quad (2.79)$$

où $\theta^2 \|r\|^2 = \alpha^2 \min(Xs)$ grâce à (2.74), et où $\|(XS)^{-1/2}\|^2 = \|(XS)^{-1}\| = \|XS\|^{-1}$ puisque XS est une matrice diagonale.

Ainsi, par (2.78) et (2.79), nous obtenons

$$\begin{aligned} \|\theta X^{-1}d_x\|^2 + \|\theta S^{-1}d_s\|^2 &\leq \alpha^2 \min(Xs) \|XS\|^{-1} \\ &\leq \alpha^2 \|XS\| \|XS\|^{-1} \\ &= \alpha^2. \end{aligned}$$

Remarquons ensuite que (2.76) implique

$$\tau := \max(\|\theta S^{-1}d_s\|_\infty, \|\theta X^{-1}d_x\|_\infty) \leq \alpha < 1 \quad (2.80)$$

puisque

$$\begin{aligned} \|\theta S^{-1}d_s\|_\infty^2 &\leq \|\theta S^{-1}d_s\|_2^2 \leq \|\theta S^{-1}d_s\|_2^2 + \|\theta X^{-1}d_x\|_2^2 \leq \alpha^2 < 1 \\ \|\theta X^{-1}d_x\|_\infty^2 &\leq \|\theta X^{-1}d_x\|_2^2 \leq \|\theta S^{-1}d_s\|_2^2 + \|\theta X^{-1}d_x\|_2^2 \leq \alpha^2 < 1. \end{aligned}$$

Cela entraîne

$$\begin{aligned} x^+ &= x + \theta d_x = x(e - \theta x^{-1}d_x) > 0 \\ s^+ &= s + \theta d_s = s(e - \theta s^{-1}d_s) > 0. \end{aligned}$$

Dès lors, grâce au lemme (2.2.1) et à (2.76), nous obtenons

$$\psi_n(x^+, s^+) - \psi_n(x, s) = n \log(x^+)^T s^+ - \sum_{i=1}^n \log x_i^+ s_i^+ - n \log x^T s + \sum_{i=1}^n \log x_i s_i,$$

Et d'autre part,

$$\begin{aligned}
-\sum_{i=1}^n \log x_i^+ s_i^+ + \sum_{i=1}^n \log x_i s_i &= -\sum_{i=1}^n \log\left(\frac{x_i^+ s_i^+}{x_i s_i}\right) \\
&= -\sum_{i=1}^n \log\left(\frac{x_i^+}{x_i}\right) - \sum_{i=1}^n \log\left(\frac{s_i^+}{s_i}\right) \\
&= -\sum_{i=1}^n \log(1 + (x_i)^{-1}(x_i^+ - x_i)) - \sum_{i=1}^n \log(1 + (s_i)^{-1}(s_i^+ - s_i)) \\
&= -\sum_{i=1}^n \log(1 + (x_i)^{-1}\theta d_x) - \sum_{i=1}^n \log(1 + (s_i)^{-1}\theta d_s) \\
&\leq -\theta e^T X^{-1} d_x + \frac{\|\theta X^{-1} d_x\|^2}{2(1 - \|\theta X^{-1} d_x\|_\infty)} - \theta e^T S^{-1} d_s \\
&\quad + \frac{\|\theta S^{-1} d_s\|^2}{2(1 - \|\theta S^{-1} d_s\|_\infty)} \\
&\leq -\theta e^T (S^{-1} d_s + X^{-1} d_x) + \frac{\|\theta S^{-1} d_s\|^2 + \|\theta X^{-1} d_x\|^2}{2(1 - \tau)}
\end{aligned}$$

où $\tau := \max(\|\theta S^{-1} d_s\|_\infty, \|\theta X^{-1} d_x\|_\infty) < 1$. Ainsi,

$$\psi_n(x^+, s^+) - \psi_n(x, s) \leq \theta e^T (X d_s + S d_x) - \theta e^T (S^{-1} d_s + X^{-1} d_x) + \frac{\|\theta S^{-1} d_s\|^2 + \|\theta X^{-1} d_x\|^2}{2(1 - \tau)}. \quad (2.81)$$

Par (2.80) et (2.76), nous savons que le dernier terme de (2.81) vérifie

$$\frac{\|\theta S^{-1} d_s\|^2 + \|\theta X^{-1} d_x\|^2}{2(1 - \tau)} \leq \frac{\alpha^2}{2(1 - \alpha)}. \quad (2.82)$$

Tandis que la somme des deux autres termes peut s'écrire

$$\begin{aligned}
\theta e^T (X d_s + S d_x) - \theta e^T (S^{-1} d_s + X^{-1} d_x) &= \theta(e^T (X d_s + S d_x) - e^T (S^{-1} d_s + X^{-1} d_x)) \\
&= \theta(e^T (X d_s + S d_x) - e^T (XS)^{-1} (X d_s + S d_x)) \\
&= \theta(e - (XS)^{-1} e)^T (X d_s + S d_x) \\
&= \theta(e - (XS)^{-1} e)^T (e - Xs) \quad \text{par (2.73)} \\
&= -\theta(e - Xs)^T (XS)^{-1} (e - Xs) \\
&= -\theta \|(XS)^{-1/2} (e - Xs)\|^2 \\
&= -\alpha \sqrt{\min(Xs)} \|(XS)^{-1/2} (e - Xs)\| \quad (2.83)
\end{aligned}$$

Par (2.82) et (2.83), nous obtenons finalement

$$\psi_n(x^+, s^+) - \psi_n(x, s) \leq -\alpha \sqrt{\min(Xs)} \|(XS)^{-1/2} (e - Xs)\| + \frac{\alpha^2}{2(1 - \alpha)}.$$

Théorème 2.4

Soient $x^+ = x + \theta d_x$ et $s^+ = s + \theta d_s$ où θ est donné par (2.74). Alors, si $\eta(x, s) \geq \eta$ pour une constante $\eta < 1$, il est possible de trouver α tel que θ permette de vérifier

$$\psi_n(x^+, s^+) - \psi_n(x, s) \leq -\zeta$$

pour une constante positive ζ .

preuve :

Considérons $\alpha\sqrt{\min(Xs)}\|(XS)^{-1/2}(e - Xs)\| = \alpha\sqrt{\min(Xs)}\|(XS)^{-1/2}e - (XS)^{1/2}e\|$.
Posons $z = Xs > 0$ et $\min(z) = z_1 \leq 1$ puisque $e^T z = e^T(Xs) = n$.

Considérons tout d'abord le cas où $z_1 \leq 1/2$.

Dans ce cas,

$$\begin{aligned} \alpha\sqrt{\min(z)}\|Z^{-1/2}e - Z^{1/2}e\| &\geq \alpha\sqrt{\min(z)}\|Z^{-1/2}e - Z^{1/2}e\|_\infty \\ &= \alpha\sqrt{\min(z)}|(z_1)^{-1/2} - (z_1)^{1/2}| \\ &= \alpha\sqrt{\min(z)}(z_1)^{-1/2}|1 - z_1| \\ &= \alpha|1 - z_1| \\ &\geq \alpha/2 \quad \text{car } z_1 \leq 1/2. \end{aligned}$$

Donc, grâce au lemme (2.3.7), nous pouvons choisir α tel que

$$\psi_n(x^+, s^+) - \psi(x, s) \leq -\frac{\alpha}{2} + \frac{\alpha^2}{2(1 - \alpha)} = -\zeta$$

où ζ est une constante positive.

Passons maintenant au cas où $z_1 \geq 1/2$.

Dans ce cas,

$$\alpha\sqrt{\min(z)}\|Z^{-1/2}e - Z^{1/2}e\| \geq \frac{\alpha}{\sqrt{2}}\|Z^{-1/2}e - Z^{1/2}e\|.$$

Posons alors $z_n = \max(z) \geq 1$ (puisque $e^T z = n$).

Si $z_n \geq 2$, alors

$$\begin{aligned} \frac{\alpha}{\sqrt{2}}\|Z^{-1/2}e - Z^{1/2}e\| &\geq \frac{\alpha}{\sqrt{2}}\|Z^{-1/2}e - Z^{1/2}e\|_\infty \\ &\geq \frac{\alpha}{\sqrt{2}}|(z_n)^{-1/2} - (z_n)^{1/2}| \\ &= \frac{\alpha}{\sqrt{2}}(z_n)^{1/2}|1 - (z_n)^{-1}| \\ &\geq \frac{\alpha}{2} \end{aligned}$$

car $(z_n)^{1/2} \geq \sqrt{2}$ et $|1 - (z_n)^{-1}| \geq 1/2$.

De nouveau, nous pouvons choisir α tel que

$$\psi_n(x^+, s^+) - \psi(x, s) \leq -\frac{\alpha}{2} + \frac{\alpha^2}{2(1-\alpha)} = -\zeta$$

pour ζ une constante positive.

Il reste à analyser le cas où $z_1 \geq 1/2$ et $z_n \leq 2$.

Dans ce cas,

$$\alpha \sqrt{\min(z)} \|Z^{-1/2}e - Z^{1/2}e\| = \alpha \sqrt{\min(z)} \|Z^{-1/2}(e - z)\| \geq \frac{\alpha}{2} \|e - z\|$$

car d'une part, $\sqrt{\min(z)} \geq 1/\sqrt{2}$, et d'autre part, $\|Z^{-1/2}(e - z)\| = \|e - z\|_{Z^{-1}} \geq \sqrt{\lambda_{\min}(Z^{-1})} \|e - z\| \geq \frac{1}{\sqrt{2}} \|e - z\|$, où l'on utilise l'inégalité de Rayleigh, à savoir $\lambda_{\min}(Z) \leq \frac{x^T Z x}{x^T x} \leq \lambda_{\max}(Z)$.

Donc, comme $\|e - z\| = \|e - Xs\| \geq \eta$, nous pouvons choisir α tel que

$$\psi_n(x^+, s^+) - \psi_n(x, s) \leq -\zeta$$

où $\zeta = \frac{\alpha}{2}\eta - \frac{\alpha^2}{2(1-\alpha)}$ est une constante positive.

CQFD.

Par ce théorème, nous obtenons que, jusqu'à ce que $\|e - Xs\| \leq \eta < 1$, la fonction potentielle primale-duale est réduite d'une constante dépendant de α à chaque itération.

Notons qu'il arrivera inévitablement un moment où $\|e - Xs\| \leq \eta$, car sinon $\psi_n(x, s)$ tend vers $-\infty$. Ceci est impossible puisque par définition de la fonction primale-duale (2.13), nous savons que $\psi_n(x, s) \geq n \log n$ pour toute paire admissible.

Par conséquent, en $\mathcal{O}(\psi_n(x^0, s^0) - \psi_n(x^a, s^a))$ itérations, ou de manière plus précise en $\mathcal{O}(\psi_n(x^0, s^0) - n \log n)$ itérations, on générera une paire (x, y) telle que

$$\eta(x, s) = \|Xs - e\| = \|X(c - A^T y) - e\| < \eta.$$

Ceci signifie que x et s forment une paire centrale approximative pour Ω .

Notons encore que la complexité ne dépend uniquement que de (x^0, s^0) .

Chapitre 3

Algorithme ACCPM basé sur le principe de réduction de potentiel pour un problème d'admissibilité linéaire.

3.1 Introduction

Tandis que le chapitre précédent avait pour but de développer quelques méthodes pour le calcul du centre analytique d'un polytope borné, les chapitres qui suivent traitent de la méthode ACCPM proprement dite, dont le principe a fait l'objet du premier chapitre. Dans ce chapitre plus précisément, nous étudierons un algorithme ACCPM particulier, basé sur le principe de réduction de potentiel, et destiné à résoudre le *problème d'admissibilité linéaire*.

Ce problème consiste à trouver un point admissible dans la région

$$\mathcal{C} = \{y \in \mathcal{R}^m : c - A^T y \geq 0 \text{ et } 0 \leq y \leq e\}$$

convexe, où les composantes A et c sont rationnelles, et où $\text{int}(\mathcal{C})$ est supposé non vide.

Remarquons que nous travaillons ici sous des contraintes plus fortes que celles faites au premier chapitre pour un algorithme ACCPM général, puisque notre polytope \mathcal{C} est défini ici par des contraintes linéaires, et pas simplement convexes.

3.2 Principe de l'algorithme ACCPM avec réduction de potentiel

Le principe général des algorithmes ACCPM décrit au Chapitre 1 peut ici être caractérisé de manière plus précise.

On démarre l'algorithme avec le polytope $\Omega^0 = \{y \in \mathcal{R}^m : 0 \leq y \leq e\} = [0, 1]^m$ contenant \mathcal{C} , et son centre analytique y^0 .

A chaque itération, on vérifie si $y^k \in \mathcal{C}$. Si oui, l'algorithme se termine. Sinon, on choisit une contrainte de \mathcal{C} non satisfaite par y^k .

Si cette contrainte de \mathcal{C} est parallèle à une contrainte de Ω^k , Ω^{k+1} est obtenu en translatant cette contrainte de Ω^k vers l'intérieur du polytope.

Si cette contrainte de \mathcal{C} n'est parallèle à aucune contrainte de Ω^k , Ω^{k+1} est obtenu en construisant une nouvelle contrainte parallèle à celle de \mathcal{C} et telle qu'elle intersecte Ω^k .

On recherche ensuite le centre analytique du nouveau polytope Ω^{k+1} et on passe à l'itération $k + 1$.

L'étude de la fonction potentielle duale $\mathcal{B}(\Omega^k)$ nous montrera que sa valeur diminue à chaque itération. Or, l'hypothèse selon laquelle $\text{int}(\mathcal{C}) \neq \emptyset$ impose aussi que $\mathcal{B}(\Omega^k)$ admette une borne inférieure. Dès lors, la complexité d'un tel algorithme ne peut être que finie. Nous l'étudierons de façon rigoureuse.

Dans le but d'obtenir une valeur précise de la diminution du potentiel à chaque itération, nous introduisons maintenant une propriété combinatoire des centres analytiques.

3.3 Propriété combinatoire du centre analytique

Notons y^a le centre analytique du polytope

$$\Omega = \{y \in \mathcal{R}^m : s = c - A^T y \geq 0\}, \quad (3.1)$$

où $c \in \mathcal{R}^n$ et $A \in \mathcal{R}^{m \times n}$ est de rang plein.

Nous savons grâce au premier chapitre que y^a est le point de Ω qui maximise sur Ω la fonction potentielle duale, définie par

$$\mathcal{B}(y, \Omega) = \sum_{j=1}^n \log(c_j - a_j^T y) = \sum_{j=1}^n \log s_j, \quad y \in \Omega, \quad (3.2)$$

où $s_j = c_j - a_j^T y$ est appelée variable d'écart.

On note

$$\mathcal{B}(\Omega) := \mathcal{B}(y^a, \Omega) = \sum_{j=1}^n \log s_j^a \quad (3.3)$$

où $s^a = c - A^T y^a$.

Définissons maintenant le polytope

$$\Omega_\beta^+ = \{y \in \mathcal{R}^m : c_j - a_j^T y \geq 0 \quad j = 1, \dots, n-1, \quad \beta s_n^a + a_n^T y^a - a_n^T y \geq 0\}, \quad (3.4)$$

où $\beta \in [0, 1]$.

Ce polytope est obtenu à partir de Ω en translatant la dernière contrainte de manière à ce qu'elle ne 'dépasse' pas le centre analytique.

Posons \bar{y}^a le centre analytique de Ω_β^+ . Alors, la fonction potentielle duale du nouveau polytope Ω_β^+ est

$$\mathcal{B}(\Omega_\beta^+) = \sum_{j=1}^{n-1} \log(c_j - a_j^T \bar{y}^a) + \log(\beta s_n^a + a_n^T y^a - a_n^T \bar{y}^a). \quad (3.5)$$

Cette fonction potentielle vérifie la propriété suivante.

Théorème 3.1

Nous vérifions que

$$\mathcal{B}(\Omega_\beta^+) \leq \mathcal{B}(\Omega) - (1 - \beta). \quad (3.6)$$

preuve :

Cette relation a déjà été démontrée au lemme (2.3.2) du chapitre précédent puisque Ω_β^+ s'assimile au Ω^{k+1} du lemme en question, dans le cas où $c_i^{k+1} > c_i$.

CQFD.

Définissons à présent le polytope convexe

$$\Omega_\beta^+ = \{y \in \mathcal{R}^m : c_j - a_j^T y \geq 0 \quad j = 1, \dots, n, \quad \beta \bar{r} + a_{n+1}^T y^a - a_{n+1}^T y \geq 0\} \quad (3.7)$$

où $\beta \in (0, 1)$, et $\bar{r} = \sqrt{a_{n+1}^T (A(X^a)^2 A^T)^{-1} a_{n+1}}$.

Ω_β^+ est donc obtenu à partir de Ω en ajoutant une nouvelle contrainte de pente arbitraire.

Le choix de \bar{r} n'est pas innocent. En effet, nous allons montrer qu'avec cette valeur de \bar{r} , la nouvelle coupe intersecte le polytope Ω , ce qui sera nécessaire dans l'algorithme pour être sûr que le polytope rétrécit à chaque itération, et que l'on approxime de mieux en mieux l'ensemble admissible.

Pour démontrer cela, nous avons besoin de deux lemmes.

Lemme 3.3.1

Soit (\bar{x}, \bar{s}) une paire intérieure admissible telle que $A\bar{x} = 0, \bar{x} > 0, \bar{s} = c - A^T \bar{y} > 0$, et telle que $\|e - \bar{x}\bar{s}\| \leq \theta < 1$. L'ellipsoïde

$$\{y \in \mathcal{R}^m : \|\bar{X} A^T (y - \bar{y})\| \leq 1 - \theta\} \quad (3.8)$$

est contenu dans $\{y \in \mathcal{R}^m : A^T y \leq c\}$.

preuve : [5]

Montrons tout d'abord que l'ellipsoïde

$$\{y \in \mathcal{R}^m : \|\bar{S}^{-1}A^T(y - \bar{y})\| \leq 1\} \quad (3.9)$$

est contenu dans $\{y \in \mathcal{R}^m : A^T y \leq c\}$.

En effet, prenons un point $y \in \mathcal{R}^m$ appartenant à l'ellipsoïde (3.9). Il vérifie

$$\|\bar{S}^{-1}A^T(y - \bar{y})\| \leq 1. \quad (3.10)$$

Or, $A^T y = c - s$ et $A^T \bar{y} = c - \bar{s}$. Donc (3.10) s'écrit aussi $\|\bar{S}^{-1}(\bar{s} - s)\| \leq 1$, ou

$$\left\| \frac{s}{\bar{s}} - e \right\| \leq 1 \quad (3.11)$$

où $\bar{s} > 0$ par hypothèse.

Dès lors, $s = c - A^T y \geq 0$, sinon (3.11) n'est pas vérifié puisqu'au moins une composante de $\left| \frac{s}{\bar{s}} \right|$ serait supérieure à 1.

Ainsi, $A^T y \leq c$ et $y \in \{y \in \mathcal{R}^m : A^T y \leq c\}$.

Si l'on montre maintenant que tout point de l'ellipsoïde (3.8) est tel que $\|\bar{S}^{-1}A^T(y - \bar{y})\| \leq 1$, alors ils appartiendront tous à l'ensemble $\{y \in \mathcal{R}^m : A^T y \leq c\}$ étant donné la première partie de la preuve. On aura ainsi démontré la thèse.

Or, pour tout élément y de l'ellipsoïde (3.8), on a

$$\begin{aligned} \|\bar{S}^{-1}A^T(y - \bar{y})\| &= \|(\bar{X}\bar{S})^{-1}\bar{X}A^T(y - \bar{y})\| \\ &\leq \|(\bar{X}\bar{S})^{-1}\|_{\infty} \|\bar{X}A^T(y - \bar{y})\| \\ &\leq \frac{1}{1 - \theta} \|\bar{X}A^T(y - \bar{y})\| \leq 1 \end{aligned}$$

où la deuxième inégalité vient de la condition de centralité $\|e - \bar{x}\bar{s}\| \leq \theta < 1$, qui entraîne $1 - \theta \leq \bar{x}_i \bar{s}_i \leq 1 + \theta \quad \forall i \in \underline{n}$, et donc

$$\frac{1}{1 + \theta} \leq \frac{1}{\bar{x}_i \bar{s}_i} \leq \frac{1}{1 - \theta} \quad \forall i \in \underline{n}.$$

Le lemme est ainsi démontré.

CQFD.

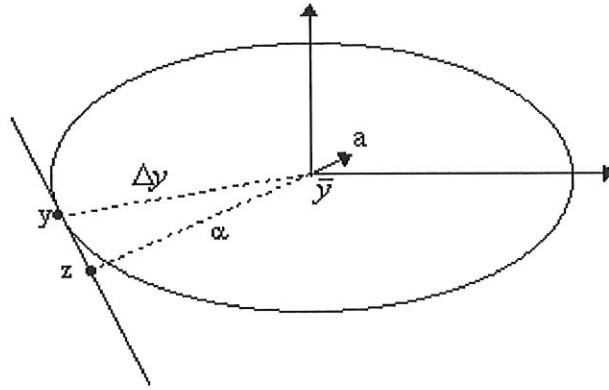
L'ellipsoïde (3.8) est appelé *ellipsoïde de Dikin*, et constitue une région de confiance autour de \bar{y} .

Lemme 3.3.2

Une coupe intersecte l'ellipsoïde de Dikin si la distance de \bar{y} à la frontière de la coupe ne dépasse pas $(1 - \theta)\bar{r}$, où $\bar{r} = \sqrt{a^T(A\bar{X}^2A^T)^{-1}a}$, et où a est le vecteur normal de la coupe ajoutée, et tel que $\|a\| = 1$ [2].

preuve :

Le problème revient à chercher la longueur α représentée en dimension deux dans la figure qui suit, puisque $\text{spdg } \|a\| = 1$.



Pour cela, nous allons d'abord évaluer $\Delta y = y - \bar{y}$, solution du problème

$$\min\{a^T \Delta y : \|\bar{X}A^T \Delta y\| \leq 1 - \theta\}, \quad (3.12)$$

et qui consiste à trouver une droite de vecteur normal ' a ' la plus éloignée possible de \bar{y} , mais ayant toujours une intersection avec l'ellipsoïde.

Pour trouver la solution du problème (3.12), effectuons tout d'abord un changement de notations, tel que $t = \Delta y = y - \bar{y}$. Pour la facilité, nous noterons également $\bar{X}A^T = D$.

Le problème devient

$$\begin{cases} \min a^T t & (a \neq 0) \\ \text{s.c. } \|Dt\|^2 \leq (1 - \theta)^2. \end{cases} \quad (3.13)$$

Pour résoudre ce problème, nous utilisons les conditions de Kuhn-Tucker. Le problème étant convexe, sa solution sera celle des conditions.

La condition de qualification de contraintes de Slater est satisfaite puisque il existe un \hat{t} tel que $\|D\hat{t}\|^2 < (1 - \theta)^2$ (l'intérieur de l'ellipsoïde n'est pas vide).

Sachant que le lagrangien du problème est

$$\mathcal{L}(t; \lambda) = a^T t + \lambda(\|Dt\|^2 - (1 - \theta)^2),$$

les conditions de Kuhn-Tucker s'écrivent

$$\begin{cases} \nabla_t \mathcal{L}(t; \lambda) = a + 2\lambda D^T D t = 0 \\ \lambda(\|Dt\|^2 - (1 - \theta)^2) = 0 \\ \lambda \geq 0 \\ \|Dt\|^2 - (1 - \theta)^2 \leq 0. \end{cases}$$

La contrainte est active, sinon $\lambda = 0$, et $a = 0$, ce qui est impossible par hypothèse. Dès lors $\|Dt\|^2 = (1 - \theta)^2$, ce qui signifie que nous nous trouvons sur la frontière.

Le système de Kuhn-Tucker devient

$$\begin{cases} a + 2\lambda D^T D t = 0 \\ \lambda \geq 0 \end{cases} \quad (3.14)$$

ou encore

$$\begin{cases} D(D^T D)^{-1} a + 2\lambda D t = 0 \\ \lambda \geq 0 \end{cases} \quad (3.15)$$

où $D^T D$ est inversible car $D = \bar{X} A^T$ est de rang plein.

La première équation de (3.15) permet donc d'obtenir

$$\|D(D^T D)^{-1} a\| = 2\lambda(1 - \theta),$$

ou

$$2\lambda = \frac{\|D(D^T D)^{-1} a\|}{(1 - \theta)}. \quad (3.16)$$

Introduite dans la première équation de (3.14), (3.16) implique que

$$a + \frac{\|D(D^T D)^{-1} a\|}{1 - \theta} D^T D t = 0.$$

Or, $D^T D$ est inversible, et donc

$$t = -\frac{(1 - \theta)}{\|D(D^T D)^{-1} a\|} (D^T D)^{-1} a.$$

Comme $D = \bar{X} A^T$ et $t = \Delta y$, on obtient la solution du problème (3.12), à savoir

$$\begin{aligned} \Delta y &= -\frac{(1 - \theta)}{\sqrt{a^T (A \bar{X}^2 A^T)^{-1} a}} (A \bar{X}^2 A^T)^{-1} a \\ &= -\frac{(1 - \theta)}{\bar{r}} (A \bar{X}^2 A^T)^{-1} a. \end{aligned}$$

Ayant trouvé la solution du problème (3.12), revenons à notre problème de départ, qui consiste à obtenir la valeur de α représentée sur la figure précédente.

Nous savons que l'égalité $-\alpha(a^T a) = a^T \Delta y$ doit être satisfaite, puisque z et y appartiennent à la même droite de vecteur normal a . Donc,

$$-\alpha \|a\|^2 = a^T \Delta y = -\frac{1 - \theta}{\bar{r}} a^T (A \bar{X}^2 A^T)^{-1} a = -(1 - \theta) \bar{r}.$$

Comme $\text{spdg } \|a\| = 1$, nous obtenons $\alpha = (1 - \theta)\bar{r}$.

CQFD.

Grâce aux deux lemmes précédents, nous pouvons maintenant justifier le choix de \bar{r} dans la définition de Ω_β^+ (cfr 3.7).

Le polytope Ω ayant y^a pour centre analytique, nous pouvons définir un ellipsoïde de Dikin centré en y^a , donné par

$$\{y \in \mathcal{R}^m : \|X^a A^T(y - y^a)\| \leq 1\} \quad (3.17)$$

où $\theta = 0$ puisque nous savons qu'au centre analytique, $x^a s^a = e$.

Le lemme (3.3.1) assure que cet ellipsoïde (3.17) est contenu dans $\Omega = \{y \in \mathcal{R}^m : A^T y \leq c\}$. Et par le lemme (3.3.2), une coupe intersectera cet ellipsoïde si la distance de y^a à la frontière de la coupe est inférieure ou égale à \bar{r} , où $\bar{r} = (a_{n+1}^T (A(X^a)^2 A^T)^{-1} a_{n+1})^{1/2}$, et où a_{n+1} est le vecteur normal de la nouvelle coupe.

Dès lors, la coupe choisie pour définir Ω_β^+ en (3.7) intersectera l'ellipsoïde de Dikin centré en y^a , et donc aussi Ω .

Passons maintenant à la relation existant entre les fonctions potentielles duales associées à Ω et Ω_β^+ défini en (3.7). Cette relation est établie par un théorème semblable au théorème (3.1).

Théorème 3.2

Nous vérifions que

$$\mathcal{B}(\Omega_\beta^+) \leq \mathcal{B}(\Omega) + \log(4\bar{r}) - (1.5 - \beta) \quad (3.18)$$

où

$$\mathcal{B}(\Omega_\beta^+) = \sum_{j=1}^n \log(c_j - a_j^T \bar{y}^a) + \log(\beta\bar{r} + a_{n+1}^T y^a - a_{n+1}^T \bar{y}^a). \quad (3.19)$$

preuve :

Puisque y^a est centre analytique de Ω , il existe $x^a > 0$ tel que $X^a s^a = X^a(c - A^T y^a) = e$ et $Ax^a = 0$, puisque y^a est solution des conditions d'optimalité (1.2).

Notons $c_{n+1} = \beta\bar{r} + a_{n+1}^T y^a$, et posons $\bar{s}^a = c - A^T \bar{y}^a > 0$, les n premiers écarts définis au nouveau centre analytique \bar{y}^a . L'écart à la nouvelle contrainte, quant à lui, vérifie

$$\begin{aligned} \bar{s}_{n+1}^a &= a_{n+1}^T (y^a - \bar{y}^a) + \beta\bar{r} \\ &= a_{n+1}^T (A(X^a)^2 A^T)^{-1} (A(X^a)^2 A^T)(y^a - \bar{y}^a) + \beta\bar{r} \\ &= a_{n+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)^2 (A^T y^a - A^T \bar{y}^a) + \beta\bar{r} \end{aligned}$$

$$\begin{aligned}
&= a_{n+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)^2 (-c + A^T y^a + c - A^T \bar{y}^a) + \beta \bar{r} \\
&= a_{n+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)^2 (\bar{s}^a - s^a) + \beta \bar{r} \\
&= a_{n+1}^T (A(X^a)^2 A^T)^{-1} A(X^a) (X^a \bar{s}^a - e) + \beta \bar{r} \\
&\leq \|a_{n+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)\| \|X^a \bar{s}^a - e\| + \beta \bar{r} \\
&= (\|X^a \bar{s}^a - e\| + \beta) \bar{r}
\end{aligned} \tag{3.20}$$

où l'unique inégalité est obtenue par Cauchy-Schwartz, et où l'égalité (3.20) se justifie de la manière suivante :

$$\begin{aligned}
\|a_{n+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)\| &= \|X^a A^T (A(X^a)^2 A^T)^{-1} a_{n+1}\| \\
&= \sqrt{a_{n+1}^T (A(X^a)^2 A^T)^{-1} A X^a X^a A^T (A(X^a)^2 A^T)^{-1} a_{n+1}} \\
&= \sqrt{a_{n+1}^T (A(X^a)^2 A^T)^{-1} a_{n+1}} = \bar{r}.
\end{aligned}$$

Nous vérifions aussi, puisque $Ax^a = 0$, que

$$\begin{aligned}
e^T X^a \bar{s}^a &= e^T X^a (c - A^T \bar{y}^a) = e^T X^a c - (AX^a e)^T \bar{y}^a \\
&= e^T X^a c = e^T X^a (c - A^T y^a) = e^T X^a s^a = e^T e \\
&= n.
\end{aligned} \tag{3.21}$$

Donc, grâce aux définitions de $\mathcal{B}(\Omega_\beta^+)$ (cfr 3.19), et $\mathcal{B}(\Omega)$ (cfr 3.3),

$$\begin{aligned}
\frac{\exp \mathcal{B}(\Omega_\beta^+)}{\bar{r} \exp \mathcal{B}(\Omega)} &= \frac{\exp(\sum_{j=1}^n \log \bar{s}_j^a + \log \bar{s}_{n+1}^a)}{\bar{r} \exp(\sum_{j=1}^n \log s_j^a)} \\
&= \frac{\bar{s}_{n+1}^a}{\bar{r}} \prod_{j=1}^n \frac{\bar{s}_j^a}{s_j^a} \\
&= \frac{\bar{s}_{n+1}^a}{\bar{r}} \prod_{j=1}^n \bar{s}_j^a x_j^a \\
&\leq (\|X^a \bar{s}^a - e\| + \beta) \prod_{j=1}^n \bar{s}_j^a x_j^a
\end{aligned} \tag{3.22}$$

où l'inégalité (3.22) est induite par (3.20).

Posons $\alpha = X^a \bar{s}^a > 0$.

Ainsi, nous pouvons encore borner le terme de droite de (3.22), tout en tenant compte de (3.21), en résolvant le problème

$$\begin{cases} \max \psi(\alpha) = (\|\alpha - e\| + \beta) \prod_{j=1}^n \alpha_j \\ \text{s.c. } e^T \alpha = n \text{ et } \alpha > 0. \end{cases} \tag{3.23}$$

La solution de (3.23) est donnée par α^* tel que $\alpha_1^* = \delta > 1$, et $\alpha_2^* = \dots = \alpha_n^* = \frac{n-\delta}{n-1}$ (voir [11] Chapitre 2).

Dès lors,

$$\begin{aligned}
\psi(\alpha) &\leq \left(\sum_{i=2}^n \left(\frac{n-\delta}{n-1} - 1 \right)^2 + (\delta - 1)^2 + \beta \right) \delta \left(\frac{n-\delta}{n-1} \right)^{n-1} \\
&= \left((\delta - 1) \sqrt{\frac{n}{n-1}} + \beta \right) \delta \left(\frac{n-\delta}{n-1} \right)^{n-1} \\
&= 4 \sqrt{\frac{n}{n-1}} \frac{\delta - 1 + \sqrt{\frac{n-1}{n}} \beta \delta}{2} \left(\frac{n-\delta}{n-1} \right)^{n-1} \\
&\leq 4 \sqrt{\frac{n}{n-1}} \left(\frac{n - 0.5 + \sqrt{\frac{n-1}{n}} \beta}{n+1} \right)^{n+1} \tag{3.24}
\end{aligned}$$

$$< \frac{4}{\exp(1.5 - \beta)} \tag{3.25}$$

où les inégalités (3.24) et (3.25) ont pu être vérifiées graphiquement pour un n donné, mais pour lesquelles la démonstration théorique formelle n'a pu être obtenue.

Par (3.22) et (3.25), nous obtenons

$$\frac{\exp \mathcal{B}(\Omega_\beta^+)}{\exp \mathcal{B}(\Omega)} \leq \frac{4\bar{r}}{\exp(1.5 - \beta)}, \tag{3.26}$$

ou

$$\mathcal{B}(\Omega_\beta^+) \leq \mathcal{B}(\Omega) + \log(4\bar{r}) - (1.5 - \beta). \tag{3.27}$$

CQFD.

3.4 Algorithme ACCPM pour la résolution du problème d'admissibilité linéaire

Comme nous l'avons déjà précisé dans l'introduction, le problème d'admissibilité linéaire consiste à trouver un point admissible de l'ensemble convexe

$$\mathcal{C} = \{y \in \mathcal{R}^m : c - A^T y \geq 0 \text{ et } 0 \leq y \leq e\},$$

où les composantes de A et c sont rationnelles, et où l'on suppose $\text{int}(\mathcal{C})$ non vide. On supposera de plus que $\text{spdg} \|a_j\| = 1$, où a_j est la $j^{\text{ème}}$ colonne de A .

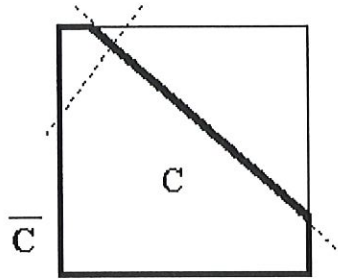
Le fait que $\text{int}(\mathcal{C})$ soit non vide peut se traduire plus spécifiquement de la manière suivante :

Pour tout polytope $\bar{\mathcal{C}} = \{y \in \mathcal{R}^m : \bar{c} - \bar{A}^T y \geq 0 \text{ et } 0 \leq y \leq e\}$, où \bar{c} est un sous-vecteur de c et \bar{A} est une sous-matrice de A , il existe un point $y \in \bar{\mathcal{C}}$ tel que

$$\bar{c} - \bar{A}^T y \geq 2^{-L} e \text{ et } 2^{-L} e \leq y \leq (1 - 2^{-L}) e \quad (3.28)$$

pour un certain $L > 1$ fixé.

Remarquons que $\mathcal{C} \subset \bar{\mathcal{C}}$, pour tout sous-système défini par \bar{c} et \bar{A} , comme le suggère la figure suivante.



Dès lors, il est clair que si \mathcal{C} a un intérieur non vide, alors pour tout \bar{c} et \bar{A} , $\bar{\mathcal{C}}$ aussi, ce que traduit (3.28).

Il en découle que , pour $y \in \bar{\mathcal{C}}$ vérifiant (3.28),

$$\mathcal{B}(\bar{\mathcal{C}}) \geq \sum_{j=1}^l \log(\bar{c}_j - \bar{a}_j^T y) + \sum_{j=1}^m \log y_j + \sum_{j=1}^m \log(1 - y_j) \quad (3.29)$$

$$\geq \sum_{j=1}^l \log(2^{-L}) + \sum_{j=1}^m \log(2^{-L}) + \sum_{j=1}^m \log(2^{-L}) \quad (3.30)$$

où (3.30) est induit par (3.28), et où l est le nombre d'inégalités de la forme $\bar{c}_j - \bar{a}_j^T y \geq 0$ définissant $\bar{\mathcal{C}}$.

Si l'on note $q = 2m + l$ le nombre d'inégalités de $\bar{\mathcal{C}}$, nous obtenons alors

$$\mathcal{B}(\bar{\mathcal{C}}) \geq q \log(2^{-L}) \geq q \log_2(2^{-L}) = -qL. \quad (3.31)$$

Décrivons à présent l'algorithme de réduction de potentiel utilisant des centres analytiques exacts.

3.4.1 Algorithme ACCPM utilisant des centres analytiques exacts

BEGIN

Initialisation

Soit $A^0 = (I \quad -I) \in \mathcal{R}^{m \times 2m}$, $c^0 = \begin{pmatrix} e \\ 0 \end{pmatrix} \in \mathcal{R}^{2m}$
 et $y^0 = 0.5e \in \mathcal{R}^m$, $s^0 = c^0 - (A^0)^T y^0 = 0.5e \in \mathcal{R}^{2m}$, et $x^0 = 2e \in \mathcal{R}^{2m}$.
 Posons également $n = 2m$.

Nous vérifions que $A^0 x^0 = 0$ et $\|X^0 s^0 - e\| = 0$.
 En d'autres termes, (x^0, y^0) forme la paire centrale de

$$\Omega^k = \{y \in \mathcal{R}^m : c^k - (A^k)^T y \geq 0\}, \quad \text{où } k = 0.$$

$k^{\text{ème}}$ itération

A la $k^{\text{ème}}$ itération, nous vérifions si le centre analytique courant y^k satisfait à toutes les inégalités de \mathcal{C} .

- a. Si oui, l'algorithme se termine, et y^k est le point admissible recherché.
- b. Sinon, il existe un indice \bar{j} pour lequel $c_{\bar{j}} - a_{\bar{j}}^T y^k < 0$.

Dans le cas b., deux alternatives sont possibles.

1. Soit $a_{\bar{j}}$ est déjà une colonne de A^k , c'est à dire que la contrainte de \mathcal{C} non vérifiée par y^k est parallèle à une des contraintes du polytope courant Ω^k .
2. Soit $a_{\bar{j}}$ n'est pas une colonne de A^k , c'est à dire que la contrainte de \mathcal{C} non vérifiée par y^k n'est parallèle à aucune des contraintes du polytope courant Ω^k .

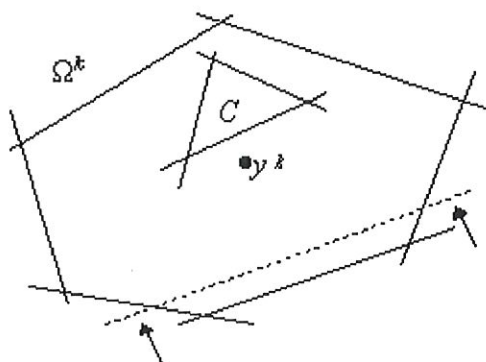
Dans le premier cas, on va construire Ω^{k+1} en remplaçant la contrainte de Ω^k parallèle à celle d'indice \bar{j} de \mathcal{C} , par une translatée. Les autres contraintes restent identiques.

Donc, nous mettons à jour

$$\begin{aligned} c_{\bar{j}}^{k+1} &= \beta(c_{\bar{j}}^k - a_{\bar{j}}^T y^k) + a_{\bar{j}}^T y^k = \beta s_{\bar{j}}^k + a_{\bar{j}}^T y^k & \beta \in (0, 1) \\ c_j^{k+1} &= c_j^k & \text{si } j \neq \bar{j} \end{aligned} \quad (3.32)$$

et

$$A^{k+1} = A^k \quad (3.33)$$



Remarquons que $\mathcal{C} \subset \Omega^{k+1}$.

Dans le cas 2., nous allons construire Ω^{k+1} en ajoutant une contrainte parallèle à la contrainte d'indice \bar{j} de \mathcal{C} .

Puisqu'il y a ajout d'une contrainte, le vecteur (la matrice) c^{k+1} (A^{k+1}) aura une composante (une colonne) de plus que c^k (A^k).

Donc, nous mettons à jour

$$c^{k+1} = \begin{pmatrix} c^k \\ \beta\bar{r} + a_{\bar{j}}^T y^k \end{pmatrix} \quad \text{et} \quad A^{k+1} = (A^k, a_{\bar{j}}) \quad (3.34)$$

où

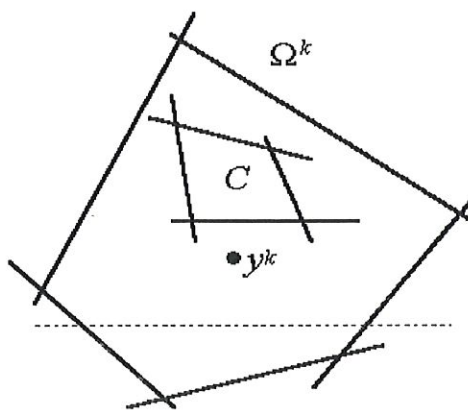
$$\bar{r} = \sqrt{a_{\bar{j}}^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}}, \quad (3.35)$$

avec x^k , la partie primale de la paire centrale (x^k, y^k) .

On pose de plus $n := n + 1$, où n représente le nombre de colonnes de A^k .

Comme nous l'avons vu, cette valeur de \bar{r} assure que la coupe ajoutée intersecte le polytope Ω^k . Remarquons de plus que $y^k \in \Omega^{k+1}$ car $c^k - (A^k)^T y^k \geq 0$ et $\beta\bar{r} + a_{\bar{j}}^T y^k - a_{\bar{j}}^T y^k = \beta\bar{r} \geq 0$.

Dès lors, la coupe n'intersectera pas \mathcal{C} , et donc $\mathcal{C} \subset \Omega^{k+1}$.



La $k^{\text{ème}}$ itération s'achève par la mise à jour du centre. Elle consiste à calculer le centre analytique y^{k+1} de Ω^{k+1} en utilisant la méthode de Newton pure à partir de y^k , et aller à

l'itération $k + 1$.

Nous développerons cette étape après avoir analysé la variation de la fonction potentielle $\mathcal{B}(\Omega^k)$ à chaque itération.

END.

Théorème 3.3

Dans le premier cas de l'algorithme, où $a_{\bar{j}}$ est déjà une colonne de A^k , la fonction potentielle duale entre deux itérés successifs vérifie

$$\mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) - (1 - \beta). \quad (3.36)$$

preuve :

La preuve de ce théorème découle de celle du théorème (3.1), qui envisage le même type de translation.

CQFD.

Dans le cas 2. de l'algorithme, où $a_{\bar{j}}$ n'est pas une colonne de A^k , nous savons par le théorème (3.2) que

$$\mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) \leq \log(4\bar{r}) - (1.5 - \beta). \quad (3.37)$$

Remarquons que le terme de droite de (3.37) dépend de l'itération k puisque \bar{r} en dépend. Ce terme peut toutefois être borné par une constante puisque l'on montre que $\bar{r} \leq \frac{1}{2\sqrt{2}}$. Pour cela, il suffit de remarquer que les deux inégalités suivantes sont satisfaites.

Propriété 3.4.1

Nous vérifions que

$$A^k(X^k)^2(A^k)^T = A^k(S^k)^{-2}(A^k)^T \succeq (Y^k)^{-2} + (I - Y^k)^{-2}. \quad (3.38)$$

preuve :

En réalité, on montre que

$$A^k(S)^{-2}(A^k)^T = (Y)^{-2} + (I - Y)^{-2} + \sum_{j=1}^{n(k)} \frac{a_j a_j^T}{(s_{2m+j})^2}, \quad (3.39)$$

où $n(k)$ représente le nombre de coupes ajoutées (et non translattées) au polytope initial Ω^0 pour obtenir Ω^k , et où $s = c^k - (A^k)^T y$ est le vecteur d'écart correspondant à un point $y \in \Omega^k$.

et (3.42) est satisfait.

Par conséquent, $A^k(S)^{-2}(A^k)^T = (Y)^{-2} + (I - Y)^{-2} + \sum_{j=1}^{n(k)} \frac{a_j a_j^T}{(s_{2m+j})^2}$. Cette égalité étant valable en tout point y de Ω^k , elle l'est aussi au centre analytique y^k , et le dernier terme étant semi défini positif, la propriété (3.4.1) est démontrée.

CQFD.

Propriété 3.4.2

Nous vérifions que

$$(Y^k)^{-2} + (I - Y^k)^{-2} \succeq 8I. \quad (3.43)$$

preuve :

En réalité, cette propriété est satisfaite en tout point de Ω^k . En effet, sachant que $y \in (0, e)$, il faut montrer que $\forall y_i \in (0, 1)$,

$$\sum_{i=1}^m \frac{1}{y_i^2} d_i^2 + \sum_{i=1}^m \frac{1}{(1 - y_i)^2} d_i^2 \geq 8 \sum_{i=1}^m d_i^2 \quad \forall d \in \mathcal{R}^m,$$

i.e.

$$\frac{(1 - y_i)^2 + y_i^2}{y_i^2(1 - y_i)^2} \geq 8 \quad \forall i \in \underline{m},$$

i.e.

$$(1 - y_i)^2 + y_i^2 \geq 8y_i^2(1 - y_i)^2 \quad \forall i \in \underline{m}.$$

Or, par la relation entre moyenne géométrique et arithmétique (cfr 2.55), nous avons que

$$\frac{y_i^2 + (1 - y_i)^2}{2} \geq [y_i^2(1 - y_i)^2]^{1/2} = y_i(1 - y_i) \quad \forall i \in \underline{m}.$$

Dès lors, il reste à montrer que

$$2y_i(1 - y_i) \geq 8y_i^2(1 - y_i)^2 \quad \text{si } 0 < y_i < 1. \quad (3.44)$$

Or, en simplifiant par $y_i(1 - y_i)$, (3.44) devient $4y_i^2 - 4y_i + 1 \geq 0$. Donc, il faut montrer que $(2y_i - 1)^2 \geq 0$ si $0 < y_i < 1 \quad \forall i$, ce qui est toujours satisfait.

CQFD.

Les propriétés (3.4.1) et (3.4.2) nous permettent de déduire que $A^k(S^k)^{-2}(A^k)^T - 8I$ est semi définie positive. Dès lors, $\bar{r}^2 = a_j^T (A^k(X^k)^2(A^k)^T)^{-1} a_j \leq \frac{1}{8} a_j^T a_j = \frac{\|a_j\|^2}{8} = \frac{1}{8}$, puisque $\text{spdg } \|a_j\| = 1$.

Ainsi, $\bar{r} \leq \frac{1}{2\sqrt{2}}$, et par (3.37), nous obtenons le théorème suivant.

Théorème 3.4

Dans le second cas de l'algorithme, où $a_{\bar{j}}$ n'est pas une colonne de A^k , la fonction potentielle duale entre deux itérés successifs vérifie

$$\mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) + \log(\sqrt{2}) - (1.5 - \beta). \quad (3.45)$$

Les théorèmes (3.3) et (3.4) nous montrent que, dans les deux cas de l'algorithme, la fonction potentielle duale diminue d'une constante pour un β approprié. Cette diminution prend par exemple une valeur de 0.25 si $\beta=0.75$ dans le premier cas, et $\beta=0.9$ dans le second cas.

Montrons à présent que le calcul de y^{k+1} , dans l'étape de mise à jour du centre, peut s'effectuer assez aisément.

En réalité, nous montrons que y^k est déjà un centre analytique approximatif de Ω^{k+1} , et est donc aussi dans la région de convergence quadratique de Ω^{k+1} . y^{k+1} peut ainsi être obtenu à partir de y^k en effectuant des pas de Newton purs, avec convergence quadratique.

Lemme 3.4.1

Dans les deux cas de l'algorithme, il existe un $\bar{x} > 0$ tel que

$$A^{k+1}\bar{x} = 0 \quad \text{et} \quad \|\bar{X}\bar{s} - e\| \leq \delta < 1,$$

où $\bar{s} = c^{k+1} - (A^{k+1})^T y^k \in \mathcal{R}^n$.

NB : \bar{s} étant le vecteur des écarts de l'ancien centre analytique aux nouvelles contraintes, ce lemme signifie que y^k est un δ -centre pour Ω^{k+1} . y^k est donc dans la région de convergence quadratique pour la méthode de Newton duale, comme le confirment les points *i*) et *iv*) du théorème (2.1), au chapitre précédent.

preuve :

Dans le premier cas, où $a_{\bar{j}}$ est déjà une colonne de A^k , la mise à jour effectuée à la $k^{\text{ème}}$ itération est celle décrite en (3.32) et (3.33). Dès lors,

$$\begin{cases} \bar{s}_j = c_j^{k+1} - (a_j^{k+1})^T y^k = c_j^k - (a_j^k)^T y^k = s_j^k \text{ pour } j \neq \bar{j} \\ \bar{s}_{\bar{j}} = c_{\bar{j}}^{k+1} - (a_{\bar{j}}^{k+1})^T y^k = \beta s_{\bar{j}}^k + a_{\bar{j}}^T y^k - a_{\bar{j}}^T y^k = \beta s_{\bar{j}}^k. \end{cases}$$

Posons $\bar{x} = x^k > 0$. Alors $A^{k+1}\bar{x} = A^k x^k = 0$, puisque x^k est la partie primale de la paire centrale (x^k, y^k) associée à Ω^k .

Et pour $\beta \in (0, 1)$,

$$\begin{aligned} \|\bar{X}\bar{s} - e\|^2 &= \|(x_1^k s_1^k - 1, \dots, x_{\bar{j}}^k s_{\bar{j}}^k \beta - 1, \dots, x_n^k s_n^k - 1)\|^2 \\ &= (x_{\bar{j}}^k s_{\bar{j}}^k \beta - 1)^2 \\ &= (\beta - 1)^2 \end{aligned}$$

car $x_i^k s_i^k = 1 \forall i$ puisque (x^k, s^k) est la paire centrale de Ω^k .
Dès lors, $\|\bar{X}\bar{s} - e\| = (1 - \beta) < 1$.

Dans le second cas, où $a_{\bar{j}}$ n'est pas une colonne de A^k , la mise à jour à la $k^{\text{ème}}$ itération est celle décrite par (3.34). Dès lors,

$$\begin{cases} \bar{s}_j = c_j^{k+1} - (a_j^{k+1})^T y^k = c_j^k - (a_j^k)^T y^k = s_j^k & j = 1, \dots, n \\ \bar{s}_{n+1} = c_{n+1}^{k+1} - (a_{n+1}^{k+1})^T y^k = \beta \bar{r} + a_{\bar{j}}^T y^k - a_{\bar{j}}^T y^k = \beta \bar{r}, \end{cases} \quad (3.46)$$

puisque $A^{k+1} = (A^k, a_{\bar{j}})$.

Posons $\alpha = \frac{\beta}{1+\beta^2} < 1$ et $\Delta x = -\frac{\alpha}{\bar{r}} X^k (A^k)^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}$. Alors

$$\|\Delta x\| = \frac{\alpha}{\bar{r}} \bar{r} = \alpha < 1 \quad (3.47)$$

puisque $\bar{r} = \sqrt{a_{\bar{j}}^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}}$ (cfr (3.35)).

Définissons $\bar{x} = \begin{pmatrix} X^k (e + \Delta x) \\ \alpha / \bar{r} \end{pmatrix} > 0$.

Alors

$$\begin{aligned} A^{k+1} \bar{x} &= A^k X^k e + A^k X^k \Delta x + \frac{\alpha}{\bar{r}} a_{\bar{j}} \\ &= A^k X^k \Delta x + \frac{\alpha}{\bar{r}} a_{\bar{j}} \quad \text{car } A^k x^k = 0. \end{aligned} \quad (3.48)$$

Or

$$\begin{aligned} A^k X^k \Delta x &= -\frac{\alpha}{\bar{r}} [A^k (X^k)^2 (A^k)^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}] \\ &= -\frac{\alpha}{\bar{r}} a_{\bar{j}}. \end{aligned} \quad (3.49)$$

Dès lors, par (3.48), $A^{k+1} \bar{x} = 0$. De plus,

$$\begin{aligned} \|\bar{X}\bar{s} - e\|^2 &= \|X^k (e + \Delta x) s^k - e\|^2 + \left(\frac{\alpha}{\bar{r}} \beta \bar{r} - 1\right)^2 \\ &= \|X^k s^k - e + X^k s^k \Delta x\|^2 + (1 - \alpha\beta)^2 \\ &= \|\Delta x\|^2 + (1 - \alpha\beta)^2 \quad \text{car } X^k s^k = e. \\ &= \alpha^2 + (1 - \alpha\beta)^2 \quad \text{par (3.47)}. \end{aligned}$$

Or, $\alpha = \frac{\beta}{1+\beta^2}$, donc

$$\begin{aligned} \|\bar{X}\bar{s} - e\|^2 &= \left(\frac{\beta}{1+\beta^2}\right)^2 + \left(1 - \frac{\beta^2}{1+\beta^2}\right)^2 \\ &= \frac{1}{1+\beta^2} < 1. \end{aligned}$$

Nous avons donc démontré le lemme dans les deux cas de l'algorithme.

CQFD.

La théorie sur la méthode de Newton nous montre que, partant de la paire (\bar{x}, \bar{s}) dans la région de convergence quadratique, la nouvelle paire centrale (x^{k+1}, y^{k+1}) associée à Ω^{k+1} peut être générée en $\mathcal{O}(\ln L)$ pas de Newton purs (primaux-duaux) (voir, par exemple, [7]).

Nous obtenons à présent un théorème de convergence pour l'algorithme utilisant des centres analytiques exacts.

Théorème 3.5

L'algorithme ACCPM utilisant des centres analytiques exacts génère un point admissible de \mathcal{C} en $\mathcal{O}(qL)$ itérations et $\mathcal{O}(qL \ln L)$ pas de Newton, où q est le nombre d'inégalités dans le dernier polytope Ω^k , et où L est défini en (3.28).

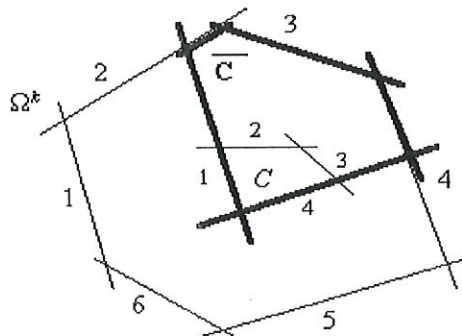
preuve :

Notons tout d'abord que $\mathcal{B}(\Omega^0) \leq 0$ car

$$\mathcal{B}(\Omega^0) = \sum_{j=1}^{2m} \log s_j^0 = 2m \log(0.5) \leq 0.$$

A chaque itération, soit nous translatons une contrainte, soit nous en ajoutons une nouvelle. La fonction potentielle est réduite d'une constante donnée par (3.36) dans le premier cas, et par (3.45) dans le second.

Par construction, nous savons aussi que tant que l'algorithme tourne, $\mathcal{C} \subset \Omega^k \forall k \geq 0$. Considérons alors le polytope $\bar{\mathcal{C}}$, obtenu à partir du Ω^k final en gardant les contraintes non parallèles à celles de \mathcal{C} , et en remplaçant les autres par leurs homologues parallèles, propres à \mathcal{C} (cfr dessin).



Nous avons $\bar{\mathcal{C}} = \{y \in \mathcal{R}^m : c(c^k) - (A^k)^T y \geq 0\}$, où $c(c^k)$ est le vecteur c^k associé à Ω^k , dont nous avons modifié les composantes correspondant aux contraintes que l'on a translatées sur \mathcal{C} . Ces nouvelles composantes forment en réalité une partie des composantes de c , le

vecteur des termes indépendants associés à \mathcal{C} .

Par exemple, pour notre dessin,

$$c^k = (c_1^k, c_2^k, c_3^k, c_4^k, c_5^k, c_6^k)^T \quad \text{et} \quad c(c^k) = (c_1, c_2^k, c_3^k, c_4^k, c_4, c_6^k).$$

Dès lors, les composantes $c_i(c^k)$ de $c(c^k)$, et c_i^k de c^k sont liées par la relation

$$c_i(c^k) \leq c_i^k \quad i = 1, \dots, 2m + n(k),$$

où $n(k)$ est le nombre de contraintes ajoutées (et non translattées!) depuis l'itération 0 (NB : $n(k) \leq k$ car nous n'ajoutons pas de contraintes à chaque itération).

De plus, $\mathcal{C} \subset \bar{\mathcal{C}} \subset \Omega^k$. Ceci implique que

$$\mathcal{B}(\bar{\mathcal{C}}) = \sum_{j=1}^{2m+n(k)} (c_j(c^k) - (a_j^k)^T y) \leq \sum_{j=1}^{2m+n(k)} (c_j^k - (a_j^k)^T y) \leq \mathcal{B}(\Omega^k). \quad (3.50)$$

Nous pouvons encore exprimer $\bar{\mathcal{C}}$ sous la forme

$$\bar{\mathcal{C}} = \{y \in \mathcal{R}^m : \bar{c}^k - (\bar{A}^k)^T y \geq 0 \text{ et } \bar{c} - (\bar{A})^T y \geq 0\},$$

où \bar{c}^k (\bar{A}^k) est un sous-vecteur (une sous-matrice) de c^k (A^k), et \bar{c} (\bar{A}) est un sous-vecteur (une sous-matrice) de c (A).

Nous remarquons que $\text{int}(\bar{\mathcal{C}}) \neq \emptyset$ puisque $\mathcal{C} \subset \bar{\mathcal{C}}$ et $\text{int}(\mathcal{C}) \neq \emptyset$ par hypothèse. Donc, il existe un élément y de $\bar{\mathcal{C}}$ tel que $\bar{c}^k - (\bar{A}^k)^T y \geq 2^{-L}e$ et $\bar{c} - (\bar{A})^T y \geq 2^{-L}e$.

Par un raisonnement similaire à (3.29), nous obtenons

$$\mathcal{B}(\bar{\mathcal{C}}) \geq -\bar{q}L, \quad (3.51)$$

où \bar{q} est le nombre de contraintes de $\bar{\mathcal{C}}$.

Or $\dim(c(c^k)) = \dim(c^k)$. Dès lors, il y a autant de contraintes dans $\bar{\mathcal{C}}$ que dans Ω^k , et $\bar{q} = q$.

Dès lors, (3.50) et (3.51) entraînent

$$\mathcal{B}(\Omega^k) \geq \mathcal{B}(\bar{\mathcal{C}}) \geq -qL, \quad (3.52)$$

où q est le nombre de contraintes définissant Ω^k .

Cependant, à chaque itération de l'algorithme, la fonction potentielle duale $\mathcal{B}(\Omega^k)$ diminue d'une constante.

Dans le premier cas, où nous translattons une contrainte, la réduction de $\mathcal{B}(\Omega^k)$ est telle que $\mathcal{B}(\Omega^k) \leq \mathcal{B}(\Omega^{k-1}) - (1 - \beta)$. Par récurrence, cela implique que $\mathcal{B}(\Omega^k) \leq \mathcal{B}(\Omega^0) - k(1 - \beta)$. Il arrivera donc inévitablement une itération k pour laquelle $\mathcal{B}(\Omega^k) \leq -qL$, ce qui est en contradiction avec (3.52), qui est vérifié tant que $\mathcal{C} \subset \Omega^k$, c'est à dire aussi tant que l'algorithme tourne. L'algorithme s'arrêtera donc au plus tard à l'itération K telle que $\forall k \geq K$,

$$\mathcal{B}(\Omega^k) \leq -qL.$$

Il suffit donc que $\mathcal{B}(\Omega^0) - k(1 - \beta) \leq -qL$, c'est à dire que

$$k \geq \frac{qL}{1 - \beta} + \frac{\mathcal{B}(\Omega^0)}{1 - \beta}. \quad (3.53)$$

Dans le second cas, où nous ajoutons une contrainte, la réduction de $\mathcal{B}(\Omega^k)$ est telle que $\mathcal{B}(\Omega^k) \leq \mathcal{B}(\Omega^{k-1}) + \log \sqrt{2} - (1.5 - \beta)$. Par récurrence, cela implique que $\mathcal{B}(\Omega^k) \leq \mathcal{B}(\Omega^0) + k(\log \sqrt{2} - (1.5 - \beta))$.

De nouveau, nous allons atteindre une itération k pour laquelle $\mathcal{B}(\Omega^k) \leq -qL$, ce qui contredit (3.52). L'algorithme prendra fin au plus tard à l'itération K telle que $\forall k \geq K$, $\mathcal{B}(\Omega^k) \leq -qL$.

Il suffit donc que $\mathcal{B}(\Omega^0) + k(\log \sqrt{2} - (1.5 - \beta)) \leq -qL$, c'est à dire que

$$k \geq \frac{qL}{(1.5 - \beta) - \log \sqrt{2}} + \frac{\mathcal{B}(\Omega^0)}{(1.5 - \beta) - \log \sqrt{2}}. \quad (3.54)$$

Une borne supérieure sur le nombre d'itérations total de l'algorithme utilisant des centres analytiques exacts peut être obtenue en additionnant le nombre d'itérations nécessaires si un seul des cas survient à toutes les itérations.

Comme les bornes (3.53) et (3.54) sont toutes deux en $\mathcal{O}(qL)$, le nombre total d'itérations de l'algorithme est en $\mathcal{O}(qL)$.

Comme à chaque itération, le nombre de pas de Newton nécessaires pour générer le nouveau centre analytique est en $\mathcal{O}(\ln L)$, la génération d'un point admissible de \mathcal{C} grâce à l'algorithme avec centres analytiques exacts se fera en $\mathcal{O}(qL \ln L)$ pas de Newton.

CQFD.

3.4.2 Algorithme ACCPM utilisant des centres analytiques approximatifs

Dans cette section, nous allons montrer qu'il est inutile d'utiliser des centres analytiques exacts dans l'algorithme ACCPM. Nous pouvons en effet très bien générer un point admissible de \mathcal{C} par un algorithme de réduction de potentiel pour la génération de coupes qui utilise des centres approximatifs. Ceci est avantageux puisque, intuitivement, le nombre de pas de Newton pour atteindre un centre approximatif est moindre que celui nécessaire pour obtenir un centre analytique exact. Le temps d'exécution de l'algorithme s'en trouvera donc diminué.

Lemme 3.4.2

Soit $\Omega = \{y \in \mathcal{R}^m : s = c - A^T y \geq 0\}$, et posons (x^k, y^k) une paire intérieure admissible de Ω , qui satisfait

$$Ax^k = 0, \quad (s^k)^T x^k = c^T x^k = n, \quad \text{et} \quad \|X^k s^k - e\| \leq \gamma, \quad (3.55)$$

où $\gamma < 1$ et $s^k = c - A^T y^k$. Alors,

$$\mathcal{B}(\Omega) \geq \sum_{j=1}^n \log(s_j^k) \geq \mathcal{B}(\Omega) - \frac{\gamma^2}{2(1-\gamma)}. \quad (3.56)$$

preuve :

Par le lemme (2.2.1) du chapitre précédent, nous avons

$$\begin{aligned} \sum_{j=1}^n \log(x_j^k s_j^k) &= \sum_{j=1}^n \log(1 + x_j^k s_j^k - 1) \\ &= e^T X^k s^k - e^T e - \frac{\|X^k s^k - e\|^2}{2(1 - \|X^k s^k - e\|_\infty)} \\ &\geq e^T X^k s^k - n - \frac{\gamma^2}{2(1-\gamma)} \\ &= -\frac{\gamma^2}{2(1-\gamma)} \end{aligned} \quad (3.57)$$

Notons (x^a, y^a) la paire centrale de Ω .

Sachant que $\sum_{j=1}^n \log(x_j^a s_j^a) = \sum_{j=1}^n \log 1 = 0$, (3.57) peut encore s'écrire

$$\sum_{j=1}^n \log(x_j^a s_j^a) - \sum_{j=1}^n \log(x_j^k s_j^k) \leq \frac{\gamma^2}{2(1-\gamma)}. \quad (3.58)$$

Le terme de gauche de (3.58) peut se décomposer en

$$\sum_{j=1}^n \log(x_j^a) - \sum_{j=1}^n \log(x_j^k) + \sum_{j=1}^n \log(s_j^a) - \sum_{j=1}^n \log(s_j^k). \quad (3.59)$$

Comme y^a maximise la fonction potentielle duale sur Ω , nous avons

$$\sum_{j=1}^n \log(s_j^a) - \sum_{j=1}^n \log(s_j^k) \geq 0, \quad (3.60)$$

et donc la première inégalité de (3.56) est démontrée puisque $\mathcal{B}(\Omega) := \sum_{j=1}^n \log s_j^a$.

D'autre part, nous savons que x^a est solution du problème (1.4). Donc

$$\sum_{j=1}^n \log x_j^a - \sum_{j=1}^n \log x_j^k \geq 0, \quad (3.61)$$

puisque (3.55) assure que x^k satisfait les contraintes du problème (1.4).

Dès lors, par (3.58), (3.59) et (3.61), nous obtenons

$$\sum_{j=1}^n \log(s_j^a) - \sum_{j=1}^n \log(s_j^k) \leq \frac{\gamma^2}{2(1-\gamma)}$$

i.e.

$$\sum_{j=1}^n \log(s_j^k) \geq \mathcal{B}(\Omega) - \frac{\gamma^2}{2(1-\gamma)},$$

ce qui démontre la deuxième inégalité.

CQFD.

Ce lemme indique que la valeur du potentiel en un centre approximatif de Ω caractérisé par (3.55) diffère de celle obtenue en un centre analytique exact d'une petite constante. De plus, associé à (3.52), il nous permet de vérifier, tant que l'algorithme tourne, que

$$\sum_{j=1}^n \log s_j^k \geq \mathcal{B}(\Omega^k) - \frac{\gamma^2}{2(1-\gamma)} \geq -qL - \frac{\gamma^2}{2(1-\gamma)}, \quad (3.62)$$

où s^k est ici associé à y^k , le centre analytique approximatif de Ω^k .

Algorithme ACCPM utilisant des centres approximatifs

BEGIN

Initialisation

Elle est tout à fait identique à celle de l'algorithme utilisant des centres analytiques exacts (cfr section (3.4.1)).

$k^{\text{ème}}$ itération

On vérifie si le centre analytique approximatif courant y^k , répondant aux contraintes (3.55), satisfait toutes les inégalités de \mathcal{C} . Si oui, l'algorithme se termine, sinon, $\exists \bar{j}$ tel que $c_{\bar{j}} - a_{\bar{j}}^T y^k < 0$.

Deux cas peuvent à nouveau se présenter.

Si $a_{\bar{j}}$ est déjà une colonne de A^k , la mise à jour est la suivante :

$$\begin{cases} c_{\bar{j}}^{k+1} = s_{\bar{j}}^k - \frac{(1-\beta)}{a_{\bar{j}}^k} + a_{\bar{j}}^T y^k \\ c_j^{k+1} = c_j^k \quad j \neq \bar{j}. \end{cases} \quad (3.63)$$

et

$$A^{k+1} = A^k \quad (3.64)$$

Remarquons que cette mise à jour est la même que celle utilisée pour l'algorithme avec centres analytiques exacts, puisque dans ce cas, nous avons (cfr (3.32) et (3.33))

$$c_{\bar{j}}^{k+1} = \beta s_{\bar{j}}^k + a_{\bar{j}}^T y^k = \beta s_{\bar{j}}^k + c_{\bar{j}}^k - s_{\bar{j}}^k$$

$$\begin{aligned}
&= c_j^k - (1 - \beta)s_j^k \\
&= s_j^k + a_j^T y^k - (1 - \beta)s_j^k \\
&= s_j^k + a_j^T y^k - (1 - \beta)/x_j^k
\end{aligned}$$

puisque dans le premier algorithme, (x^k, s^k) était une paire centrale exacte de Ω^k , et donc $x^k s^k = e$.

Si $a_{\bar{j}}$ n'est pas une colonne de A^k , la mise à jour est alors

$$c^{k+1} = \begin{pmatrix} c^k \\ \beta \bar{r} + a_{\bar{j}}^T y^k \end{pmatrix} \quad \text{et} \quad A^{k+1} = (A^k, a_{\bar{j}}), \quad (3.65)$$

où $\bar{r} = \sqrt{a_{\bar{j}}^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}}$, avec x^k le point primal associé à y^k , centre analytique approximatif.

La $k^{\text{ème}}$ itération s'achève par la mise à jour du centre approximatif. Cette étape consiste à calculer la paire (x^{k+1}, y^{k+1}) , où y^{k+1} est un centre analytique approximatif de Ω^{k+1} tel que $\|X^{k+1} s^{k+1} - e\| \leq \gamma$ (où $s^{k+1} = c^{k+1} - (A^{k+1})^T y^{k+1}$), en utilisant la méthode de Newton pure à partir de y^k , γ -centre de Ω^k et déjà centre approximatif de Ω^{k+1} . Nous passons ensuite à l'itération $k + 1$.

END.

Avant de clarifier cette étape de mise à jour, montrons que, comme pour l'algorithme utilisant des centres analytiques exacts, la fonction potentielle diminue d'une constante à chaque itération, quelque soit le cas traité, c'est à dire que $a_{\bar{j}}$ soit une colonne de A^k ou pas.

Théorème 3.6

Soit (x^k, y^k) une paire centrale approximative de Ω^k vérifiant

$$A^k x^k = 0, \quad (s^k)^T x^k = (c^k)^T x^k = n, \quad \text{et} \quad \|X^k s^k - e\| \leq \gamma.$$

Alors, dans le premier cas de l'algorithme, pour lequel $a_{\bar{j}}$ est une colonne de A^k ,

$$\mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) \leq (\beta - 1) + \frac{\gamma^2}{2(1 - \gamma)}. \quad (3.66)$$

preuve :

Considérons \bar{x}^a, \bar{y}^a la paire centrale exacte de Ω^{k+1} . Nous avons

$$\begin{aligned}
e^T X^k \bar{s}^a &= e^T X^k (c^{k+1} - (A^{k+1})^T \bar{y}^a) \\
&= e^T X^k (c^{k+1} - (A^k)^T \bar{y}^a) \\
&= e^T X^k c^{k+1} \quad \text{car } A^k x^k = 0 \\
&= e^T X^k (c_1^k, \dots, s_j^k + a_j^T y^k - (1 - \beta)/x_j^k, \dots, c_n^k)^T \tag{3.67}
\end{aligned}$$

$$\begin{aligned}
&= e^T X^k c^k - x_j^k (1 - \beta)/x_j^k \\
&= n - 1 + \beta \tag{3.68}
\end{aligned}$$

Par le lemme (3.4.2), nous savons que $\mathcal{B}(\Omega^k) \geq \sum_{j=1}^n \log s_j^k$. Donc,

$$\frac{\exp \mathcal{B}(\Omega^{k+1})}{\exp \mathcal{B}(\Omega^k)} \leq \frac{\prod_{j=1}^n \bar{s}_j^a}{\prod_{j=1}^n s_j^k} = \prod_{j=1}^n (\bar{s}_j^a x_j^k) \prod_{j=1}^n \left(\frac{1}{s_j^k x_j^k} \right) \tag{3.69}$$

$$\leq \left(\frac{1}{n} \sum_{j=1}^n \bar{s}_j^a x_j^k \right)^n \prod_{j=1}^n \left(\frac{1}{s_j^k x_j^k} \right) \tag{3.70}$$

$$= \left(\frac{n-1+\beta}{n} \right)^n \prod_{j=1}^n \left(\frac{1}{s_j^k x_j^k} \right) \tag{3.71}$$

$$\leq \exp(\beta-1) \prod_{j=1}^n \left(\frac{1}{s_j^k x_j^k} \right), \tag{3.72}$$

où l'inégalité (3.70) vient de la relation entre moyenne géométrique et arithmétique (2.55), et où (3.72) vient, elle, de la relation $\ln(1 + \frac{x}{n}) \leq \frac{x}{n}$, i.e. $(1 + \frac{x}{n})^n \leq e^x$, $x \geq 0$.

Donc,

$$\begin{aligned}
\mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) &\leq (\beta-1) + \sum_{j=1}^n \log\left(\frac{1}{s_j^k x_j^k}\right) \\
&= (\beta-1) - \sum_{j=1}^n \log(s_j^k x_j^k) \\
&\leq (\beta-1) + \frac{\gamma^2}{2(1-\gamma)} \quad \text{par (3.57)}.
\end{aligned}$$

CQFD.

Avant d'établir un théorème similaire dans le deuxième cas de l'algorithme, nous avons besoin de la propriété suivante.

Propriété 3.4.3

Nous vérifions que

$$A^k(X^k)^2(A^k)^T = A^k(S^k)^{-1}(S^k X^k)^2(S^k)^{-1}(A^k)^T \quad (3.73)$$

$$\succeq (\min x_j^k s_j^k)^2((Y^k)^{-2} + (I - Y^k)^{-2}). \quad (3.74)$$

preuve :

Le procédé à utiliser est tout à fait similaire à celui qui a permis de démontrer la propriété (3.4.1).

CQFD.

Théorème 3.7

Soit (x^k, y^k) une paire centrale approximative de Ω^k vérifiant $A^k x^k = 0$, $(s^k)^T x^k = (c^k)^T x^k = n$ et $\|X^k s^k - e\| \leq \gamma$. Alors, dans le deuxième cas de l'algorithme, pour lequel $a_{\bar{j}}$ n'est pas une colonne de A^k ,

$$\mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) \leq \log \sqrt{2} - \log(1 - \gamma) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)}. \quad (3.75)$$

preuve :

Considérons $\bar{s}^a = c^k - (A^k)^T \bar{y}^a > 0$ le vecteur composé des n premiers écarts aux contraintes de Ω^{k+1} associés à \bar{y}^a , le centre analytique exact de Ω^{k+1} . L'écart à la dernière contrainte, quant à lui, vérifie

$$\begin{aligned} \bar{s}_{n+1}^a &= c_{n+1}^{k+1} - (a_{n+1}^{k+1})^T \bar{y}^a \\ &= \beta \bar{r} + a_{\bar{j}}^T y^k - a_{\bar{j}}^T \bar{y}^a \\ &= a_{\bar{j}}^T (y^k - \bar{y}^a) + \beta \bar{r} \\ &= a_{\bar{j}}^T (A^k(X^k)^2(A^k)^T)^{-1} (A^k(X^k)^2(A^k)^T) (y^k - \bar{y}^a) + \beta \bar{r} \\ &= a_{\bar{j}}^T (A^k(X^k)^2(A^k)^T)^{-1} A^k(X^k)^2 ((A^k)^T y^k - (A^k)^T \bar{y}^a) + \beta \bar{r} \\ &= a_{\bar{j}}^T (A^k(X^k)^2(A^k)^T)^{-1} A^k(X^k)^2 (-c^k + (A^k)^T y^k + c^k - (A^k)^T \bar{y}^a) + \beta \bar{r} \\ &= a_{\bar{j}}^T (A^k(X^k)^2(A^k)^T)^{-1} A^k(X^k)^2 (\bar{s}^a - s^k) + \beta \bar{r} \\ &= a_{\bar{j}}^T (A^k(X^k)^2(A^k)^T)^{-1} A^k X^k (X^k \bar{s}^a - X^k s^k) + \beta \bar{r} \\ &\leq \|a_{\bar{j}}^T (A^k(X^k)^2(A^k)^T)^{-1} A^k X^k\| \|X^k \bar{s}^a - X^k s^k\| + \beta \bar{r} \\ &= (\|X^k \bar{s}^a - X^k s^k\| + \beta) \bar{r} \end{aligned} \quad (3.76)$$

$$\begin{aligned} &\leq (\|X^k \bar{s}^a - e\| + \|X^k s^k - e\| + \beta) \bar{r} \\ &\leq (\|X^k \bar{s}^a - e\| + \gamma + \beta) \bar{r} \end{aligned} \quad (3.77)$$

où (3.76) est obtenue par un raisonnement similaire à celui utilisé pour (3.20), avec \bar{r} défini ici par $\bar{r} = \sqrt{a_j^T (A^k (X^k)^2 (A^k)^T)^{-1} a_j}$, comme nous l'avons vu dans la mise à jour (3.65).

Donc,

$$\begin{aligned} \frac{\exp \mathcal{B}(\Omega^{k+1})}{\bar{r} \exp \mathcal{B}(\Omega^k)} &\leq \frac{\bar{s}_{n+1}^a \prod_{j=1}^n \bar{s}_j^a}{\bar{r} \prod_{j=1}^n s_j^k} \quad \text{car } \mathcal{B}(\Omega^k) \geq \sum_{j=1}^n \log s_j^k \quad (\text{cfr}(3.56)) \\ &= \frac{\bar{s}_{n+1}^a}{\bar{r}} \left(\prod_{j=1}^n \bar{s}_j^a x_j^k \right) \left(\prod_{j=1}^n \frac{1}{x_j^k s_j^k} \right) \\ &\leq (\|X^k \bar{s}^a - e\| + \gamma + \beta) \left(\prod_{j=1}^n \bar{s}_j^a x_j^k \right) \left(\prod_{j=1}^n \frac{1}{x_j^k s_j^k} \right) \quad \text{par (3.77)}. \end{aligned} \quad (3.78)$$

De plus,

$$e^T X^k \bar{s}^a = e^T X^k (c^k - (A^k)^T \bar{y}^a) = e^T X^k c^k = n. \quad (3.79)$$

De la même façon que celle utilisée dans le théorème (3.2), nous allons essayer d'évaluer le maximum du terme droit de (3.78), en tenant compte de (3.79).

Posons $\alpha = X^k \bar{s}^a > 0$.

Le problème à résoudre est le suivant

$$\begin{cases} \max \psi(\alpha) = (\|\alpha - e\| + \gamma + \beta) \prod_{j=1}^n \alpha_j \prod_{j=1}^n \frac{1}{s_j^k x_j^k} \\ \text{s.c. } e^T \alpha = n \text{ et } \alpha > 0. \end{cases} \quad (3.80)$$

Nous savons que la solution du problème (3.80) est atteinte en α tel que $\alpha_1 = \delta > 1$ et $\alpha_2 = \dots = \alpha_n = \frac{n-\delta}{n-1} > 0$. D'où

$$\begin{aligned} \psi(\alpha) &\leq ((\delta - 1) \sqrt{\frac{n}{n-1}} + \gamma + \beta) \delta \left(\frac{n-\delta}{n-1} \right)^{n-1} \prod_{j=1}^n (s_j^k x_j^k)^{-1} \\ &\leq \frac{4}{\exp(1.5 - \beta - \gamma)} \prod_{j=1}^n (s_j^k x_j^k)^{-1}, \end{aligned}$$

par un raisonnement semblable à celui du théorème (3.2).

Dès lors, (3.78) implique $\frac{\exp \mathcal{B}(\Omega^{k+1})}{\exp \mathcal{B}(\Omega^k)} \leq \frac{4\bar{r}}{\exp(1.5 - \beta - \gamma)} \prod_{j=1}^n (s_j^k x_j^k)^{-1}$, et donc

$$\mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) \leq \log(4\bar{r}) - (1.5 - \beta - \gamma) - \sum_{j=1}^n \log(s_j^k x_j^k) \quad (3.81)$$

$$\leq \log(4\bar{r}) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)} \quad \text{par (3.57)}. \quad (3.82)$$

Remarquons que le terme droit de (3.82) dépend de l'itération k puisque

$\bar{r} = \sqrt{a_j^T (A^k (X^k)^2 (A^k)^T)^{-1} a_j}$ en dépend aussi. Ce terme peut toutefois être borné par une

constante puisque l'on montre que $\bar{r} \leq \frac{1}{8(1-\gamma)^2}$. En effet, grâce aux propriétés (3.4.2) et (3.4.3), nous vérifions que

$$\begin{aligned} A^k(X^k)^2(A^k)^T &\succeq (\min x_j^k s_j^k)^2((Y^k)^{-2} + (I - Y^k)^{-2}) \\ &\succeq (1 - \gamma)^2((Y^k)^{-2} + (I - Y^k)^{-2}) \\ &\succeq 8(1 - \gamma)^2 I \end{aligned}$$

puisque $\|X^k s^k - e\|_\infty \leq \|X^k s^k - e\| \leq \gamma$ par hypothèse.

Cela signifie que $A^k(X^k)^2(A^k)^T - 8(1 - \gamma)^2 I$ est semi définie positive. D'où

$$\bar{r}^2 = a_j^T (A^k(X^k)^2(A^k)^T)^{-1} a_j \leq \frac{\|a_j\|^2}{8(1 - \gamma)^2} = \frac{1}{8(1 - \gamma)^2}.$$

On peut dès lors borner supérieurement le terme droit de (3.82), et obtenir

$$\begin{aligned} \mathcal{B}(\Omega^{k+1}) - \mathcal{B}(\Omega^k) &\leq \log(4\bar{r}) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)} \\ &\leq \log \frac{2}{\sqrt{2}(1 - \gamma)} - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)} \\ &= \log \sqrt{2} - \log(1 - \gamma) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)}. \end{aligned}$$

CQFD.

Remarquons que, comme dans le cas de l'algorithme avec centres analytiques exacts, la décroissance du potentiel n'est assurée que pour des valeurs appropriées de β et γ . Ces valeurs doivent être telles que les termes de droite de (3.66) et (3.75) soient négatifs.

Exemple :

Pour que le terme droit de (3.66) soit négatif, prenons $\beta = 1/2$. Dans ce cas, pour γ suffisamment petit, nous vérifions que $(\beta - 1) + \frac{\gamma^2}{2(1-\gamma)} < 0$.

ex : $\gamma = 0.3 : (\beta - 1) + \frac{\gamma^2}{2(1-\gamma)} = -0,4357 < 0$.

Pour que le terme droit de (3.75) soit négatif, prenons $\beta = 1/4$. Dans ce cas, pour γ suffisamment petit, nous vérifions que $\log(\sqrt{2}) - \log(1 - \gamma) - (1.5 - \gamma - \beta) + \frac{\gamma^2}{2(1-\gamma)} < 0$.

ex : $\gamma = 0.3 : \log(\sqrt{2}) - \log(1 - \gamma) - (1.5 - \gamma - \beta) + \frac{\gamma^2}{2(1-\gamma)} = -0.1825 < 0$.

Après avoir analysé en détail la première étape de l'itération k de l'algorithme, clarifions la deuxième partie, qui consiste à mettre à jour le centre analytique approximatif.

Pour cela, nous utilisons le lemme suivant, semblable au lemme (3.4.1), établi dans le cas de l'algorithme avec centres analytiques exacts.

Lemme 3.4.3

Soit (x^k, y^k) une paire centrale approximative de Ω^k telle que

$$A^k x^k = 0, \quad (s^k)^T x^k = n, \quad \|X^k s^k - e\| \leq \gamma \quad (3.83)$$

où $s^k = c^k - (A^k)^T y^k$.

Alors, dans les deux cas de l'algorithme, il existe $\bar{x} > 0$ tel que

$$A^{k+1} \bar{x} = 0 \quad \text{et} \quad \|\bar{X} \bar{s} - e\| \leq \lambda \quad \text{pour une constante } \lambda < 1,$$

où $\bar{s} = c^{k+1} - (A^{k+1})^T y^k \in \mathcal{R}^n$.

preuve :

Dans le cas où $a_{\bar{j}}$ est déjà une colonne de A^k , la mise à jour, donnée par (3.63) et (3.64), est telle

$$\bar{s}_j = c_j^{k+1} - (a_j^{k+1})^T y^k = c_j^k - (a_j^k)^T y^k = s_j^k \quad \text{si } j \neq \bar{j}.$$

et

$$\bar{s}_{\bar{j}} = c_{\bar{j}}^{k+1} - (a_{\bar{j}}^{k+1})^T y^k = s_{\bar{j}}^k - \frac{(1-\beta)}{x_{\bar{j}}^k} + a_{\bar{j}}^T y^k - a_{\bar{j}}^T y^k = s_{\bar{j}}^k - \frac{(1-\beta)}{x_{\bar{j}}^k},$$

puisque $a_{\bar{j}}^{k+1} = a_{\bar{j}}^k = a_{\bar{j}}$.

Posons $\bar{x} = x^k > 0$. Dès lors, $A^{k+1} \bar{x} = A^k x^k = 0$, et

$$\begin{aligned} \|\bar{X} \bar{s} - e\| &= \|X^k s^k - e + \bar{X} \bar{s} - X^k s^k\| \\ &\leq \|X^k s^k - e\| + \|\bar{X} \bar{s} - X^k s^k\| \\ &\leq \gamma + (1-\beta) \end{aligned}$$

car $\|X^k s^k - e\| \leq \gamma$ et

$$\begin{aligned} \|\bar{X} \bar{s} - X^k s^k\| &= \|(x_1^k s_1^k - x_1^k s_1^k, \dots, x_{\bar{j}}^k [s_{\bar{j}}^k - (1-\beta)/x_{\bar{j}}^k] - x_{\bar{j}}^k s_{\bar{j}}^k, \dots, x_n^k s_n^k - x_n^k s_n^k)^T\| \\ &= |-(1-\beta)| = (1-\beta). \end{aligned}$$

Pour avoir $\|\bar{X} \bar{s} - e\| < 1$, il faut donc choisir $0 < \gamma < \beta < 1$.

Dans le cas où $a_{\bar{j}}$ n'est pas une colonne de A^k , la mise à jour (3.65) est telle que $\bar{s} = ((s^k)^T \beta \bar{r})^T$.

En effet, \bar{s} est identique à celui obtenu en (3.46) pour l'algorithme avec centres analytiques exacts puisque la mise à jour est identique (cfr (3.34)).

Soit $\alpha < 1$ et $\Delta x = -\frac{\alpha}{\bar{r}} X^k (A^k)^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}$, où $\bar{r} = \sqrt{a_{\bar{j}}^T (A^k (X^k)^2 (A^k)^T)^{-1} a_{\bar{j}}}$.

Alors, de nouveau,

$$\|\Delta x\| = \frac{\alpha}{\bar{r}} \bar{r} = \alpha < 1. \quad (3.84)$$

Posons maintenant

$$\bar{x} = \begin{pmatrix} X^k(e + \Delta x) \\ \alpha/\bar{r} \end{pmatrix} > 0.$$

Alors, $A^{k+1}\bar{x} = A^k X^k e + A^k X^k \Delta x + \frac{\alpha}{\bar{r}} a_j = A^k X^k \Delta x + \frac{\alpha}{\bar{r}} a_j = 0$ (par (3.49)), et

$$\begin{aligned} \|\bar{X}\bar{s} - e\|^2 &= \|X^k S^k(e + \Delta x) - e\|^2 + \left(\beta\bar{r}\frac{\alpha}{\bar{r}} - 1\right)^2 \\ &\leq (\|X^k S^k e - e\| + \|X^k S^k \Delta x\|)^2 + (1 - \alpha\beta)^2 \\ &\leq (\|X^k S^k e - e\| + \|X^k S^k\| \|\Delta x\|)^2 + (1 - \alpha\beta)^2. \end{aligned} \quad (3.85)$$

Or, $\|X^k S^k\| \leq 1 + \gamma$ puisque, $X^k S^k$ étant diagonale, $\|X^k S^k\| = \lambda_{max}(X^k S^k) = \|x^k s^k\|_\infty = x_i^k s_i^k \leq |x_i^k s_i^k - 1| + 1 \leq \gamma + 1$, où i est ici l'indice correspondant à la composante maximale du vecteur $x^k s^k$.

Comme $\|X^k S^k\| \leq 1 + \gamma$, et par (3.83), (3.84) et (3.85), nous obtenons

$$\begin{aligned} \|\bar{X}\bar{s} - e\|^2 &\leq (\|X^k S^k e - e\| + \|X^k S^k\| \|\Delta x\|)^2 + (1 - \alpha\beta)^2 \\ &\leq (\gamma + (1 + \gamma)\alpha)^2 + (1 - \alpha\beta)^2 \\ &\leq 1 + \gamma^2 - \frac{(\beta - \gamma(1 + \gamma))^2}{(1 + \gamma)^2 + \beta^2} \end{aligned} \quad (3.86)$$

pour $\alpha = \frac{\beta - \gamma(1 + \gamma)}{(1 + \gamma)^2 + \beta^2}$, puisqu'il suffit de vérifier, par une simple distribution, que $(\gamma + (1 + \gamma)\alpha)^2 + (1 - \alpha\beta)^2 \leq 1 + \gamma^2 - \alpha(\beta - \gamma(1 + \gamma))$.

Le terme droit de (3.86) sera inférieur à l'unité si l'on choisit adéquatement β et γ .

Exemple :

Prenons $\beta = 1/2$.

Il faut choisir $\gamma \in (0, 1)$ tel que $\gamma^2 - \frac{(\beta - \gamma(1 + \gamma))^2}{(1 + \gamma)^2 + \beta^2} < 0$.

Après transformation, cela revient à prendre $\gamma \in (0, 1)$ tel que $\frac{5}{4}\gamma^2 + \gamma - \frac{1}{4} < 0$, c'est à dire $0 < \gamma < \frac{1}{5}$. Il faut donc choisir γ suffisamment petit.

CQFD.

Le lemme (3.4.3) montre que, même s'il n'est pas parfaitement centré, y^k est toutefois dans la région de convergence quadratique de Ω^{k+1} pour la méthode de Newton (cfr théorème (2.1) *i*) et *iv*)).

Une paire centrale approximative (x^{k+1}, y^{k+1}) telle que $\|X^{k+1} s^{k+1} - e\| \leq \gamma$ peut ainsi être obtenue après un nombre constant de pas de Newton à partir de (x^k, y^k) . Ce nombre est d'autant plus grand que la précision souhaitée est importante.

Nous concluons ce chapitre par un théorème donnant la complexité de l'algorithme dont il fut question dans cette deuxième partie, c'est à dire l'algorithme ACCPM avec réduction de potentiel utilisant des centres approximatifs.

Théorème 3.8

L'algorithme ACCPM avec réduction de potentiel utilisant des centres approximatifs génère un point admissible de \mathcal{C} en $\mathcal{O}(qL)$ itérations et $\mathcal{O}(qL)$ pas de Newton, où q est le nombre d'inégalités du polytope final Ω^k , et où L est défini en (3.28).

preuve :

Par (3.62), nous savons que tant que l'algorithme tourne,

$$\sum_{j=1}^n \log s_j^k \geq -qL - \frac{\gamma^2}{2(1-\gamma)},$$

où s^k est l'écart associé à y^k , un centre approximatif de Ω^k . Or, nous montrons qu'après $\mathcal{O}(qL)$ itérations,

$$\sum_{j=1}^n \log s_j^k \leq -qL - \frac{\gamma^2}{2(1-\gamma)}.$$

En effet, il suffit de constater qu'après $\mathcal{O}(qL)$ itérations,

$$\mathcal{B}(\Omega^k) \leq -qL - \frac{\gamma^2}{2(1-\gamma)},$$

puisque, par le lemme (3.4.2), $\mathcal{B}(\Omega) \geq \sum_{j=1}^n \log(s_j^k)$.

Or, dans le premier cas de l'algorithme, où a_j est une colonne de A^k , la décroissance de la fonction potentielle entre deux itérés successifs est donnée par (3.66), où β et γ prennent des valeurs telles que la décroissance de $\mathcal{B}(\Omega^k)$ est assurée. Donc par récurrence

$$\mathcal{B}(\Omega^k) \leq \mathcal{B}(\Omega^0) + k((\beta - 1) + \frac{\gamma^2}{2(1-\gamma)}).$$

Dès lors, il arrivera inévitablement une itération k pour laquelle $\mathcal{B}(\Omega^k) \leq -qL - \frac{\gamma^2}{2(1-\gamma)}$.

L'algorithme s'arrêtera donc au plus tard à l'itération K telle que $\forall k \geq K$, $\mathcal{B}(\Omega^k) \leq -qL - \frac{\gamma^2}{2(1-\gamma)}$.

Pour que cette inégalité soit satisfaite, il suffit que $\mathcal{B}(\Omega^0) + k((\beta - 1) + \frac{\gamma^2}{2(1-\gamma)}) \leq -qL - \frac{\gamma^2}{2(1-\gamma)}$, c'est à dire que

$$k \geq \frac{(-\mathcal{B}(\Omega^0) - qL - \frac{\gamma^2}{2(1-\gamma)})}{\beta - 1 + \frac{\gamma^2}{2(1-\gamma)}} \quad (3.87)$$

puisque la décroissance de la fonction potentielle est assurée par des valeurs de β et γ telles que $(\beta - 1) + \frac{\gamma^2}{2(1-\gamma)} < 0$.

Dans le second cas de l'algorithme, où a_j n'est pas une colonne de A^k , la décroissance

de $\mathcal{B}(\Omega^k)$ entre deux itérés successifs est donnée par (3.75) où β et γ prennent des valeurs qui assurent la décroissance de $\mathcal{B}(\Omega^k)$. Donc par récurrence

$$\mathcal{B}(\Omega^k) \leq \mathcal{B}(\Omega^0) + k(\log \sqrt{2} - \log(1 - \gamma) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)}).$$

Dès lors, il arrivera une itération pour laquelle $\mathcal{B}(\Omega^k) \leq -qL - \frac{\gamma^2}{2(1 - \gamma)}$.

L'algorithme prendra donc fin au plus tard à l'itération K telle que $\forall k \geq K$, $\mathcal{B}(\Omega^k) \leq -qL - \frac{\gamma^2}{2(1 - \gamma)}$. Pour qu'une telle inégalité soit satisfaite, il suffit que

$$\mathcal{B}(\Omega^0) + k(\log \sqrt{2} - \log(1 - \gamma) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)}) \leq -qL - \frac{\gamma^2}{2(1 - \gamma)},$$

c'est à dire que

$$k \geq \frac{(-\mathcal{B}(\Omega^0) - qL - \frac{\gamma^2}{2(1 - \gamma)})}{\log \sqrt{2} - \log(1 - \gamma) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)}} \quad (3.88)$$

puisque β et γ sont choisis de telle sorte que $\mathcal{B}(\Omega^k)$ diminue à chaque itération, c'est à dire que $\log \sqrt{2} - \log(1 - \gamma) - (1.5 - \beta - \gamma) + \frac{\gamma^2}{2(1 - \gamma)} < 0$.

Une borne supérieure sur le nombre d'itérations de l'algorithme utilisant des centres approximatifs peut donc être obtenue en additionnant le nombre d'itérations nécessaires si seul un des cas survient à toutes les itérations. Comme les bornes (3.87) et (3.88) sont en $\mathcal{O}(qL)$, le nombre total d'itérations est borné par un nombre en $\mathcal{O}(qL)$.

Comme à chaque itération, le nombre de pas de Newton nécessaires pour générer le nouveau centre approximatif tel que $\|X^k s^k - e\| \leq \gamma$ est constant, la génération d'un point admissible de \mathcal{C} par l'algorithme ACCPM avec réduction de potentiel utilisant des centres approximatifs s'effectuera en un nombre de pas de Newton de l'ordre de $\mathcal{O}(qL)$.

CQFD.

Chapitre 4

Algorithme ACCPM basé sur le principe de réduction de potentiel pour un problème d'admissibilité convexe

4.1 Introduction

Dans le chapitre précédent, nous avons étudié un algorithme ACCPM pour la résolution du problème d'admissibilité linéaire. A travers ce nouveau chapitre, nous développerons un algorithme ACCPM destiné à résoudre un problème d'admissibilité convexe, qui consiste à trouver un point admissible dans

$$\mathcal{C} = \{y \in \mathcal{R}^m : f_i(y) \leq 0 \ i = 1, 2, \dots\},$$

où chaque $f_i : \mathcal{R}^m \rightarrow \mathcal{R}$ est convexe, où $\text{int}(\mathcal{C})$ est supposé non vide, et où \mathcal{C} est inclus dans un polytope $\Omega^0 = \{y \in \mathcal{R}^m : 0 \leq y \leq e\} = [0, 1]^m$.

Le principe de cet algorithme est proche de celui développé pour le problème d'admissibilité linéaire, en ce sens que l'on construit toujours des polytopes Ω^k ($k = 1, 2, \dots$) à partir de Ω^0 , et ce jusqu'à ce que le centre analytique y^k du polytope courant appartienne à l'ensemble admissible \mathcal{C} .

Cependant, tandis que l'algorithme destiné à résoudre le problème d'admissibilité linéaire translatait ou ajoutait une contrainte selon que la contrainte \bar{j} non vérifiée par y^k était parallèle ou non à une contrainte de Ω^k , il n'en sera pas de même pour l'algorithme destiné à la résolution du problème d'admissibilité convexe. En effet, ce dernier ne fera qu'ajouter des contraintes à chaque itération, puisque les contraintes de \mathcal{C} ne sont plus nécessairement linéaires (et donc pas parallèles à celles de Ω^k).

De plus, la contrainte ajoutée passera toujours par un centre analytique approximatif du polytope courant. On ne s'encombrera donc plus du terme $\beta\bar{r}$ présent dans l'algorithme pour le problème d'admissibilité linéaire (cfr 3.65).

Notons encore que le vecteur normal de la contrainte ajoutée sera choisi comme annoncé au premier chapitre.

Outre les raisonnements intermédiaires, destinés à établir une relation entre les fonctions potentielles $\mathcal{B}(\Omega^k)$ à chaque itération, et proches de ceux suivis au chapitre précédent, nous obtiendrons un résultat de complexité algorithmique pour l'algorithme ACCPM destiné à la résolution du problème d'admissibilité convexe. Elle sera plus précise que celle obtenue pour le problème d'admissibilité linéaire, d'où son intérêt.

4.2 Préliminaires

Avant d'étudier l'algorithme proprement dit, établissons, en utilisant les mêmes notations, une petite extension du théorème (2.1), démontré au Chapitre 2.

Lemme 4.2.1

Soit (y, s) une paire admissible intérieure, y^a le centre analytique de Ω . Soit également x un point intérieur de \mathcal{X}_Ω , et x^a le minimiseur du potentiel primal tel que $c^T x^a = n$. Alors, si $\delta(s) < 1$,

$$\mathcal{B}(y^a) - \mathcal{B}(y) \leq \frac{\delta(s)^2}{1 - \delta(s)^2} \quad \text{i.e.} \quad \mathcal{B}(\Omega) - \mathcal{B}(y) \leq \frac{\delta(s)^2}{1 - \delta(s)^2}.$$

preuve :

Soit $y^+ = y + \Delta y$ où $\Delta y = -(AS^{-2}A^T)^{-1}AS^{-1}e$ représente la direction de Newton duale associée au problème $\max_{y \in \Omega} \sum_{j=1}^n \log s_j$. Comme $\mathcal{B}(y) = \sum_{j=1}^n \log s_j$ est une fonction concave,

$$\mathcal{B}(y^+) - \mathcal{B}(y) \leq \nabla \mathcal{B}(y)^T (y^+ - y) = -(AS^{-1}e)^T \Delta y = e^T S^{-1} A^T (AS^{-2}A^T)^{-1} AS^{-1}e = \delta(s)^2 \quad (4.1)$$

Considérons la séquence générée par ce processus, telle que $y^0 = y$ et $y^{k+1} = y^k + \Delta(y^k)$. Elle est telle que $\mathcal{B}(y^{k+1}) - \mathcal{B}(y^k) \leq \delta(s^k)^2$.

Associée à la relation (2.21) du théorème (2.1), nous avons que

$$\mathcal{B}(y^{k+1}) - \mathcal{B}(y^0) \leq \sum_{j=1}^k \delta(s)^{2^j} \leq \sum_{j=1}^k (\delta(s)^2)^j \leq \frac{\delta(s)^2}{1 - \delta(s)^2}$$

puisque $\delta(s) < 1$. En passant à la limite,

$$\mathcal{B}(y^a) - \mathcal{B}(y) \leq \frac{\delta(s)^2}{1 - \delta(s)^2}.$$

CQFD.

4.3 Algorithme ACCPM pour la résolution du problème d'admissibilité convexe

BEGIN

Initialisation

$$A^0 = (I \quad -I) \in \mathcal{R}^{m \times 2m}, \quad c^0 = (e^T \quad 0)^T \in \mathcal{R}^{2m}$$

$$y^0 = 0.5e \in \mathcal{R}^m, \quad s^0 = c^0 - (A^0)^T y^0 = 0.5e \in \mathcal{R}^{2m}, \quad \text{et } x^0 = 2e \in \mathcal{R}^{2m}.$$

De toute évidence, $A^0 x^0 = 0$ et $\|X^0 s^0 - e\| = 0$.

En d'autres mots, y^0 et x^0 forment la paire centrale de $\Omega^k = \{y \in \mathcal{R}^m : c^k - (A^k)^T y \geq 0\}$ avec $k = 0$.

$k^{\text{ème}}$ itération

Vérifier si le centre approximatif y^k de $\Omega^k \in \text{int}(\mathcal{C})$.

Si oui, l'algorithme se termine et y^k est le point admissible recherché.

Sinon,

1) générer un hyperplan $\{y \in \mathcal{R}^m : a_{k+1}^T y \leq a_{k+1}^T y^k\} \supset \mathcal{C}$ tel que $\|a_{k+1}\| = 1$, et poser

$$\Omega^{k+1} = \{y \in \mathcal{R}^m : c^{k+1} - (A^{k+1})^T y \geq 0\} \quad (4.2)$$

$$\text{où } A^{k+1} = (A^k, a_{k+1}) \text{ et } c^{k+1} = \begin{pmatrix} c^k \\ a_{k+1}^T y^k \end{pmatrix}.$$

2) Mettre à jour le centre approximatif.

Cela signifie générer $(y^{k+1}, s^{k+1}, x^{k+1})$, où y^{k+1} est η -centre de Ω^{k+1} , à partir de (y^k, s^k, x^k) η -centre de Ω^k , en utilisant la méthode qui sera développée dans la section (4.5).

3) Passer à l'itération $k + 1$.

END.

4.4 Réduction de potentiel

Comme nous l'avons fait au chapitre précédent, nous établissons ici une relation entre les fonctions potentielles duale $\mathcal{B}(\Omega)$ et $\mathcal{B}(\Omega^+)$, où Ω^+ est le polytope obtenu en ajoutant à Ω une contrainte passant par un centre analytique approximatif de Ω .

Soit y^k (s^k) un point intérieur de Ω (\mathcal{S}_Ω) et $x^k > 0$ avec $Ax^k = 0$ tels que $\|X^k s^k - e\| \leq \eta$ pour $\eta \in (0, 1)$.

Construisons une coupe passant par y^k . Nous ajoutons donc l'inégalité $a^T y \leq a^T y^k$ à Ω , pour obtenir le nouveau polytope

$$\Omega^+ = \{y : A^T y \leq c, a^T y \leq a^T y^k\}. \quad (4.3)$$

Nous remarquons que cette coupe sectionne Ω de manière plus 'profonde' que celle ajoutée pour obtenir Ω_β^+ au chapitre précédent (cfr 3.7), même si ici, nous travaillons avec des centres analytiques approximatifs. Cela signifie que nous enlevons une plus grande portion de Ω pour obtenir Ω^+ . En effet, le terme $\beta\bar{r}$, où $\beta \in (0, 1)$ et $\bar{r} = \sqrt{a^T(A(S^k)^{-2}A^T)^{-1}a}$, qui apparaissait dans la définition de Ω_β^+ , disparaît ici.

Un théorème semblable au théorème (3.2) peut ainsi être établi.

Théorème 4.1

Soit y^a le centre analytique exact de Ω , et s^a l'écart associé. Posons $\bar{r} = \sqrt{a^T(A(S^a)^{-2}A^T)^{-1}a}$. Alors

$$\mathcal{B}(\Omega^+) \leq \mathcal{B}(\Omega) + \log(\bar{r}) - \kappa \quad (4.4)$$

pour une constante κ dépendant seulement de η .

De plus, si $0 < \eta \leq 1/100$, alors nous avons $\kappa > 0$.

preuve :

Soit y^a et \bar{y}^a les centres analytiques exacts respectifs de Ω et Ω^+ .

Soit $\bar{s}^a = c - A^T \bar{y}^a$ les n premiers écarts aux contraintes de Ω^+ associés au centre analytique \bar{y}^a et $\bar{s}_{n+1}^a = a^T y^k - a^T \bar{y}^a$ le dernier écart associé au même centre analytique. Nous avons

$$\begin{aligned} \bar{s}_{n+1}^a &= a^T(y^k - \bar{y}^a) \\ &= a^T(A(S^a)^{-2}A^T)^{-1}(A(S^a)^{-2}A^T)(y^k - \bar{y}^a) \\ &= a^T(A(S^a)^{-2}A^T)^{-1}A(S^a)^{-2}(A^T y^k - A^T \bar{y}^a) \\ &= a^T(A(S^a)^{-2}A^T)^{-1}A(S^a)^{-2}(-c + A^T y^k + c - A^T \bar{y}^a) \\ &= a^T(A(S^a)^{-2}A^T)^{-1}A(S^a)^{-2}(\bar{s}^a - s^k) \\ &= a^T(A(S^a)^{-2}A^T)^{-1}A(S^a)^{-1}((S^a)^{-1}\bar{s}^a - (S^a)^{-1}s^k) \\ &\leq \|a^T(A(S^a)^{-2}A^T)^{-1}A(S^a)^{-1}\| \|(S^a)^{-1}\bar{s}^a - (S^a)^{-1}s^k\| \\ &= \bar{r} \|(S^a)^{-1}\bar{s}^a - e + e - (S^a)^{-1}s^k\| \\ &\leq \bar{r} (\|(S^a)^{-1}\bar{s}^a - e\| + \|e - (S^a)^{-1}s^k\|) \\ &\leq \bar{r} (\|(S^a)^{-1}\bar{s}^a - e\| + \frac{\eta}{1-2\eta}) \end{aligned} \quad (4.5)$$

où la dernière inégalité vient de la relation (2.23) du théorème (2.1) puisque

$$\|(S^a)^{-1}s^k - e\| = \|S^k(S^a)^{-1}((S^k)^{-1}s^a - e)\|$$

$$\begin{aligned}
&\leq \|S^k(S^a)^{-1}\| \|(S^k)^{-1}s^a - e\| \\
&\leq \lambda_{max}(S^k(S^a)^{-1}) \frac{\eta}{1-\eta} \\
&\leq \frac{1-\eta}{1-2\eta} \frac{\eta}{1-\eta} = \frac{\eta}{1-2\eta}
\end{aligned} \tag{4.6}$$

car $\|(S^k)^{-1}s^a - e\| \leq \frac{\eta}{1-\eta}$ et donc $\frac{s_i^k}{s_i^a} \leq \frac{1-\eta}{1-2\eta}$ $i = 1, \dots, n$, si $\eta < 0.5!$ (ce que l'on peut supposer sans perte de généralité).

Donc,

$$\frac{\exp \mathcal{B}(\Omega^+)}{\bar{r} \exp \mathcal{B}(\Omega)} = \frac{\bar{s}_{n+1}^a}{\bar{r}} \prod_{j=1}^n \frac{\bar{s}_j^a}{s_j^a} \leq (\|(S^a)^{-1}\bar{s}^a - e\| + \frac{\eta}{1-2\eta}) \prod_{j=1}^n \frac{\bar{s}_j^a}{s_j^a}. \tag{4.7}$$

Notons que nous avons aussi $e^T(S^a)^{-1}\bar{s}^a = n$ car $e^T(S^a)^{-1}\bar{s}^a = e^T(X^a)\bar{s}^a = n$ (cfr 3.21).

Dès lors, une borne supérieure pour (4.7) peut être obtenue par le même raisonnement que celui tenu dans la preuve du théorème (3.2) avec $\alpha = (S^a)^{-1}\bar{s}^a = X^a\bar{s}^a > 0$, et où β est remplacé par $\frac{\eta}{1-2\eta}$.

On obtient ainsi une relation similaire à $\psi(\alpha) \leq \frac{4}{\exp(1.5-\beta)}$ obtenue au théorème (3.2), à savoir

$$\begin{aligned}
(\|(S^a)^{-1}\bar{s}^a - e\| + \frac{\eta}{1-2\eta}) \prod_{j=1}^n \frac{\bar{s}_j^a}{s_j^a} &\leq 4 \exp(-1.5 + \frac{\eta}{1-2\eta}) \\
&= 4 \exp(\frac{-3(1-2\eta) + 2\eta}{2(1-2\eta)}) \\
&= 4 \exp(\frac{8\eta - 3}{2(1-2\eta)}).
\end{aligned} \tag{4.8}$$

Par conséquent, (4.7) et (4.8) entraînent

$$\mathcal{B}(\Omega^+) - \mathcal{B}(\Omega) \leq \log(\bar{r}) + \log 4 + \frac{8\eta - 3}{2(1-2\eta)}.$$

Posons $\kappa = -\log 4 - \frac{8\eta-3}{2(1-2\eta)}$. Nous vérifions que si $\eta \in]0, 1/100]$, alors $\kappa > 0$, le résultat souhaité.

CQFD.

4.5 Mise à jour du centre approximatif

Afin d'éclaircir l'étape de mise à jour de l'algorithme décrit précédemment, nous montrons ici comment construire une paire (x^{k+1}, y^{k+1}) telle que $y^{k+1} \in \Omega^+$, $(A, a)x^{k+1} = 0$, $x^{k+1} > 0$, et vérifiant $\|X^{k+1}s^{k+1} - e\| \leq \eta < 1$, à partir d'une paire (x^k, y^k) où $y^k \in \Omega$, $Ax^k = 0$, $x^k > 0$ et vérifiant $\|X^k s^k - e\| \leq \eta < 1$.

En d'autres mots, nous cherchons à construire un η -centre approximatif de Ω^+ , connaissant

un η -centre approximatif de Ω .

Posons

$$r^k = \sqrt{a^T(A(S^k)^{-2}A^T)^{-1}a} \quad (4.9)$$

$$\Delta y = -\left(\frac{\beta}{r^k}\right)(A(S^k)^{-2}A^T)^{-1}a \quad (4.10)$$

$$\Delta s = \left(\frac{\beta}{r^k}\right)A^T(A(S^k)^{-2}A^T)^{-1}a = -A^T\Delta y \quad (4.11)$$

$$\Delta x = -\left(\frac{\beta}{r^k}\right)(S^k)^{-2}A^T(A(S^k)^{-2}A^T)^{-1}a. \quad (4.12)$$

où $\beta \in (0, 1)$. Prenons ensuite $y = y^k + \Delta y$, et

$$x = \begin{pmatrix} x^k + \Delta x \\ \beta/r^k \end{pmatrix} = \begin{pmatrix} x^k - \frac{\beta}{r^k}(S^k)^{-2}A^T(A(S^k)^{-2}A^T)^{-1}a \\ \beta/r^k \end{pmatrix}.$$

Nous vérifions que

$$s := \begin{pmatrix} c - A^T(y^k + \Delta y) \\ a^T y^k - a^T(y^k + \Delta y) \end{pmatrix} = \begin{pmatrix} s^k + \Delta s \\ \beta/r^k \end{pmatrix} = \begin{pmatrix} s^k + \frac{\beta}{r^k}A^T(A(S^k)^{-2}A^T)^{-1}a \\ \beta/r^k \end{pmatrix}.$$

Nous obtenons donc

$$(A, a)x = A(x^k + \Delta x) + \left(\frac{\beta}{r^k}\right)a = Ax^k - \left(\frac{\beta}{r^k}\right)a + \left(\frac{\beta}{r^k}\right)a = Ax^k = 0,$$

et

$$\begin{cases} s^k + \left(\frac{\beta}{r^k}\right)A^T(A(S^k)^{-2}A^T)^{-1}a = S^k(e + p^k) \\ x^k - \left(\frac{\beta}{r^k}\right)(S^k)^{-2}A^T(A(S^k)^{-2}A^T)^{-1}a = (S^k)^{-1}(X^k S^k e - p^k) \end{cases}$$

où $p^k = \left(\frac{\beta}{r^k}\right)(S^k)^{-1}A^T(A(S^k)^{-2}A^T)^{-1}a$, et $\|p^k\| = \beta$.

De plus, nous avons

$$x_j^k s_j^k \geq 1 - \eta \quad j = 1, \dots, n \quad (4.13)$$

puisque $\|X^k s^k - e\| \leq \eta$ par hypothèse. Donc, si l'on choisit η et β tels que $1 - \eta - \beta > 0$, alors

$$s = \begin{pmatrix} S^k(e + p^k) \\ \beta/r^k \end{pmatrix} > 0 \quad \text{et} \quad x = \begin{pmatrix} (S^k)^{-1}(X^k S^k e - p^k) \\ \beta/r^k \end{pmatrix} > 0. \quad (4.14)$$

Finalement,

$$Xs - e = \begin{pmatrix} X^k s^k - e \\ 0 \end{pmatrix} - \begin{pmatrix} (p^k)^2 \\ 1 - \beta^2 \end{pmatrix} + \begin{pmatrix} (X^k S^k - I)p^k \\ 0 \end{pmatrix},$$

où le vecteur $(p^k)^2 = ((p_1^k)^2, (p_2^k)^2, \dots, (p_n^k)^2)^T$.

Dès lors,

$$\begin{aligned}\|Xs - e\| &\leq \|X^k s^k - e\| + \sqrt{\|(p^k)^2\|^2 + (1 - \beta^2)^2} + \|X^k s^k - e\| \|p^k\| \\ &\leq \eta + \sqrt{\beta^4 + (1 - \beta^2)^2} + \eta\beta,\end{aligned}$$

où la deuxième inégalité vient de l'hypothèse $\|X^k s^k - e\| \leq \eta$, et de la définition de p^k .

Posons $\beta = \frac{1}{\sqrt{2}}$ et $\eta = 0.15$. Alors

$$\|Xs - e\| \leq \gamma = 0.15 + \frac{1}{\sqrt{2}} + \frac{0.15}{\sqrt{2}} < 1,$$

et si l'on pose $\eta = 1/100$ (pour être sûr de la réduction du potentiel), nous avons $\gamma = 1/100 + 1.01/\sqrt{2} < 1$. Nous vérifions facilement que $1 - \eta - \beta > 0$ a lieu.

Désormais, ayant trouvé (x, y, s) tel que $\|Xs - e\| \leq \gamma < 1$, nous pouvons appliquer la procédure de Newton duale pure proposée au théorème (2.1 *iv*)) pour générer de manière quadratique une paire (y^{k+1}, s^{k+1}) et $x^{k+1} = x(s^{k+1})$ tels que $(A, a)x^{k+1} = 0$, $x^{k+1} > 0$, $s^{k+1} = (c^T, a^T y^k)^T - (A, a)^T y^{k+1} > 0$, et $\|X^{k+1} s^{k+1} - e\| \leq \eta < \gamma < 1$.

Grâce au théorème (2.1 *iv*)) et aux valeurs de η et β envisagées précédemment (i.e. $\eta = 1/100$ et $\gamma = 1/100 + 1.01/\sqrt{2}$), cette paire (y^{k+1}, s^{k+1}) sera générée en quatre pas de Newton purs puisque $\gamma^{16} \leq \eta$.

Dans l'algorithme destiné à résoudre le problème d'admissibilité convexe, nous avons utilisé cette mise à jour du centre approximatif à chaque itération. Et grâce au théorème (4.1), la séquence des polytopes emboîtés Ω^k générés par l'algorithme satisfait $\mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) + \log(\bar{r}^k) - \kappa$, où $\bar{r}^k = (a_{k+1}^T (A(S^a)^{-2} A^T)^{-1} a_{k+1})^{1/2}$, κ est la constante définie au théorème (4.1), y^a le centre analytique exact de Ω^k , et a_{k+1} le vecteur normal de la coupe ajoutée à l'itération k .

Remarquons que (y^a, s^a) ne sera utilisé que dans l'analyse de la convergence, et pas dans l'algorithme proprement dit puisque nous travaillons avec des centres approximatifs.

4.6 Complexité algorithmique

Pour obtenir la complexité de l'algorithme ACCPM pour le problème d'admissibilité convexe, nous avons besoin de plusieurs résultats intermédiaires.

Soit $\mathcal{B}(\Omega^k)$ la fonction potentielle duale calculée au centre analytique exact y^a de Ω^k . Nous avons

$$\mathcal{B}(\Omega^k) = \sum_{j=1}^{2m+k} (c^k - (A^k)^T y^a)_j.$$

Les relations qui suivent sont vérifiées, à condition que l'algorithme n'ait pas encore pris fin.

$$1) \forall k \geq 0, \quad \mathcal{C} \subset \Omega^k \text{ (car chaque nouvelle coupe est choisie de la sorte)} \quad (4.15)$$

$$2) \forall k \geq 0, \quad \mathcal{B}(\Omega^{k+1}) \leq \mathcal{B}(\Omega^k) + \frac{1}{2} \log(\bar{r}^k)^2 - \kappa, \quad (4.16)$$

où $(\bar{r}^k)^2 = a_{k+1}^T (A^k (S^a)^{-2} (A^k)^T)^{-1} a_{k+1}$, $s^a = c^k - (A^k)^T y^a$, et κ est une constante positive (par le théorème (4.1)).

Lemme 4.6.1

Pour tout $k \geq 0$, et tant que $\mathcal{C} \subset \Omega^k$,

$$\mathcal{B}(\Omega^k) \geq (2m + k) \log \varepsilon. \quad (4.17)$$

preuve :

Par hypothèse, $\mathcal{C} \subset \Omega^k$. Donc, Ω^k contient une boule de dimension pleine et de rayon $\varepsilon > 0$, puisque cela était vrai pour \mathcal{C} . Notons \bar{y} le centre de cette boule.

Alors, $c^k - (A^k)^T \bar{y} \geq \varepsilon e$, et donc

$$\mathcal{B}(\Omega^k) = \sum_{j=1}^{2m+k} \log(c^k - (A^k)^T y^a)_j \geq \sum_{j=1}^{2m+k} \log(c^k - (A^k)^T \bar{y})_j \geq \sum_{j=1}^{2m+k} \log \varepsilon = (2m + k) \log \varepsilon,$$

où la première inégalité est due au fait que y^a , centre analytique de Ω^k , est par définition le point qui maximise la fonction potentielle duale $\mathcal{B}(y, \Omega^k)$ sur Ω^k .

CQFD.

Lemme 4.6.2

Soit $s = c^k - (A^k)^T y$ pour un $y \in \Omega^k$. Alors

- 1) $0 \leq s_j \leq 1 \quad j = 1, 2, \dots, 2m.$
- 2) $0 \leq s_j \leq \sqrt{m} \quad j = 2m + 1, \dots, 2m + k.$

preuve :

Par définition, $s_j = c_j^k - (a_j^k)^T y$.

Lors de l'initialisation, les seules contraintes du système sont celles formant Ω^0 , c'est à dire

$$\begin{cases} y_j \leq 1 & j = 1, \dots, m \\ -y_j \leq 0 & j = 1, \dots, m. \end{cases}$$

Nous obtenons donc $2m$ contraintes initiales, que nous pouvons encore écrire

$$\begin{cases} s_j = 1 - y_j \geq 0 & j = 1, \dots, m \\ s_j = y_{j-m} \geq 0 & j = m, \dots, 2m. \end{cases}$$

Comme $0 \leq y_j \leq 1$ $j = 1, \dots, m$, nous avons $0 \leq s_j \leq 1$ $j = 1, \dots, 2m$.

A chaque itération, nous ajoutons une contrainte, de sorte que pour Ω^k , nous obtenons k contraintes supplémentaires d'indices $j = 2m+1, \dots, 2m+k$ de la forme $s_j = a_{j-2m}^T y^{j-2m-1} - a_{j-2m}^T y \geq 0$. Or,

$$s_j = a_{j-2m}^T (y^{j-2m-1} - y) \leq \|a_{j-2m}\| \|y^{j-2m-1} - y\| = \|y^{j-2m-1} - y\| \leq \sqrt{m} \quad (4.18)$$

où l'égalité vient de l'hypothèse $\|a_j\| = 1 \forall j$, et où la dernière inégalité est due au fait que $0 \leq y^{j-2m-1}, y \leq e$, puisque y^{j-2m-1} et y appartiennent à Ω^0 , le polytope initial. Et donc, $\|y^{j-2m-1} - y\| \leq \|e\| = \sqrt{m}$.

CQFD.

Le lemme (4.6.1) nous montre que tant que $\mathcal{C} \subset \Omega^k$, nous avons $\mathcal{B}(\Omega^k) \geq (2m+k) \log \varepsilon$. Dès lors, si nous prouvons qu'il existe un K tel que $\forall k \geq K$, $\mathcal{B}(\Omega^k) < (2m+k) \log \varepsilon$, l'algorithme se terminera au plus après ces K itérations puisque nous aurons alors $\mathcal{C} \not\subset \Omega^k$. En effet, par la relation (4.15), nous savons que si l'algorithme n'a pas atteint son terme, nous avons $\mathcal{C} \subset \Omega^k$. Dès lors, par contraposition, si $\mathcal{C} \not\subset \Omega^k$, l'algorithme se terminera.

Par (4.16), montrer que $\mathcal{B}(\Omega^k) < (2m+k) \log \varepsilon$ pour un certain k revient à trouver une borne supérieure sur \bar{r}^k . Pour cela, le lemme suivant sera utile.

Lemme 4.6.3

Soit $s = c^k - (A^k)^T y$ pour $y \in \Omega^k$ et $B^0 = 8I$, $B^{k+1} = B^k + \frac{1}{m} a_{k+1} a_{k+1}^T$. Alors,

$$A^k S^{-2} (A^k)^T \succeq B^k \quad \text{i.e.} \quad A^k S^{-2} (A^k)^T - B^k \text{ sdp.} \quad (4.19)$$

preuve :

Soit $Y = \text{diag}(y)$, où $y \in \Omega^k$.

$$\begin{aligned} A^k S^{-2} (A^k)^T &= Y^{-2} + (I - Y)^{-2} - \sum_{j=1}^k \frac{a_j a_j^T}{s_{2m+j}^2} \quad \text{par(3.39) où } n(k) = k. \\ &\succeq Y^{-2} + (I - Y)^{-2} - \frac{1}{m} \sum_{j=1}^k a_j a_j^T \quad \text{par le lemme (4.6.2)} \\ &\succeq 8I + \frac{1}{m} \sum_{j=1}^k a_j a_j^T \quad \text{par la propriété (3.4.2)} \\ &= B^k. \end{aligned}$$

CQFD.

Lemme 4.6.4

Soit $s^a = c^k - (A^k)^T y^a$ le vecteur des écarts associés au centre analytique exact y^a de Ω^k , et $(w^k)^2 = a_{k+1}^T (B^k)^{-1} a_{k+1}$. Alors

$$(w^k)^2 \geq a_{k+1}^T (A^k (S^a)^{-2} (A^k)^T)^{-1} a_{k+1} = (\bar{r}^k)^2. \quad (4.20)$$

preuve :

Grâce au lemme (4.6.3),

$$(w^k)^2 = a_{k+1}^T (B^k)^{-1} a_{k+1} \geq a_{k+1}^T (A^k (S^a)^{-2} (A^k)^T)^{-1} a_{k+1} = (\bar{r}^k)^2. \quad (4.21)$$

CQFD.

Ce lemme signifie que toute borne supérieure sur la série des $(w^k)^2$ sera aussi une borne supérieure sur celle des $(\bar{r}^k)^2$.

Lemme 4.6.5

Nous vérifions que

$$\sum_{j=0}^k (w^j)^2 \leq 2m^2 \log\left(1 + \frac{k+1}{8m^2}\right).$$

preuve :

Notons que

$$\det B^{k+1} = \det\left(B^k + \frac{1}{m} a_{k+1} a_{k+1}^T\right) = \det B^k \left(1 + \frac{a_{k+1}^T (B^k)^{-1} a_{k+1}}{m}\right) = \det B^k \left(1 + \frac{(w^k)^2}{m}\right),$$

où la deuxième égalité vient de la propriété selon laquelle $\det(I + uv^T) = 1 + v^T u$ avec ici $u = \frac{1}{m} (B^k)^{-1} a_{k+1}$, et $v = a_{k+1}$. Donc,

$$\log \det B^{k+1} = \log \det B^k + \log\left(1 + \frac{(w^k)^2}{m}\right). \quad (4.22)$$

Or, on prouve que

$$\frac{(w^k)^2}{m} \leq \frac{1}{8} a_{k+1}^T a_{k+1} = \frac{1}{8}. \quad (4.23)$$

En effet, démontrer (4.23) revient à montrer que $\frac{a_{k+1}^T (B^k)^{-1} a_{k+1}}{m} \leq \frac{1}{8} a_{k+1}^T a_{k+1}$, où $B^{k+1} = B^k + \frac{1}{m} a_{k+1} a_{k+1}^T$, et $B^0 = 8I$.

Or, par la formule de Sherman-Morrison, nous avons que, pour tout $k \geq 1$,

$$\begin{aligned}(B^k)^{-1} &= (B^{k-1} + \frac{1}{m}a_k a_k^T)^{-1} \\ &= (B^{k-1})^{-1} - \frac{[(B^{k-1})^{-1}a_k a_k^T (B^{k-1})^{-1}]/m}{1 + \frac{a_k^T (B^{k-1})^{-1} a_k}{m}},\end{aligned}$$

où $1 + \frac{a_k^T (B^{k-1})^{-1} a_k}{m} > 0$ car B^{k-1} est définie positive par définition (et donc aussi $(B^{k-1})^{-1}$), et où $[(B^{k-1})^{-1}a_k a_k^T (B^{k-1})^{-1}]/m$ est définie positive. Dès lors $(B^k)^{-1} \preceq (B^{k-1})^{-1}$.

Par récurrence, nous obtenons $(B^k)^{-1} \preceq (B^0)^{-1} = \frac{1}{8}I$. Ainsi

$$\frac{a_{k+1}^T (B^k)^{-1} a_{k+1}}{m} \leq \frac{a_{k+1}^T (B^0)^{-1} a_{k+1}}{m} = \frac{\frac{1}{8}a_{k+1}^T a_{k+1}}{m} \leq \frac{1}{8}a_{k+1}^T a_{k+1},$$

et (4.23) est ainsi démontré.

Cette dernière propriété nous permet d'obtenir

$$\log(1 + \frac{(w^k)^2}{m}) \geq \frac{(w^k)^2}{m} - \frac{(\frac{(w^k)^2}{m})^2}{2(1 - \frac{(w^k)^2}{m})} \quad (4.24)$$

$$= \frac{(w^k)^2}{m} (1 - \frac{\frac{(w^k)^2}{m}}{2(1 - \frac{(w^k)^2}{m})}) \quad (4.25)$$

$$\geq \frac{(w^k)^2}{2m} \quad (4.26)$$

où (4.24) vient de la relation $\log(1 + \lambda) \geq (\lambda - \frac{\lambda^2}{2(1-|\lambda|)})$ si $\lambda < 1$, et où (4.26) est obtenue grâce à (4.23).

En effet, démontrer (4.26) revient à montrer que

$$\frac{(w^k)^2}{m} (1 - \frac{\frac{(w^k)^2}{m}}{2(1 - \frac{(w^k)^2}{m})}) = \frac{(w^k)^2}{m} - \frac{\frac{(w^k)^4}{m^2}}{2(1 - \frac{(w^k)^2}{m})} \geq \frac{(w^k)^2}{m} - \frac{(w^k)^2}{2m} = \frac{(w^k)^2}{2m},$$

ou encore que

$$\frac{\frac{(w^k)^4}{m^2}}{2(1 - \frac{(w^k)^2}{m})} \leq \frac{(w^k)^2}{2m}. \quad (4.27)$$

Or, cette dernière inégalité est satisfaite puisque

$$\frac{\frac{(w^k)^4}{m^2}}{2(1 - \frac{(w^k)^2}{m})} = \frac{\frac{(w^k)^2}{m^2}}{2(\frac{1}{(w^k)^2} - \frac{1}{m})} = \frac{(w^k)^2}{2(\frac{m^2}{(w^k)^2} - m)} \leq \frac{(w^k)^2}{16m - 2m} \leq \frac{(w^k)^2}{2m}, \quad (4.28)$$

où la première inégalité est obtenue par (4.23), qui implique que $\frac{m^2}{(w^k)^2} \geq 8m$.

Par (4.22) et (4.26), nous obtenons par récurrence

$$\log \det B^{k+1} \geq \log \det B^0 + \sum_{j=0}^k \frac{(w^j)^2}{2m} \quad (4.29)$$

$$= m \log 8 + \sum_{j=0}^k \frac{(w^j)^2}{2m} \quad (4.30)$$

Or,

$$\frac{1}{m} \log \det B^{k+1} \leq \log \frac{\text{trace} B^{k+1}}{m} = \log \left(8 + \frac{k+1}{m^2} \right), \quad (4.31)$$

car d'une part, $(\det B^{k+1})^{\frac{1}{m}} = (\prod_{i=1}^m \lambda_i)^{\frac{1}{m}} \leq (\sum_{i=1}^m \frac{\lambda_i}{m}) = \frac{\text{trace} B^{k+1}}{m}$ par (2.55), où les λ_i sont les valeurs propres de B , et d'autre part,

$$\begin{aligned} \text{trace} B^{k+1} &= \text{trace} B^0 + \text{trace} \left(\frac{1}{m} \sum_{j=1}^{k+1} a_j a_j^T \right) \\ &= 8m + \frac{1}{m} \sum_{j=1}^{k+1} \text{trace} a_j a_j^T \\ &= 8m + \frac{1}{m} \sum_{j=1}^{k+1} \left(\sum_{i=1}^m a_{ij}^2 \right) \\ &= 8m + \frac{1}{m} \sum_{j=1}^{k+1} \|a_j\|^2 \\ &= 8m + \frac{1}{m} (k+1) \quad \text{puisque } \|a_j\|^2 = 1 \quad \forall j. \end{aligned} \quad (4.32)$$

Donc,

$$\begin{aligned} \sum_{j=0}^k \frac{(w^j)^2}{2m} &\leq \log \det B^{k+1} - m \log 8 \quad \text{par (4.30)} \\ &\leq m \log \left(8 + \frac{k+1}{m^2} \right) - m \log 8, \end{aligned}$$

et

$$\sum_{j=0}^k (w^j)^2 \leq 2m^2 \log \left(1 + \frac{k+1}{8m^2} \right).$$

CQFD.

Théorème 4.2

L'algorithme ACCPM pour le problème d'admissibilité convexe génère une solution admissible de \mathcal{C} au plus en un nombre d'itérations K tel que $\forall k \geq K$,

$$\frac{\varepsilon^2}{m} > \frac{\frac{1}{2} + 2m \log\left(1 + \frac{k+1}{8m^2}\right)}{2m + k + 1} e^{-2\kappa \frac{k+1}{k+1+2m}}. \quad (4.33)$$

preuve :

Par la relation (4.16) et le lemme (4.6.1), tant que l'algorithme ne se termine pas,

$$\begin{aligned} (2m + k + 1) \log \varepsilon &\leq \mathcal{B}(\Omega^{k+1}) \\ &\leq \mathcal{B}(\Omega^0) + \frac{1}{2} \sum_{j=0}^k \log(\bar{r}^j)^2 - (k+1)\kappa \\ &= 2m \log \frac{1}{2} + \frac{1}{2} \sum_{j=0}^k \log(\bar{r}^j)^2 - (k+1)\kappa. \end{aligned}$$

Donc

$$\begin{aligned} \log \varepsilon + \frac{k+1}{2m+k+1} \kappa &\leq \frac{1}{2(2m+k+1)} \left[2m \log \frac{1}{4} + \sum_{j=0}^k \log(\bar{r}^j)^2 \right] \\ &\leq \frac{1}{2} \log \frac{2m \frac{1}{4} + \sum_{j=0}^k (\bar{r}^j)^2}{2m+k+1} \quad (4.34) \\ &\leq \frac{1}{2} \log \frac{\frac{m}{2} + \sum_{j=0}^k (\bar{w}^j)^2}{2m+k+1} \quad \text{par le lemme (4.6.4)} \\ &\leq \frac{1}{2} \log \frac{\frac{m}{2} + 2m^2 \log\left(1 + \frac{k+1}{8m^2}\right)}{2m+k+1} \quad \text{par le lemme (4.6.5)} \end{aligned}$$

ou

$$\frac{\varepsilon^2}{m} \leq \frac{\frac{1}{2} + 2m \log\left(1 + \frac{k+1}{8m^2}\right)}{2m+k+1} e^{-2\kappa \frac{k+1}{k+1+2m}}. \quad (4.35)$$

L'inégalité (4.34) est obtenue par la propriété de concavité du logarithme, à savoir

$$\log\left(\sum \lambda_i a_i\right) \geq \sum \lambda_i \log a_i, \quad \text{si } \lambda_i > 0 \text{ et } \sum \lambda_i = 1. \quad (4.36)$$

Nous vérifions en effet que les coefficients $\frac{2m}{2m+k+1}$ et $\frac{1}{2m+k+1}$ associés respectivement à $\frac{1}{4}$ et $(\bar{r}^j)^2$ sont positifs et satisfont l'égalité $\frac{2m}{2m+k+1} + \sum_{j=0}^k \frac{1}{2m+k+1} = 1$.

Par la contraposée, $\frac{\varepsilon^2}{m} > \frac{\frac{1}{2} + 2m \log\left(1 + \frac{k+1}{8m^2}\right)}{2m+k+1} e^{-2\kappa \frac{k+1}{k+1+2m}}$ entraîne que $\mathcal{B}(\Omega^{k+1}) < (2m+k+1) \log \varepsilon$, et donc l'algorithme s'arrête.

CQFD.

Ce théorème implique que la complexité de l'algorithme ACCPM, estimée en nombre d'appels à l'oracle (i.e. en nombre d'itérations k) est en $\mathcal{O}^*(\frac{m^2}{\varepsilon^2})$, où \mathcal{O}^* signifie que les termes d'ordre inférieur sont ignorés.

Cette affirmation a pu être prouvée dans le cas où $\varepsilon > 1/2$, c'est à dire dans le cas où la boule contenue dans l'ensemble admissible \mathcal{C} a un rayon supérieur à $1/2$. Sous cette hypothèse, l'inégalité (4.33) sera satisfaite si

$$\frac{\varepsilon^2}{m} > \frac{\frac{1}{2} + 2m(1 + \frac{k+1}{8m^2})}{2m + k + 1}, \quad (4.37)$$

puisque $e^{-2\kappa \frac{k+1}{k+1+2m}} < 1$ et $\log(1 + \frac{k+1}{8m^2}) < 1 + \frac{k+1}{8m^2}$.

Or, (4.37) peut encore s'écrire

$$2m + k + 1 > \frac{m}{2\varepsilon^2} + \frac{2m^2}{\varepsilon^2} + \frac{k+1}{4\varepsilon^2} = \frac{2m + 8m^2 + k + 1}{4\varepsilon^2}.$$

On obtient donc

$$k(4\varepsilon^2 - 1) > 2m + 8m^2 + 1 - 8m\varepsilon^2 - 4\varepsilon^2,$$

ou

$$k > \frac{2m + 8m^2 + 1 - 8m\varepsilon^2 - 4\varepsilon^2}{4\varepsilon^2 - 1} \quad (\text{car } \varepsilon > 1/2),$$

où le terme de droite est effectivement en $\mathcal{O}^*(\frac{m^2}{\varepsilon^2})$ si l'on ignore les termes d'ordre inférieur.

Chapitre 5

Algorithme ACCPM générant plusieurs coupes pour le problème d'admissibilité convexe.

5.1 Introduction

Ce chapitre propose une généralisation de l'algorithme ACCPM étudié au chapitre précédent. Il consiste à améliorer la vitesse de convergence en générant plusieurs coupes à chaque itération, au lieu d'une seule comme précédemment.

Après avoir décrit l'algorithme proprement dit, nous établirons une relation entre les fonctions potentielles $\mathcal{B}(\Omega^+)$ et $\mathcal{B}(\Omega)$, où Ω^+ est obtenu en ajoutant plus d'une coupe à Ω , passant toutes par un centre analytique approximatif de Ω . Ensuite, nous bornerons le nombre total de coupes nécessaires pour générer un point admissible de \mathcal{C} , ainsi que le nombre de pas de Newton nécessaires.

Ces deux nombres seront tous deux du même ordre, c'est à dire polynomiaux en m et $1/\varepsilon$, où m est la dimension de l'espace contenant \mathcal{C} , et ε la précision utilisée pour le problème.

5.2 Algorithme ACCPM générant plusieurs coupes

Décrivons l'algorithme ACCPM générant plusieurs coupes à chaque itération. Rappelons que $\mathcal{C} \subset \Omega^0 = \{y \in \mathcal{R}^m : 0 \leq y \leq e\}$ est à l'intérieur non vide, et qu'il existe un oracle qui, pour chaque $y^k \in \mathcal{R}^m$, confirme que $y^k \in \mathcal{C}$, ou génère des coupes $\{y \in \mathcal{R}^m : a_i^T y \leq a_i^T y^k\} \supset \mathcal{C}$, avec $\|a_i\| = 1$.

Algorithme :

BEGIN

Initialisation

Posons $A^0 = (I \ -I) \in \mathcal{R}^{m \times 2m}$, $c^0 = (e^T \ 0)^T \in \mathcal{R}^{2m}$, $y^0 = 0.5e \in \mathcal{R}^m$, $s^0 = c^0 - (A^0)^T y^0 = 0.5e \in \mathcal{R}^{2m}$, et enfin $x^0 = 2e \in \mathcal{R}^{2m}$.

Posons encore $\delta = \frac{1}{4}$, $k = 0$, et $n(0) = 0$.

$k^{\text{ème}}$ itération

Soit y^k un δ -centre de $\Omega^k = \{y \in \mathcal{R}^m : c^k - (A^k)^T y \geq 0\}$, $s^k = c^k - (A^k)^T y^k$ l'écart associé, et x^k le point primal admissible associé, obtenu comme au lemme (3.4.3). On suppose de plus que $(c^k)^T x^k = 2m + n(k)$.

L'oracle vérifie si $y^k \in \mathcal{C}$ ou pas. Si oui, l'algorithme s'arrête et y^k est le point admissible recherché. Sinon, l'algorithme génère des hyperplans $\{y \in \mathcal{R}^m : a_{n(k)+i}^T y \leq a_{n(k)+i}^T y^k\} \supset \mathcal{C}$ pour $i = 1, \dots, \eta(k)$, avec $\|a_{n(k)+i}\| = 1$, et nous posons

$$\Omega^{k+1} = \{y \in \mathcal{R}^m : c^{k+1} - (A^{k+1})^T y \geq 0\}$$

où

$$A^{k+1} = (A^k, a_{n(k)+1}, \dots, a_{n(k)+\eta(k)})$$

$$c^{k+1} = \begin{pmatrix} c^k \\ a_{n(k)+1}^T y^k \\ \vdots \\ a_{n(k)+\eta(k)}^T y^k \end{pmatrix} \in \mathcal{R}^{2m+n(k)+\eta(k)}.$$

La deuxième partie de la $k^{\text{ème}}$ itération consiste à mettre à jour le δ -centre, c'est à dire à calculer un δ -centre de Ω^{k+1} .

On pose ensuite $n(k+1) = n(k) + \eta(k)$, et on passe à l'itération $k+1$ (i.e. $k = k+1$).

END.

Ici, $\eta(k)$ représente le nombre de coupes ajoutées à chaque itération k , et $n(k)$ représente le nombre de coupes au début de l'itération k , sans tenir compte des $2m$ contraintes initiales formant Ω^0 .

Nous développerons la mise à jour du δ -centre plus tard dans notre analyse.

5.3 Réduction de potentiel

Soit $\Omega = \{y \in \mathcal{R}^m : c - A^T y \geq 0\}$, où nous supposons $m \geq 2$, défini par $n (> m)$, borné, à intérieur non vide.

Envisageons tout d'abord le cas où les coupes ajoutées passeraient toutes par le centre analytique exact de Ω . Notons que ceci est inconcevable en pratique, comme nous l'avons déjà fait remarquer au Chapitre 3.

Ajoutons donc η coupes passant par y^a ($\eta \geq 1$), le centre analytique de Ω . Cela signifie ajouter η nouvelles inégalités $a_{n+i}^T y \leq a_{n+i}^T y^a$, $i = 1, \dots, \eta$ à Ω . Nous obtenons le polytope

$$\Omega^+ = \{y \in \mathcal{R}^m : A^T y \leq c, a_{n+i}^T y \leq a_{n+i}^T y^a, i = 1, \dots, \eta\},$$

où nous supposons que $\|a_{n+i}\| = 1$, $i = 1, \dots, \eta$.

Définissons $r_{n+i} = \sqrt{a_{n+i}^T (A(S^a)^{-2} A^T)^{-1} a_{n+i}}$, $i = 1, \dots, \eta$.

Nous pouvons dès lors établir la relation suivante.

Théorème 5.1

Soit \bar{y}^a le centre analytique de Ω^+ et

$$\bar{s}^a = c - A^T \bar{y}^a \in \mathcal{R}^n, \quad \bar{s}_{n+i}^a = a_{n+i}^T y^a - a_{n+i}^T \bar{y}^a, \quad i = 1, \dots, \eta.$$

Alors

$$\begin{aligned} \mathcal{B}(\Omega^+) &= \sum_{j=1}^{n+\eta} \log \bar{s}_j^a \\ &\leq \mathcal{B}(\Omega) + \sum_{i=1}^{\eta} \log r_{n+i} + \frac{\eta}{2} \log\left(\frac{n}{n-1}\right) + (\eta+1) \log(\eta+1) - \frac{\eta(\eta+2)}{\eta+1}. \end{aligned}$$

preuve :

La preuve de ce théorème suit de celle du corollaire (5.3.1), démontré plus loin dans ce chapitre, et qui établit une relation équivalente à celle du théorème (5.1), dans le cas où les coupes sont ajoutées en un centre analytique approximatif.

Le théorème (5.1) découle donc du corollaire (5.3.1) en prenant $\delta = 0$.

CQFD.

Soit maintenant (y^k, s^k) un δ -centre de Ω . Ajoutons η (≥ 1) coupes à Ω , passant par y^k . Nous obtenons le nouveau polytope

$$\Omega^+ = \{y \in \mathcal{R}^m : A^T y \leq c, a_{n+i}^T y \leq a_{n+i}^T y^k, \quad i = 1, \dots, \eta\}, \quad (5.1)$$

où $\text{spdg} \|a_{n+i}\| = 1, i = 1, \dots, \eta$.

Pour obtenir l'équivalent du théorème (5.1) dans le cas où les coupes ajoutées passent par un centre analytique approximatif, les deux lemmes suivants nous seront fort utiles.

Lemme 5.3.1

Soit y un δ -centre de Ω , s l'écart associé, et x le point primal correspondant, obtenu comme au lemme (3.4.3), avec $s = c - A^T y$.

Soit aussi y^a le centre analytique exact de Ω , et s^a l'écart associé. Alors

$$0 \leq \mathcal{B}(\Omega) - \sum_{j=1}^n \log s_j \leq \frac{\delta^2}{2(1-\delta)} \quad (5.2)$$

$$0 \leq -\sum_{j=1}^n \log x_j - \mathcal{B}(\Omega) \leq \frac{\delta^2}{2(1-\delta)} \quad (5.3)$$

et

$$\|S^{-1}s^a - e\| \leq \frac{\delta}{1-\delta} \quad \text{et} \quad \|(S^a)^{-1}s - e\| \leq \frac{\delta}{1-2\delta}. \quad (5.4)$$

preuve :

La première inégalité de (5.2) vient du fait que, par définition, $\mathcal{B}(\Omega) = \max_{s=c-A^T y > 0} (\sum_{j=1}^n \log s_j)$, tandis que la deuxième inégalité de (5.2) a déjà fait l'objet d'une preuve (cfr lemme (3.4.2)).

Pour démontrer la première inégalité de (5.3), il faut se rappeler que x^a est le point qui maximise $\sum_{j=1}^n \log x_j$ sur \mathcal{X}_Ω , et s^a le point qui maximise $\sum_{j=1}^n \log s_j$ sur \mathcal{S}_Ω . Dès lors, (x^a, s^a) est la paire qui maximise $\sum_{j=1}^n \log x_j s_j$ sur $(\mathcal{X}_\Omega \times \mathcal{S}_\Omega)$. Or, au centre analytique, cette fonction s'annule car $x^a s^a = e$. Donc, pour toute paire admissible, $\sum_{j=1}^n \log x_j s_j \leq 0$, et en particulier pour $s = s^a$. Quant à la deuxième inégalité de (5.3), elle peut être démontrée en utilisant une des inégalités établie dans le lemme (3.4.2), à savoir (3.57). En effet, (3.57) implique que

$$-\sum_{j=1}^n \log x_j - \mathcal{B}(\Omega) \leq -\sum_{j=1}^n \log x_j - \sum_{j=1}^n \log s_j \leq \frac{\delta^2}{2(1-\delta)}.$$

Il reste à démontrer les deux inégalités de (5.4).

La première suit du point v au théorème (2.1), puisque $\eta(x, s) = \|Xs - e\| \leq \delta$. Et la seconde a déjà été démontrée au sein de la preuve du théorème (4.1), en (4.6). Remarquons que cette dernière inégalité n'est vérifiée que sous l'hypothèse $\delta < 0.5$ (ce qui peut être sous-entendu spdg).

CQFD.

Lemme 5.3.2

Considérons le problème de maximisation

$$\begin{cases} \max f(\alpha) = \|\alpha - e\|^\eta \prod_{j=1}^n \alpha_j \\ \text{s.c. } e^T \alpha = n, \alpha > 0. \end{cases}$$

Soit α^* la solution optimale. Alors

$$f(\alpha^*) \leq \left(\frac{n}{n-1}\right)^{\frac{\eta}{2}} (\eta+1)^{\eta+1} \exp\left(\frac{-\eta(\eta+2)}{\eta+1}\right). \quad (5.5)$$

preuve :

De la même façon que pour la preuve du théorème (3.2) (cfr (3.23)), $f(\alpha)$ est maximum pour α^* tel que $\text{spdg } \alpha_1^* = \nu > 1$ et $\alpha_2^* = \dots = \alpha_n^* = \frac{n-\nu}{n-1} > 0$, pour $1 < \nu < n$. D'où

$$\begin{aligned} f(\alpha^*) &= \|\alpha^* - e\|^\eta \prod_{j=1}^n \alpha_j^* \\ &= \left(\sum_{j=2}^n \left(\frac{n-\nu}{n-1} - 1\right)^2 + (\nu-1)^2\right)^{\frac{\eta}{2}} \nu \left(\frac{n-\nu}{n-1}\right)^{n-1} \\ &= (\nu-1)^\eta \left(\frac{n}{n-1}\right)^{\frac{\eta}{2}} \nu \left(\frac{n-\nu}{n-1}\right)^{n-1} \\ &= \left(\frac{n}{n-1}\right)^{\frac{\eta}{2}} (\eta+1)^{\eta+1} \left(\frac{\nu-1}{\eta+1}\right)^\eta \frac{\nu}{\eta+1} \left(\frac{n-\nu}{n-1}\right)^{n-1} \end{aligned} \quad (5.6)$$

$$\leq \left(\frac{n}{n-1}\right)^{\frac{\eta}{2}} (\eta+1)^{\eta+1} \left(\frac{n-\frac{\eta}{\eta+1}}{n+\eta}\right)^{n+\eta} \quad (5.7)$$

$$\leq \left(\frac{n}{n-1}\right)^{\frac{\eta}{2}} (\eta+1)^{\eta+1} \exp\left(\frac{-\eta(\eta+2)}{\eta+1}\right),$$

où (5.7) est obtenu par le même raisonnement que celui tenu en (3.24), où η était égal à un, mais où ici $\beta = 0$.

CQFD.

Nous pouvons maintenant établir l'équivalent du théorème (5.1) dans le cas où les coupes ajoutées passent par un centre approximatif.

Corollaire 5.3.1

Soit \bar{y}^a le centre analytique exact de Ω^+ , et $\bar{s}^a = c - A^T \bar{y}^a \in \mathcal{R}^n$, $\bar{s}_{n+i}^a = a_{n+i}^T y^k - a_{n+i}^T \bar{y}^a$, $i = 1, \dots, \eta$. Alors

$$\begin{aligned} \mathcal{B}(\Omega^+) &= \sum_{j=1}^{n+\eta} \log \bar{s}_j^a \\ &\leq \mathcal{B}(\Omega) + \sum_{i=1}^{\eta} \log r_{n+i} + \eta \log \left(\left(\frac{n}{n-1}\right)^{\frac{1}{2}} (\eta+1)^{\frac{\eta+1}{\eta}} \exp\left(\frac{-\eta-2}{\eta+1}\right) + \frac{\delta}{1-2\delta} \right) \end{aligned} \quad (5.8)$$

où $r_{n+i} = \sqrt{a_{n+i}^T (A(S^a)^{-2} A^T)^{-1} a_{n+i}}$ $i = 1, \dots, \eta$.

preuve :

Soit (y^a, s^a) le centre analytique exact de Ω . Pour tout $i = 1, \dots, \eta$, nous avons

$$\bar{s}_{n+i}^a = a_{n+i}^T (y^k - \bar{y}^a) \leq r_{n+i} (\|(S^a)^{-1} \bar{s}^a - e\| + \frac{\delta}{1-2\delta}),$$

de la même façon que dans le théorème (4.1) (cfr (4.5)). Donc,

$$\begin{aligned} \frac{\exp \mathcal{B}(\Omega^+)}{\prod_{i=1}^{\eta} (r_{n+i}) \exp \mathcal{B}(\Omega)} &= \prod_{i=1}^{\eta} \frac{\bar{s}_{n+i}^a}{r_{n+i}} \prod_{j=1}^n \frac{\bar{s}_j^a}{s_j^a} \\ &\leq (\|(S^a)^{-1} \bar{s}^a - e\| + \frac{\delta}{1-2\delta})^{\eta} \prod_{j=1}^n \frac{\bar{s}_j^a}{s_j^a}. \end{aligned}$$

Posons $\alpha = (S^a)^{-1} \bar{s}^a > 0$. Nous vérifions d'une part que $e^T \alpha = n$, car $e^T (S^a)^{-1} \bar{s}^a = e^T (X^a) \bar{s}^a = n$ (cfr (3.21)), et d'autre part que $\prod_{j=1}^n \alpha_j \leq (\sum_{j=1}^n \frac{\alpha_j}{n})^n = 1$. Dès lors,

$$\begin{aligned} \frac{\exp \mathcal{B}(\Omega^+)}{\prod_{i=1}^{\eta} (r_{n+i}) \exp \mathcal{B}(\Omega)} &\leq (\|\alpha - e\| + \frac{\delta}{1-2\delta})^{\eta} \prod_{j=1}^n \alpha_j \\ &= (\|\alpha - e\| (\prod_{j=1}^n \alpha_j)^{\frac{1}{\eta}} + \frac{\delta}{1-2\delta} (\prod_{j=1}^n \alpha_j)^{\frac{1}{\eta}})^{\eta} \\ &\leq (\|\alpha - e\| (\prod_{j=1}^n \alpha_j)^{\frac{1}{\eta}} + \frac{\delta}{1-2\delta})^{\eta} \end{aligned}$$

Par le lemme (5.3.2),

$$\|\alpha - e\|^{\eta} \prod_{j=1}^n \alpha_j \leq (\frac{n}{n-1})^{\frac{\eta}{2}} (\eta+1)^{\eta+1} \exp(\frac{-\eta(\eta+2)}{\eta+1}),$$

et donc,

$$\frac{\exp \mathcal{B}(\Omega^+)}{\exp \mathcal{B}(\Omega)} \leq ((\frac{n}{n-1})^{\frac{1}{2}} (\eta+1)^{\frac{\eta+1}{\eta}} \exp(\frac{-(\eta+2)}{\eta+1}) + \frac{\delta}{1-2\delta})^{\eta} \prod_{i=1}^{\eta} r_{n+i},$$

dont on peut facilement déduire (5.8).

CQFD.

5.4 Complexité algorithmique

Le cheminement suivi ici est très semblable à celui suivi au chapitre précédent dans la section de même nom. Nous allons adapter les résultats obtenus pour le cas de la génération d'une seule coupe par itération à celui de la génération de plusieurs coupes par itération.

Grâce au théorème (5.1), au corollaire (5.3.1), et sachant que $\delta = \frac{1}{4}$, $\eta(k) \geq 1$, et $n \geq 4$ (puisque $A^0 \in \mathcal{R}^{m \times 2m}$, où $m \geq 2$ par hypothèse), il est évident que les relations suivantes sont satisfaites tant que l'algorithme n'a pas pris fin.

$$1) \forall k \geq 0, \quad \mathcal{C} \subset \Omega^k \quad \text{car chaque coupe est choisie de la sorte.} \quad (5.9)$$

$$2) \forall k \geq 0,$$

$$\begin{aligned} \mathcal{B}(\Omega^{k+1}) &\leq \mathcal{B}(\Omega^k) + \sum_{i=1}^{\eta(k)} \log r_{n(k)+i} + \eta(k) \log \left(\left(\frac{4}{3} \right)^{\frac{1}{2}} (\eta(k) + 1)^{\frac{\eta(k)+1}{n(k)}} \exp \left(\frac{-\eta(k) - 2}{\eta(k) + 1} \right) + \frac{\delta}{1 - 2\delta} \right) \\ &\leq \mathcal{B}(\Omega^k) + \sum_{i=1}^{\eta(k)} \log r_{n(k)+i} + \eta(k) \log(\eta(k) + 1) \end{aligned} \quad (5.10)$$

où $r_{n(k)+i} = \sqrt{a_{n(k)+i}^T (A^k (S^a)^{-2} (A^k)^T)^{-1} a_{n(k)+i}}$, avec (y^a, s^a) le centre analytique exact de Ω^k .

Remarquons que le centre analytique exact de Ω^k , bien qu'il n'intervienne pas dans l'algorithme, sera utile dans l'étude de la complexité.

5.4.1 Borne sur le nombre total de coupes

Nous essayons ici d'obtenir une borne supérieure sur le nombre total de coupes à ajouter durant tout l'algorithme pour être sûr de générer un point admissible de \mathcal{C} .

Grâce aux relations (5.9) et (5.10) établies ci-dessus, et par un raisonnement similaire à celui tenu au lemme (4.6.1), nous avons

Lemme 5.4.1

Pour $k \geq 0$ et tant que $\mathcal{C} \subset \Omega^k$,

$$\mathcal{B}(\Omega^k) \geq (2m + n(k)) \log \varepsilon.$$

Le lemme (5.4.1) montre que tant que $\mathcal{C} \subset \Omega^k$, nous avons $\mathcal{B}(\Omega^k) \geq (2m + n(k)) \log \varepsilon$. Dès lors, si l'on prouve qu'il existe un K tel que $\forall k \geq K$, $\mathcal{B}(\Omega^k) < (2m + n(k)) \log \varepsilon$, alors

$\mathcal{C} \not\subset \Omega^k$ ($k \geq K$), et l'algorithme se termine (puisque (5.9) n'a lieu que si l'algorithme n'a pas atteint son terme).

Grâce à (5.10), montrer que $\mathcal{B}(\Omega^k) < (2m + n(k)) \log \varepsilon$ pour un certain k revient à trouver une borne supérieure sur les $r_{n(k)+i}$.

Lemme 5.4.2

Soit $s = c^k - (A^k)^T y \geq 0$ pour tout $y \in \Omega^k$. Alors,

- 1) $0 \leq s_j \leq 1 \quad j = 1, 2, \dots, 2m.$
- 2) $0 \leq s_j \leq \sqrt{m} \quad j = 2m + 1, \dots, 2m + n(k).$

preuve :

La première relation a déjà été prouvée au lemme (4.6.2).

Quant à la preuve de la deuxième relation, elle s'inspire également de celle du lemme (4.6.2), à la différence près que le y^{j-2m-1} de la relation (4.18) doit être remplacé par y^i ($0 \leq i \leq k$), puisque désormais, il existe plusieurs coupes (i.e. plusieurs a_{j-2m}) passant par le centre approximatif y^i d'un même polytope Ω^i .

Nous avons pour $j = 2m + 1, \dots, 2m + n(k)$, et $0 \leq i \leq k$,

$$0 \leq a_{j-2m}^T y^i - a_{j-2m}^T y \leq \|a_{j-2m}\| \|y^i - y\| = \|y^i - y\| \leq \sqrt{m}$$

puisque $y, y^i \in \Omega^0$ (i.e. $-e \leq y^i - y \leq e$).

CQFD.

Lemme 5.4.3

Soit $s = c^k - (A^k)^T y > 0$ pour $y \in \Omega^k$, $B^0 = 8I$, et

$$B^{k+1} = B^k + \frac{1}{m} \sum_{i=1}^{\eta(k)} a_{n(k)+i}^T a_{n(k)+i}.$$

Alors

$$A^k S^{-2} (A^k)^T \succeq B^k \quad \text{i.e.} \quad A^k S^{-2} (A^k)^T - B^k \text{ sdp}$$

preuve :

Nous pouvons utiliser le même raisonnement qu'au lemme (4.6.3), en tenant compte toutefois des petits changements occasionnés par l'ajout simultané de plusieurs coupes. En effet, le nombre de coupes ajoutées à Ω^0 pour obtenir Ω^k est $n(k) \geq k$, et non plus k , puisque nous générons plusieurs coupes par itération. De plus,

$$\begin{aligned}
B^k &= B^{k-1} + \frac{1}{m} \sum_{i=1}^{\eta(k-1)} a_{n(k-1)+i} a_{n(k-1)+i}^T \\
&= B^0 + \frac{1}{m} \left(\sum_{i=1}^{\eta(0)} a_{n(0)+i} a_{n(0)+i}^T + \dots + \sum_{i=1}^{\eta(k-1)} a_{n(k-1)+i} a_{n(k-1)+i}^T \right) \\
&= 8I + \frac{1}{m} \left(\sum_{j=n(0)+1}^{n(0)+\eta(0)} a_j a_j^T + \sum_{j=n(1)+1}^{n(1)+\eta(1)} a_j a_j^T + \dots + \sum_{j=n(k-1)+1}^{n(k-1)+\eta(k-1)} a_j a_j^T \right) \\
&= 8I + \sum_{j=1}^{n(k)} a_j a_j^T \quad \text{car } n(0) = 0 \text{ et } n(k) = n(k-1) + \eta(k-1).
\end{aligned}$$

CQFD.

Lemme 5.4.4

Soit y^a le centre analytique exact de Ω^k , $s^a = c^k - (A^k)^T y^a$ le vecteur des écarts associé et, pour $i = 1, \dots, \eta(k)$, $(w_{n(k)+i})^2 = a_{n(k)+i}^T (B^k)^{-1} a_{n(k)+i}$.

Alors, pour $i = 1, \dots, \eta(k)$,

$$w_{n(k)+i}^2 \geq a_{n(k)+i}^T (A^k (S^a)^{-2} (A^k)^T)^{-1} a_{n(k)+i} := r_{n(k)+i}^2.$$

preuve :

La preuve de ce lemme est une conséquence directe du lemme (5.4.3).

CQFD.

Dès lors, trouver une borne supérieure sur la série des $(r_{n(k)+i})^2$ revient à chercher une borne supérieure sur la série des $(w_{n(k)+i})^2$.

Lemme 5.4.5

Soit $1 \leq \eta(k) \leq \eta \leq m$ pour tout k . Alors

$$\sum_{j=1}^{n(k+1)} (w_j)^2 \leq \frac{18m^2}{15} \log\left(1 + \frac{n(k+1)}{8m^2}\right).$$

preuve :

$$\det B^{k+1} = \det\left(B^k + \frac{1}{m} \sum_{i=1}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T\right) \quad (5.11)$$

$$= \det(B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T + \frac{1}{m} a_{n(k)+1} a_{n(k)+1}^T) \quad (5.12)$$

$$= \det(I + \frac{1}{m} (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T)^{-1} a_{n(k)+1} a_{n(k)+1}^T) \\ \det(B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T) \quad (5.13)$$

$$= (1 + \frac{1}{m} a_{n(k)+1}^T (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T)^{-1} a_{n(k)+1}) \\ \det(B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T) \quad (5.14)$$

$$= (1 + \frac{w^2}{m}) \det(B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T) \quad (5.15)$$

où $w^2 = a_{n(k)+1}^T (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T)^{-1} a_{n(k)+1}$.

L'égalité (5.14) vient de la formule $\det(I + uv^T) = 1 + v^T u$,

où ici $u = \frac{1}{m} (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T)^{-1} a_{n(k)+1}$, et $v = a_{n(k)+1}$.

On vérifie que $w^2 \leq a_{n(k)+1}^T (B^k)^{-1} a_{n(k)+1} = (w_{n(k)+1})^2$, car $B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T \succeq B^k$, puisque le second terme est défini positif.

De plus, nous pouvons montrer que la plus grande valeur propre λ_{max} de la matrice $M = I + \frac{1}{m} \sum_{i=2}^{\eta(k)} (B^k)^{-1/2} a_{n(k)+i} a_{n(k)+i}^T (B^k)^{-1/2}$ est telle que $\lambda_{max} \leq 9/8$. La matrice M étant symétrique par construction, cela revient à montrer qu $\|M\|_2 = \lambda_{max} \leq \frac{9}{8}$, ou encore que $\sup_{\|y\|=1} \|My\|_2 \leq \frac{9}{8}$.

Soit donc $y \in \mathcal{R}^m$ tel que $\|y\| = 1$. Nous avons

$$y^T (I + \frac{1}{m} \sum_{i=2}^{\eta(k)} (B^k)^{-1/2} a_{n(k)+i} a_{n(k)+i}^T (B^k)^{-1/2}) y \\ = \|y\|^2 + \frac{1}{m} \sum_{i=2}^{\eta(k)} (y^T (B^k)^{-1/2} a_{n(k)+i})^2 \\ \leq \|y\|^2 + \frac{1}{m} \sum_{i=2}^{\eta(k)} \|y\|^2 \| (B^k)^{-1/2} a_{n(k)+i} \|^2$$

$$\begin{aligned}
&= 1 + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i}^T (B^k)^{-1} a_{n(k)+i} \quad \text{puisque } \|y\| = 1 \\
&\leq 1 + \frac{1}{m} \sum_{i=2}^{\eta(k)} \frac{1}{8} \tag{5.16}
\end{aligned}$$

$$= 1 + \frac{\eta(k) - 1}{8m} \leq \frac{9}{8} \quad \text{car } \eta(k) < m + 1 \text{ par hypothèse} \tag{5.17}$$

où (5.16) est vérifiée puisque

$$a_{n(k)+i}^T (B^k)^{-1} a_{n(k)+i} = \|(B^k)^{-1/2} a_{n(k)+i}\|^2 \leq \|(B^k)^{-1/2}\|^2 \|a_{n(k)+i}\|^2 \leq \frac{1}{8} \|a_{n(k)+i}\|^2 = \frac{1}{8}, \tag{5.18}$$

vu que $B^k \succeq B^0 = 8I$, et $\|a_{n(k)+i}\| = 1$ par hypothèse .

Ainsi,

$$\begin{aligned}
w^2 &= a_{n(k)+1}^T (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T)^{-1} a_{n(k)+1} \\
&= a_{n(k)+1}^T (B^k)^{-1/2} (I + \frac{1}{m} \sum_{i=2}^{\eta(k)} (B^k)^{-1/2} a_{n(k)+i} a_{n(k)+i}^T (B^k)^{-1/2})^{-1} (B^k)^{-1/2} a_{n(k)+1} \\
&\geq a_{n(k)+1}^T (B^k)^{-1/2} (\frac{8}{9} I) (B^k)^{-1/2} a_{n(k)+1} \\
&= \frac{8}{9} (w_{n(k)+1})^2 \tag{5.19}
\end{aligned}$$

où l'unique inégalité est due au fait que $\lambda_{max}(M) \leq \frac{9}{8}$, ce qui prouve que

$$(I + \frac{1}{m} \sum_{i=2}^{\eta(k)} (B^k)^{-1/2} a_{n(k)+i} a_{n(k)+i}^T (B^k)^{-1/2}) \preceq (9/8) I.$$

Par (5.15) et (5.19), nous obtenons

$$\det B^{k+1} \geq (1 + \frac{8}{9m} (w_{n(k)+1})^2) \det (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T),$$

où, grâce au même développement que celui utilisé pour obtenir (5.15),

$$\det (B^k + \frac{1}{m} \sum_{i=2}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T) = (1 + \frac{(w')^2}{m}) \det (B^k + \frac{1}{m} \sum_{i=3}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T),$$

$$\text{où } (w')^2 = a_{n(k)+2}^T (B^k + \frac{1}{m} \sum_{i=3}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T)^{-1} a_{n(k)+2} \geq \frac{8}{9} (w_{n(k)+2})^2.$$

Dès lors,

$$\det B^{k+1} \geq (1 + \frac{8}{9m} (w_{n(k)+1})^2) (1 + \frac{8}{9m} (w_{n(k)+2})^2) \det (B^k + \frac{1}{m} \sum_{i=3}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T).$$

Par récurrence, nous obtenons

$$\det B^{k+1} \geq \det B^k \prod_{i=1}^{\eta(k)} \left(1 + \frac{8}{9m} (w_{n(k)+i})^2\right),$$

et

$$\log \det B^{k+1} \geq \log \det B^k + \sum_{i=1}^{\eta(k)} \log \left(1 + \frac{8}{9m} (w_{n(k)+i})^2\right). \quad (5.20)$$

Mais, pour $i = 1, \dots, \eta(k)$,

$$\frac{8}{9m} (w_{n(k)+i})^2 = \frac{8}{9m} a_{n(k)+i}^T (B^k)^{-1} a_{n(k)+i} \quad (5.21)$$

$$\leq \frac{8}{9m} \frac{1}{8} \|a_{n(k)+i}\|^2 \quad \text{par (5.18)} \quad (5.22)$$

$$= \frac{1}{9m} \leq \frac{1}{9}, \quad (5.23)$$

d'où

$$\begin{aligned} \log \left(1 + \frac{8}{9m} (w_{n(k)+i})^2\right) &\geq \frac{8}{9m} (w_{n(k)+i})^2 - \frac{\left(\frac{8}{9m} (w_{n(k)+i})^2\right)^2}{2\left(1 - \frac{8}{9m} (w_{n(k)+i})^2\right)} \\ &= \frac{8}{9m} (w_{n(k)+i})^2 \left(1 - \frac{\frac{8}{9m} (w_{n(k)+i})^2}{2\left(1 - \frac{8}{9m} (w_{n(k)+i})^2\right)}\right) \\ &\geq \frac{15(w_{n(k)+i})^2}{18m}, \end{aligned}$$

où la première inégalité vient de la relation

$$\log(1 + \lambda) \geq \lambda - \frac{\lambda^2}{2(1 - |\lambda|)} \quad \forall \lambda < 1, \quad (5.24)$$

et où la deuxième inégalité suit de (5.23), puisque alors, $2\left(1 - \frac{8}{9m} (w_{n(k)+i})^2\right) \geq 2\left(1 - \frac{1}{9}\right) = \frac{16}{9}$, et donc

$$\frac{\frac{8}{9m} (w_{n(k)+i})^2}{2\left(1 - \frac{8}{9m} (w_{n(k)+i})^2\right)} \leq \frac{8}{9m} (w_{n(k)+i})^2 \frac{9}{16} \leq \frac{1}{16m} \quad \text{par (5.23),}$$

ce qui entraîne

$$\begin{aligned} \frac{8}{9m} (w_{n(k)+i})^2 \left(1 - \frac{\frac{8}{9m} (w_{n(k)+i})^2}{2\left(1 - \frac{8}{9m} (w_{n(k)+i})^2\right)}\right) &\geq \frac{8}{9m} (w_{n(k)+i})^2 \left(1 - \frac{1}{16m}\right) \\ &\geq \frac{8}{9m} (w_{n(k)+i})^2 \left(1 - \frac{m}{16m}\right) \\ &= \frac{8}{9m} (w_{n(k)+i})^2 \frac{15}{16} = \frac{15}{18m} (w_{n(k)+i})^2. \end{aligned}$$

Ainsi, (5.20) devient

$$\log \det B^{k+1} \geq \log \det B^k + \sum_{i=1}^{\eta(k)} \frac{15}{18m} (w_{n(k)+i})^2.$$

Par induction sur k , et sachant que $n(k) = n(k-1) + \eta(k-1)$ et $n(0) = 0$, nous obtenons

$$\begin{aligned} \log \det B^{k+1} &\geq \log \det B^0 + \sum_{j=1}^{n(k+1)} \frac{15}{18m} (w_j)^2 \\ &= m \log 8 + \sum_{j=1}^{n(k+1)} \frac{15}{18m} (w_j)^2. \end{aligned} \quad (5.25)$$

Mais

$$\frac{1}{m} \log \det B^{k+1} \leq \log \frac{\text{trace} B^{k+1}}{m} = \log \left(8 + \frac{n(k+1)}{m^2} \right), \quad (5.26)$$

par un raisonnement similaire à celui tenu pour (4.31) au chapitre précédent, mais où $k+1$ est remplacé par $n(k+1)$.

Par (5.26) et (5.25), il suit donc que

$$\sum_{j=1}^{n(k+1)} \frac{15}{18m} (w_j)^2 \leq \log \det B^{k+1} - m \log 8 \leq m \log \left(8 + \frac{n(k+1)}{m^2} \right) - m \log 8,$$

ou

$$\sum_{j=1}^{n(k+1)} (w_j)^2 \leq \frac{18m^2}{15} \log \left(1 + \frac{n(k+1)}{8m^2} \right).$$

CQFD.

Venons en maintenant au résultat principal.

Théorème 5.2

Supposons que le nombre de coupes ajoutées à chaque itération soit compris entre 1 et $\eta (\leq m)$. Alors, l'algorithme ACCPM génère un point admissible de \mathcal{C} et prend fin aussitôt que $n(k+1)$ satisfait

$$\frac{\varepsilon^2}{(\eta+1)^2} > \frac{\frac{m}{2} + \frac{18m^2}{15} \log \left(1 + \frac{n(k+1)}{8m^2} \right)}{2m + n(k+1)}. \quad (5.27)$$

preuve :

Par le lemme (5.4.1), nous savons que tant que $\mathcal{C} \subset \Omega^k$, nous avons $\mathcal{B}(\Omega^k) \geq (2m + n(k)) \log \varepsilon$.

Dès lors, si il existe K tel que $\forall k \geq K$, $\mathcal{B}(\Omega^k) < (2m + n(k)) \log \varepsilon$, \mathcal{C} ne sera plus inclus dans Ω^k , et l'algorithme se terminera puisque par la contraposée, si l'algorithme n'a pas atteint son terme, (5.9) nous assure que $\mathcal{C} \subset \Omega^k$.

Prouvons donc qu'à partir d'une certaine itération, $\mathcal{B}(\Omega^k) < (2m + n(k)) \log \varepsilon$.
Tant que l'algorithme tourne, nous savons grâce à (5.10) et au lemme (5.4.1) que

$$\begin{aligned}
(2m + n(k+1)) \log \varepsilon &\leq \mathcal{B}(\Omega^{k+1}) \\
&\leq \mathcal{B}(\Omega^k) + \frac{1}{2} \sum_{i=1}^{\eta(k)} \log(r_{n(k)+i})^2 + \eta(k) \log(\eta(k) + 1) \\
&\leq \mathcal{B}(\Omega^k) + \frac{1}{2} \sum_{i=1}^{\eta(k)} \log(r_{n(k)+i})^2 + \eta(k) \log(\eta + 1) \\
&\leq \mathcal{B}(\Omega^{k-1}) + \frac{1}{2} \sum_{i=1}^{\eta(k-1)} \log(r_{n(k-1)+i})^2 + \eta(k-1) \log(\eta + 1) + \\
&\quad \frac{1}{2} \sum_{i=1}^{\eta(k)} \log(r_{n(k)+i})^2 + \eta(k) \log(\eta + 1) \\
&\leq \mathcal{B}(\Omega^0) + \frac{1}{2} \sum_{j=1}^{n(k+1)} \log r_j^2 + n(k+1) \log(\eta + 1) \\
&= 2m \log \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{n(k+1)} \log r_j^2 + n(k+1) \log(\eta + 1)
\end{aligned}$$

puisque $\mathcal{B}(\Omega^0) = \sum_{j=1}^{2m} \log(c^0 - (A^0)^T y^0)_j = m \log(1 - \frac{1}{2}) + m \log \frac{1}{2} = 2m \log \frac{1}{2}$.
Donc,

$$\log \varepsilon - \frac{n(k+1) \log(\eta + 1)}{2m + n(k+1)} \leq \frac{1}{2(2m + n(k+1))} [2m \log(\frac{1}{2})^2 + \sum_{j=1}^{n(k+1)} \log r_j^2],$$

et comme $2m + n(k+1) > n(k+1)$,

$$\begin{aligned}
\log \varepsilon - \log(\eta + 1) &\leq \frac{1}{2(2m + n(k+1))} [2m \log(\frac{1}{4}) + \sum_{j=1}^{n(k+1)} \log r_j^2] \\
&\leq \frac{1}{2} \log \frac{2m \frac{1}{4} + \sum_{j=1}^{n(k+1)} r_j^2}{2m + n(k+1)} \tag{5.28}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{2} \log \frac{\frac{m}{2} + \sum_{j=1}^{n(k+1)} w_j^2}{2m + n(k+1)} \quad \text{par le lemme (5.4.4)} \\
&\leq \frac{1}{2} \log \frac{\frac{m}{2} + \frac{18m^2}{15} \log(1 + \frac{n(k+1)}{8m^2})}{2m + n(k+1)} \quad \text{par le lemme (5.4.5)} \quad (5.29)
\end{aligned}$$

où l'inégalité (5.28) est obtenue en utilisant la propriété de concavité de la fonction logarithme, à savoir

$$\log(\sum \lambda_i a_i) \geq \sum \lambda_i \log a_i, \quad \text{si } \lambda_i \geq 0 \text{ et } \sum \lambda_i = 1. \quad (5.30)$$

On vérifie en effet que les coefficients $\frac{2m}{2m+n(k+1)}$ associé à $\frac{1}{4}$, et les coefficients $\frac{1}{2m+n(k+1)}$ associés aux r_j^2 pour $j = 1, \dots, n(k+1)$ sont positifs, et satisfont

$$\frac{2m}{2m + n(k+1)} + \sum_{j=1}^{n(k+1)} \frac{1}{2m + n(k+1)} = 1.$$

(5.29) peut encore s'exprimer

$$\frac{\varepsilon^2}{(\eta + 1)^2} \leq \frac{\frac{m}{2} + \frac{18m^2}{15} \log(1 + \frac{n(k+1)}{8m^2})}{2m + n(k+1)}.$$

Par la contraposée, nous obtenons (5.27).

CQFD.

Le théorème précédent implique que la complexité de l'algorithme ACCPM avec génération de plusieurs coupes, estimée par le nombre de coupes ajoutées durant l'algorithme, est en $\mathcal{O}^*(\frac{\eta^2 m^2}{\varepsilon^2})$, où la notation \mathcal{O}^* signifie que les termes d'ordre inférieur sont ignorés.

Cette affirmation a pu être prouvée dans le cas où $\varepsilon > \frac{\sqrt{3}}{2\sqrt{5}}(\eta + 1)$. En effet, sous cette hypothèse, l'inégalité (5.27) est satisfaite si

$$\frac{\varepsilon^2}{(\eta + 1)^2} > \frac{\frac{m}{2} + \frac{18m^2}{15} (1 + \frac{n(k+1)}{8m^2})}{2m + n(k+1)},$$

c'est-à-dire si

$$2m + n(k+1) > \frac{60m(\eta + 1)^2 + 144m^2(\eta + 1)^2 + 18(\eta + 1)^2 n(k+1)}{120\varepsilon^2},$$

ou encore

$$n(k+1)[120\varepsilon^2 - 18(\eta + 1)^2] > 60m(\eta + 1)^2 + 144m^2(\eta + 1)^2 - 240\varepsilon^2 m.$$

Dès lors, (5.27) sera satisfaite si

$$n(k+1) > \frac{60m(\eta + 1)^2 + 144m^2(\eta + 1)^2 - 240\varepsilon^2 m}{120\varepsilon^2 - 18(\eta + 1)^2} \quad (\text{car } \varepsilon^2 > \frac{18}{120}(\eta + 1)^2),$$

où le terme de droite est effectivement en $\mathcal{O}^*(\frac{\eta^2 m^2}{\varepsilon^2})$ si l'on ignore les termes d'ordre inférieur.

5.4.2 Borne sur le nombre total de pas de Newton

Nous essayons ici de trouver une borne supérieure sur le nombre total de pas de Newton à effectuer durant tout l'algorithme pour être sûr de générer un point admissible de \mathcal{C} .

A chaque itération de l'algorithme décrit à la section (5.2), nous devons calculer un centre analytique approximatif de Ω^{k+1} . Cela peut se faire en utilisant un algorithme de points intérieurs basé sur la méthode de Newton. Dans ce cas, le nombre de pas de Newton pour générer ce centre approximatif est borné par un nombre en $\mathcal{O}(\sqrt{2m + n(k+1)})$.

Soit k^* le premier indice k pour lequel $n(k^* + 1)$ satisfait (5.27). Alors, la complexité de l'algorithme ACCPM, estimée par le nombre total de pas de Newton, est bornée par un nombre en

$$\mathcal{O}^*((k^* + 1)\frac{\eta m}{\varepsilon}) \leq \mathcal{O}^*(n(k^* + 1)\frac{\eta m}{\varepsilon}) \leq \mathcal{O}^*\left(\frac{\eta^3 m^3}{\varepsilon^3}\right). \quad (5.31)$$

En effet, le premier terme est obtenu en multipliant le nombre d'itérations par une approximation du nombre de pas de Newton à chaque itération. $\frac{m\eta}{\varepsilon}$ est effectivement un ordre de grandeur pour le nombre de pas de Newton par itération, puisque ce dernier est de l'ordre de $\mathcal{O}(\sqrt{2m + n(k+1)}) \leq \mathcal{O}(\sqrt{2m + n(k^* + 1)})$, où $n(k^* + 1)$ est lui en $\mathcal{O}^*(\frac{\eta^2 m^2}{\varepsilon^2})$ (grâce au théorème (5.2)). En ignorant le terme $2m$ sous la racine, il en résulte que le nombre de pas de Newton par itération est en $\mathcal{O}^*(\frac{\eta m}{\varepsilon})$.

Quant aux inégalités de (5.31), elles sont immédiates puisque nous savons que $k^* + 1 \leq n(k^* + 1)$ et que, par le théorème (5.2), $n(k^* + 1)$ est en $\mathcal{O}^*(\frac{\eta^2 m^2}{\varepsilon^2})$.

Cette section nous montrera cependant qu'un algorithme de réduction de potentiel permet d'effectuer un nombre moindre de pas de Newton. Pour obtenir ce nombre, nous devons établir plusieurs résultats intermédiaires.

Au polytope $\Omega^k = \{y \in \mathcal{R}^m : c^k - (A^k)^T y \geq 0\}$, nous associons la fonction potentielle primale

$$\mathcal{P}(x, \Omega^k) := (2m + n(k)) \log(c^k)^T x - \sum_{j=1}^{2m+n(k)} \log x_j,$$

pour tout $x > 0$ tel que $A^k x = 0$, ainsi que la fonction potentielle primale-duale

$$\psi_{2m+n(k)}(x, s) = (2m + n(k)) \log(s^T x) - \sum_{j=1}^{2m+n(k)} \log x_j s_j,$$

où $s = c^k - (A^k)^T y$.

Nous savons, depuis que nous avons défini la fonction potentielle primale-duale (cfr 2.13), qu'elle vérifie la propriété $\psi_{2m+n(k)}(x, s) \geq (2m + n(k)) \log(2m + n(k))$.

De plus,

$$\psi_{2m+n(k)}(x, s^a) = (2m + n(k)) \log(c^k)^T x - \sum_{j=1}^{2m+n(k)} \log x_j s_j^a \quad \text{car } A^k x = 0$$

$$\begin{aligned}
&= \mathcal{P}(x, \Omega^k) - \sum_{j=1}^{2m+n(k)} \log s_j^a \\
&= \mathcal{P}(x, \Omega^k) - \mathcal{B}(\Omega^k).
\end{aligned} \tag{5.32}$$

Pour obtenir une borne sur le nombre de pas de Newton de l'algorithme, il faut tout d'abord éclaircir l'étape de mise à jour du δ -centre. En réalité, la mise à jour s'effectue comme suit.

Partant d'un point $\hat{x} > 0$ tel que $A^{k+1}\hat{x} = 0$, nous appliquons un algorithme de réduction de potentiel primal pour réduire $\mathcal{P}(x, \Omega^{k+1})$, et s'approcher ainsi du centre analytique de Ω^{k+1} . Cet algorithme est détaillé dans l'article [8], et utilise une démarche similaire à celle suivie dans l'algorithme de Karmarkar décrit au Chapitre 2. Cependant, ce nouvel algorithme ne passe pas par l'intermédiaire d'une fonction \mathcal{P}' , et démarre en un point \hat{x} tel que $A^{k+1}\hat{x} = 0$, $\hat{x} > 0$, mais ne vérifiant pas spécialement la contrainte de normalisation $e^T \hat{x} = n$. Les itérés engendrés par cet algorithme permettent la réduction de $\mathcal{P}(x, \Omega^{k+1})$ à chaque itération, et cela jusqu'à ce qu'un point $(y, s) \in (\Omega^{k+1} \times \mathcal{S}_{\Omega^{k+1}})$ satisfaisant $\|\frac{1}{\mu} Xs - e\| \leq \delta$ soit généré, où $\mu = \frac{(c^{k+1})^T x}{2m+n(k+1)} = \frac{s^T x}{2m+n(k+1)}$, $s = c^{k+1} - (A^{k+1})^T y$, et $\delta < 1$. Remarquons qu'un tel point était également engendré par l'algorithme de Karmarkar. Ensuite, si (x, y, s) vérifie cette inégalité, nous posons $(y^{k+1}, s^{k+1}) = (y, s)$, et $x^{k+1} = \frac{x}{\mu}$. Notons que $(c^{k+1})^T x^{k+1} = 2m + n(k+1)$ par définition de μ . Dès lors, (x^{k+1}, y^{k+1}) est la paire correspondant à un δ -centre approximatif de Ω^{k+1} .

Le nombre de pas de Newton nécessaires à l'itération k pour générer un δ -centre de Ω^{k+1} , noté $\mathcal{N}(k)$, est donc borné par

$$\begin{aligned}
\gamma \mathcal{N}(k) &\leq \mathcal{P}(\hat{x}, \Omega^{k+1}) - \mathcal{P}(\bar{x}^a, \Omega^{k+1}) \\
&= \mathcal{P}(\hat{x}, \Omega^{k+1}) - \mathcal{B}(\Omega^{k+1}) + \mathcal{B}(\Omega^{k+1}) - \mathcal{P}(\bar{x}^a, \Omega^{k+1}) \\
&= \mathcal{P}(\hat{x}, \Omega^{k+1}) - \mathcal{B}(\Omega^{k+1}) - \psi_{2m+n(k+1)}(\bar{x}^a, \bar{s}^a) \\
&= \mathcal{P}(\hat{x}, \Omega^{k+1}) - \mathcal{B}(\Omega^{k+1}) - (2m + n(k+1)) \log(2m + n(k+1))
\end{aligned} \tag{5.33}$$

où (\bar{x}^a, \bar{s}^a) est la paire centrale associée à Ω^{k+1} , et où γ est une constante dans $(0, 1)$.

Dès lors, pour calculer un δ -centre de Ω^{k+1} , où

$$\begin{aligned}
\Omega^{k+1} &= \{y \in \mathcal{R}^m : c^{k+1} - (A^{k+1})^T y \geq 0\}, \\
A^{k+1} &= (A^k, a_{n(k)+1}, \dots, a_{n(k)+\eta(k)}) \\
c^{k+1} &= \begin{pmatrix} c^k \\ a_{n(k)+1}^T y^k \\ \vdots \\ a_{n(k)+\eta(k)}^T y^k \end{pmatrix} \in \mathcal{R}^{2m+n(k+1)},
\end{aligned}$$

à partir d'une paire (x^k, y^k) associée à un δ -centre de Ω^k , i.e. $s^k = c^k - (A^k)^T y^k > 0$, $A^k x^k = 0$, $(c^k)^T x^k = 2m + n(k)$, et $\|X^k s^k - e\| \leq \delta$, nous devons, comme annoncé, appliquer l'algorithme de réduction de potentiel primal à partir d'un point $\hat{x} \in \mathcal{R}^{2m+n(k+1)}$ tel que $\hat{x} > 0$,

et $A^{k+1}\hat{x} = 0$.

Pour obtenir ce point, nous procédons de la manière suivante.

Posons

$$a^k = \sum_{i=1}^{\eta(k)} a_{n(k)+i} \quad (5.34)$$

$$r^k = \sqrt{(a^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k} \quad (5.35)$$

$$\Delta x = -\left(\frac{\beta}{r^k}\right) X^k (S^k)^{-1} (A^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k \quad (5.36)$$

où $\beta := \frac{1}{2}$.

Nous définissons alors $\hat{x} = (x^k + \Delta x, \frac{\beta}{r^k}, \dots, \frac{\beta}{r^k})^T \in \mathcal{R}^{2m+n(k+1)}$. Nous vérifions que

$$\begin{aligned} A^{k+1}\hat{x} &= A^k(x^k + \Delta x) + \frac{\beta}{r^k} a^k = A^k x^k - \frac{\beta}{r^k} a^k + \frac{\beta}{r^k} a^k = 0, \\ (c^{k+1})^T \hat{x} &= (c^k)^T x^k + (c^k)^T \Delta x + \sum_{i=1}^{\eta(k)} (a_{n(k)+i}^T y^k)^T \frac{\beta}{r^k} \\ &= (c^k)^T x^k + (c^k - (A^k)^T y^k + (A^k)^T y^k)^T \Delta x + \frac{\beta}{r^k} (y^k)^T a^k \\ &= (c^k)^T x^k + (s^k)^T \Delta x + (y^k)^T A^k \Delta x + \frac{\beta}{r^k} (y^k)^T a^k \\ &= (c^k)^T x^k + (y^k)^T A^k \Delta x + \frac{\beta}{r^k} (y^k)^T a^k \quad \text{car } A^k x^k = 0 \\ &= (c^k)^T x^k - \frac{\beta}{r^k} (y^k)^T a^k + \frac{\beta}{r^k} (y^k)^T a^k \quad \text{par définition de } \Delta x \\ &= (c^k)^T x^k = 2m + n(k) \end{aligned}$$

et

$$\begin{aligned} x^k + \Delta x &= x^k - \left(\frac{\beta}{r^k}\right) X^k (S^k)^{-1} (A^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k \\ &= X^k (e - p^k) \end{aligned} \quad (5.37)$$

où $p^k = \left(\frac{\beta}{r^k}\right) (S^k)^{-1} (A^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k$.

Notons que $\|p^k\| \leq \frac{\beta}{\sqrt{1-\delta}}$ puisque

$$\begin{aligned} \|p^k\|^2 &= \left(\frac{\beta}{r^k}\right)^2 \|(S^k)^{-1} (A^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k\|^2 \\ &\leq \left(\frac{\beta}{r^k}\right)^2 \frac{1}{1-\delta} (a^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k \\ &= \frac{\beta^2}{1-\delta} \end{aligned} \quad (5.38)$$

où l'inégalité est due au fait que $A^k(S^k)^{-2}(A^k)^T \preceq \frac{1}{1-\delta}A^kX^k(S^k)^{-1}(A^k)^T$ puisque, par hypothèse, $\|X^k s^k - e\| \leq \delta$, et donc, $x_i^k \geq (1-\delta)(s_i^k)^{-1}$ et $X^k \succeq (1-\delta)(S^k)^{-1}$, où X^k et S^k sont diagonales.

Comme $\|p^k\| \leq \frac{\beta}{\sqrt{1-\delta}}$, et par (5.37), nous vérifions qu'avec $\beta = \frac{1}{2}$ et $\delta = \frac{1}{4}$ comme c'est le cas dans l'algorithme, le nouveau point \hat{x} est tel que $\hat{x} > 0$ puisque $\|p^k\| = \frac{\sqrt{3}}{3} < 1$.

Ayant défini le point initial \hat{x} de l'algorithme de réduction de potentiel, nous pouvons dès lors l'appliquer, jusqu'à l'obtention d'un δ -centre de Ω^{k+1} .

Remarquons que la valeur de la fonction potentielle primale au point initial \hat{x} satisfait

$$\begin{aligned}
\mathcal{P}(\hat{x}, \Omega^{k+1}) &= (2m + n(k+1)) \log(c^{k+1})^T \hat{x} - \sum_{j=1}^{2m+n(k+1)} \log \hat{x}_j \\
&= (2m + n(k+1)) \log(2m + n(k)) - \sum_{j=1}^{2m+n(k)} \log \hat{x}_j - \sum_{j=2m+n(k)+1}^{2m+n(k+1)} \log \hat{x}_j \\
&= (2m + n(k+1)) \log(2m + n(k)) - \sum_{j=1}^{2m+n(k)} \log \hat{x}_j - \eta(k) \log\left(\frac{\beta}{r^k}\right) \\
&= (2m + n(k+1)) \log(2m + n(k)) - \eta(k) \log\left(\frac{\beta}{r^k}\right) - \sum_{j=1}^{2m+n(k)} \log x_j \\
&\quad - \sum_{j=1}^{2m+n(k)} \log(1 - p_j^k) \quad \text{par définition de } \hat{x} \\
&\leq (2m + n(k+1)) \log(2m + n(k)) - \eta(k) \log\left(\frac{\beta}{r^k}\right) + \mathcal{B}(\Omega^k) \\
&\quad + \frac{\delta^2}{2(1-\delta)} - \sum_{j=1}^{2m+n(k)} \log(1 - p_j^k) \quad \text{par le lemme (5.3.1)} \\
&\leq (2m + n(k+1)) \log(2m + n(k)) - \eta(k) \log\left(\frac{\beta}{r^k}\right) + \mathcal{B}(\Omega^k) \\
&\quad + \frac{\delta^2}{2(1-\delta)} + \frac{\beta^2}{2(1-\delta-\beta\sqrt{1-\delta})} \quad \text{car } \|p^k\| \leq \frac{\beta}{\sqrt{1-\delta}} \\
&\leq (2m + n(k+1)) \log(2m + n(k)) + \eta(k) \log r^k \\
&\quad - \eta(k) \log \beta + \mathcal{B}(\Omega^k) + 1 \quad \text{puisque } \delta = 1/4 \text{ et } \beta = 1/2.
\end{aligned} \tag{5.39}$$

Par les relations (5.33) et (5.39), le nombre de pas de Newton de l'algorithme de réduction de potentiel à l'itération k , démarrant en \hat{x} , est donc borné par

$$\gamma \mathcal{N}(k) \leq \mathcal{P}(\hat{x}, \Omega^{k+1}) - \mathcal{B}(\Omega^{k+1}) - (2m + n(k+1)) \log(2m + n(k+1))$$

$$\begin{aligned}
&\leq (2m + n(k+1)) \log(2m + n(k)) - (2m + n(k+1)) \log(2m + n(k+1)) \\
&\quad + 1 + \eta(k) \log r^k - \eta(k) \log \beta + \mathcal{B}(\Omega^k) - \mathcal{B}(\Omega^{k+1}) \\
&\leq (2m + n(k+1)) \log(2m + n(k+1) - 1) - (2m + n(k+1)) \log(2m + n(k+1)) \\
&\quad + 1 + \eta(k) \log r^k - \eta(k) \log \beta + \mathcal{B}(\Omega^k) - \mathcal{B}(\Omega^{k+1}) \quad \text{car } n(k+1) \geq n(k) + 1 \\
&= (2m + n(k+1)) \log\left(1 - \frac{1}{2m + n(k+1)}\right) \\
&\quad + 1 + \eta(k) \log r^k - \eta(k) \log \beta + \mathcal{B}(\Omega^k) - \mathcal{B}(\Omega^{k+1}) \\
&\leq \eta(k) \log r^k - \eta(k) \log \beta + \mathcal{B}(\Omega^k) - \mathcal{B}(\Omega^{k+1})
\end{aligned}$$

car $\log\left(1 - \frac{1}{2m+n(k+1)}\right) \leq -\frac{1}{2m+n(k+1)}$.

Donc, le nombre total de pas de Newton pour l'algorithme complet, c'est à dire jusqu'à l'itération k^* du théorème (5.2) est

$$\begin{aligned}
\gamma \sum_{j=0}^{k^*} \mathcal{N}(j) &\leq \sum_{j=0}^{k^*} (\eta(j) \log(r^j) - \eta(j) \log \beta) + \mathcal{B}(\Omega^0) - \mathcal{B}(\Omega^{k^*+1}) \\
&= \frac{1}{2} \sum_{j=0}^{k^*} (\eta(j) \log(r^j)^2) + \mathcal{B}(\Omega^0) - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log \beta \\
&= \frac{1}{2} \sum_{j=0}^{k^*} (\eta(j) \log(r^j)^2) + 2m \log \frac{1}{2} - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log \beta
\end{aligned}$$

Posons $(w^k)^2 = (a^k)^T (B^k)^{-1} a^k$, où B^k est définie au lemme (5.4.3), et $a^k = \sum_{i=1}^{\eta(k)} a_{n(k)+i}$.

Alors, par le lemme (5.4.3),

$$\begin{aligned}
(w^k)^2 &\geq (a^k)^T (A^k (S^k)^{-2} (A^k)^T)^{-1} a^k \\
&\geq (1 - \delta) (a^k)^T (A^k X^k (S^k)^{-1} (A^k)^T)^{-1} a^k \\
&= (1 - \delta) (r^k)^2 \quad \text{par (5.35),}
\end{aligned} \tag{5.40}$$

où (5.40) est obtenu par la même justification que (5.38). Finalement, le nombre de pas de Newton de l'algorithme ACCPM est borné par

$$\begin{aligned}
\gamma \sum_{j=0}^{k^*} \mathcal{N}(j) &\leq \frac{1}{2} \sum_{j=0}^{k^*} (\eta(j) \log(w^j)^2) + \frac{1}{2} \sum_{j=0}^{k^*} (\eta(j) \log(1 - \delta)^{-1}) \\
&\quad + 2m \log \frac{1}{2} - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log \beta \\
&= \frac{1}{2} \sum_{j=0}^{k^*} (\eta(j) \log(w^j)^2) - \sum_{j=0}^{k^*} (\eta(j) \log \sqrt{1 - \delta})
\end{aligned} \tag{5.41}$$

$$+2m \log \frac{1}{2} - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log \beta \quad (5.42)$$

$$= \frac{1}{2} \sum_{j=0}^{k^*} (\eta(j) \log(w^j)^2) - n(k^* + 1) \log(\beta \sqrt{1 - \delta}) \\ + 2m \log \frac{1}{2} - \mathcal{B}(\Omega^{k^*+1}) \quad (5.43)$$

$$= \frac{1}{2} \left(\sum_{j=0}^{k^*} (\eta(j) \log(w^j)^2) + 2m \log \frac{1}{4} \right) \\ - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log(\beta \sqrt{1 - \delta}) \quad (5.44)$$

$$\leq \frac{2m + n(k^* + 1)}{2} \log \frac{\frac{2m}{4} + \sum_{j=0}^{k^*} \eta(j)(w^j)^2}{2m + n(k^* + 1)} \\ - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log(\beta \sqrt{1 - \delta}) \quad (5.45)$$

$$\leq \frac{2m + n(k^* + 1)}{2} \log \frac{\frac{m}{2} + \eta \sum_{j=0}^{k^*} (w^j)^2}{2m + n(k^* + 1)} \\ - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log(\beta \sqrt{1 - \delta}) \quad (5.46)$$

où (5.45) est obtenue par la propriété de concavité du logarithme (5.30), applicable ici puisque la somme des coefficients (positifs) de $1/4$ et des $(w^j)^2$ satisfait

$$\frac{2m}{2m + n(k^* + 1)} + \sum_{j=0}^{k^*} \frac{\eta(j)}{2m + n(k^* + 1)} = 1.$$

Partant de cette dernière relation, nous établissons le lemme suivant.

Lemme 5.4.6

Soit $1 \leq \eta(k) \leq \eta \leq m \forall k$. Alors,

$$\sum_{j=0}^k (w^j)^2 \leq \frac{14}{13} \eta m^2 \log\left(1 + \frac{n(k+1)}{8m^2}\right).$$

preuve :

Remarquons tout d'abord que

$$\sum_{i=1}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T \succeq \frac{1}{\eta(k)} a^k (a^k)^T. \quad (5.47)$$

En effet, cela revient à montrer que pour tout $x \neq 0$,

$$x^T \left(\sum_{i=1}^{\eta(k)} a_{n(k)+i} a_{n(k)+i}^T \right) x \geq \frac{1}{\eta(k)} x^T \left(\sum_{i=1}^{\eta(k)} a_{n(k)+i} \right) \left(\sum_{i=1}^{\eta(k)} a_{n(k)+i} \right)^T x,$$

ou encore

$$\sum_{i=1}^{\eta(k)} (a_{n(k)+i}^T x)^2 \geq \frac{1}{\eta(k)} \left(\sum_{i=1}^{\eta(k)} a_{n(k)+i}^T x \right)^2. \quad (5.48)$$

En posant $u_i = a_{n(k)+i}^T x$ $i = 1, \dots, \eta(k)$, (5.48) peut encore s'écrire

$$\sum_{i=1}^{\eta(k)} u_i^2 \geq \frac{1}{\eta(k)} \left(\sum_{i=1}^{\eta(k)} u_i \right)^2 \quad \text{ou} \quad \|u\|_2 \geq \frac{1}{\sqrt{\eta(k)}} \left| \sum_{i=1}^{\eta(k)} u_i \right|.$$

Cette dernière inégalité est toujours satisfaite puisque $\frac{1}{\sqrt{\eta(k)}} \left| \sum_{i=1}^{\eta(k)} u_i \right| \leq \frac{1}{\sqrt{\eta(k)}} \|u\|_1 \leq \|u\|_2$.

Grâce à (5.47) et par définition de B^{k+1} (cfr lemme (5.4.3)), $B^{k+1} \succeq B^k + \frac{1}{\eta(k)m} a^k (a^k)^T$.
Dès lors,

$$\det B^{k+1} \geq \det \left(B^k + \frac{1}{\eta(k)m} a^k (a^k)^T \right) \quad (5.49)$$

$$= \det B^k \det \left(I + \frac{1}{\eta(k)m} (B^k)^{-1} a^k (a^k)^T \right) \quad (5.50)$$

$$= \det B^k \left(1 + \frac{(w^k)^2}{\eta(k)m} \right), \quad (5.51)$$

où (5.51) vient de la formule $\det(I + uv^T) = (1 + v^T u)$, avec $u = \frac{1}{\eta(k)m} (B^k)^{-1} a^k$, et $v = a^k$.
De plus

$$\frac{(w^k)^2}{\eta(k)m} \leq \frac{1}{8\eta(k)m} \|a^k\|^2 \leq \frac{\eta(k)}{8\eta(k)m} \left(\sum_{i=1}^{\eta(k)} \|a_{n(k)+i}\|^2 \right) = \frac{\eta(k)}{8m} \leq \frac{1}{8}, \quad (5.52)$$

puisque d'une part $(w^k)^2 = \|(B^k)^{-1/2} a^k\|^2 \leq \|(B^k)^{-1/2}\|^2 \|a^k\|^2 \leq \frac{1}{8} \|a^k\|^2$ car $B^k \succeq B^0$, et d'autre part, $1 \leq \eta(k) \leq m$.

Par conséquent,

$$\begin{aligned} \log \left(1 + \frac{(w^k)^2}{\eta(k)m} \right) &\geq \frac{(w^k)^2}{\eta(k)m} - \frac{\left(\frac{(w^k)^2}{\eta(k)m} \right)^2}{2 \left(1 - \frac{(w^k)^2}{\eta(k)m} \right)} \quad \text{par (5.24)} \\ &= \frac{(w^k)^2}{\eta(k)m} \left(1 - \frac{\frac{(w^k)^2}{\eta(k)m}}{2 \left(1 - \frac{(w^k)^2}{\eta(k)m} \right)} \right) \\ &\geq \frac{13}{14} \frac{(w^k)^2}{\eta(k)m}, \end{aligned} \quad (5.53)$$

puisque (5.52) implique $2 \left(1 - \frac{(w^k)^2}{\eta(k)m} \right) \geq 2 \left(1 - \frac{1}{8} \right) = \frac{14}{8}$, et donc

$$1 - \frac{\frac{(w^k)^2}{\eta(k)m}}{2 \left(1 - \frac{(w^k)^2}{\eta(k)m} \right)} \geq 1 - \frac{(w^k)^2}{\eta(k)m} \frac{8}{14}$$

$$\begin{aligned}
&\geq 1 - \frac{\eta}{8m} \frac{8}{14} \quad \text{par (5.52)} \\
&\geq 1 - \frac{1}{14} \quad \text{car } \eta \leq m \\
&= \frac{13}{14}
\end{aligned}$$

Donc, il suit de (5.51) et (5.53) que

$$\log \det B^{k+1} \geq \log \det B^k + \frac{13(w^k)^2}{14\eta(k)m},$$

et par induction sur k , nous obtenons

$$\begin{aligned}
\log \det B^{k+1} &\geq \log \det B^0 + \sum_{j=0}^k \frac{13(w^j)^2}{14\eta(j)m} \\
&= m \log 8 + \sum_{j=0}^k \frac{13(w^j)^2}{14\eta(j)m}
\end{aligned}$$

Mais de nouveau, $\frac{1}{m} \log \det B^{k+1} \leq \log \frac{\text{trace} B^{k+1}}{m} = \log(8 + \frac{n(k+1)}{m^2})$ (cfr 5.26), et donc

$$\sum_{j=0}^k \frac{13(w^j)^2}{14\eta(j)m} \leq \log \det B^{k+1} - m \log 8 \leq m \log(1 + \frac{n(k+1)}{8m^2}), \quad (5.54)$$

ou

$$\sum_{j=0}^k (w^j)^2 \leq \frac{14}{13} \eta m^2 \log(1 + \frac{n(k+1)}{8m^2}) \quad (\text{car } \eta(j) \leq \eta).$$

CQFD.

Grâce à la relation (5.46) et au lemme (5.4.6), le nombre total de pas de Newton pour l'algorithme ACCPM complet est borné par

$$\gamma \sum_{j=0}^{k^*} \mathcal{N}(j) \leq \frac{2m + n(k^* + 1)}{2} \log \frac{\frac{2m}{4} + \frac{14\eta^2 m^2}{13} \log(1 + \frac{n(k^*+1)}{8m^2})}{2m + n(k^* + 1)} - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log(\beta\sqrt{1 - \delta}) \quad (5.55)$$

$$\leq \frac{2m + n(k^* + 1)}{2} \log \frac{\frac{\eta^2 m}{2} + \frac{18\eta^2 m^2}{15} \log(1 + \frac{n(k^*+1)}{8m^2})}{2m + n(k^* + 1)} - \mathcal{B}(\Omega^{k^*+1}) - n(k^* + 1) \log(\beta\sqrt{1 - \delta}) \quad (5.56)$$

$$\leq \frac{2m + n(k^* + 1)}{2} \log \frac{\frac{\eta^2 m}{2} + \frac{18\eta^2 m^2}{15} \log(1 + \frac{n(k^*+1)}{8m^2})}{2m + n(k^* + 1)} - (2m + n(k^* + 1)) \log \varepsilon - n(k^* + 1) \log(\beta\sqrt{1 - \delta}) \quad \text{par le lemme (5.4.1)} \quad (5.57)$$

$$= \frac{2m + n(k^* + 1)}{2} \log \frac{\frac{\eta^2 m}{2} + \frac{18\eta^2 m^2}{15} \log(1 + \frac{n(k^*+1)}{8m^2})}{\varepsilon^2(2m + n(k^* + 1))}$$

$$-n(k^* + 1) \log(\beta\sqrt{1 - \delta}) \quad (5.58)$$

$$\leq -n(k^* + 1) \log(\beta\sqrt{1 - \delta}) \quad (5.59)$$

$$= n(k^* + 1) \log \frac{1}{\beta\sqrt{1 - \delta}} \quad (5.60)$$

où (5.59) est due au fait que k^* satisfait la condition du théorème (5.2), i.e.

$$\frac{\frac{\eta^2 m}{2} + \frac{18\eta^2 m^2}{15} \log(1 + \frac{n(k^*+1)}{8m^2})}{\varepsilon^2(2m + n(k^* + 1))} \leq \frac{\frac{(\eta+1)^2 m}{2} + \frac{18(\eta+1)^2 m^2}{15} \log(1 + \frac{n(k^*+1)}{8m^2})}{\varepsilon^2(2m + n(k^* + 1))} < 1.$$

Nous obtenons finalement le théorème suivant.

Théorème 5.3

Supposons que le nombre de coupes $\eta(k)$ ajoutées à chaque itération soit situé entre 1 et η ($\leq m$). Alors, le nombre total de pas de Newton effectués par l'algorithme est borné par $n(k^ + 1) \leq \mathcal{O}^*(\frac{\eta^2 m^2}{\varepsilon^2})$ (cfr théorème (5.2)), i.e. la même borne que pour le nombre de coupes ajoutées jusqu'à la fin de l'algorithme.*

En effet, (5.60) peut être borné par $n(k^* + 1)$ puisque sous les conditions de l'algorithme (i.e. $\delta = 1/4$ et $\beta = 1/2$), on vérifie que $\log(1/(\beta\sqrt{1 - \delta})) < 1$.

Pour conclure, remarquons que le nombre d'itérations de cet algorithme est borné par le nombres de coupes ajoutées (i.e. $k^* + 1 \leq n(k^* + 1)$) puisque nous générons plus d'une coupe par itération.

Comme le nombre de coupes total est borné par $\mathcal{O}^*(\frac{\eta^2 m^2}{\varepsilon^2})$, et que le nombre de coupes par itération ne dépasse pas η , le nombre d'itérations est borné par $\mathcal{O}^*(\frac{\eta m^2}{\varepsilon^2})$. Grâce au théorème (5.3), nous remarquons encore que le nombre de pas de Newton par itération sera donc borné par η .

Conclusion

Dans ce mémoire, nous avons montré en quoi une méthode de réduction de potentiel utilisant des centres analytiques et des plans coupants (méthode ACCPM) pouvait être utile pour la résolution d'un problème d'admissibilité convexe. Ce problème consistait à trouver un point y dans \mathcal{C} convexe, ou à conclure que \mathcal{C} est vide.

Nous avons d'abord exposé l'idée de la méthode pour ensuite l'appliquer au problème d'admissibilité linéaire et convexe. Nous avons enfin présenté une stratégie pour améliorer la vitesse de convergence de cette méthode dans le cas convexe. Celle-ci consistait à effectuer plusieurs coupes au lieu d'une seule à chaque itération.

Comme annoncé dans l'introduction, la méthode ACCPM peut être utilisée pour résoudre d'autres problèmes que celui d'admissibilité convexe, notamment celui de programmation convexe.

Par manque de temps, nous nous sommes limités dans ce mémoire à l'étude du problème d'admissibilité convexe, qui à lui seul, fait déjà l'objet de nombreux articles.

Bibliographie

- [1] J.-L. GOFFIN, A. HAURIE AND J.-P. VIAL. Decomposition and non differentiable optimization with the projective algorithm. July, 1989.
- [2] J.-L. GOFFIN, J.-P. VIAL. Shallow, deep and very deep cuts in the analytic center cutting plane method. LOGILAB Technical Report 96.1, May 1996.
- [3] J.-L. GOFFIN, ZQ LUO, Y.YE. On the complexity of a column generation algorithm for convex and quasiconvex feasibility problems. *Large Scale Optimization : State of the Art*, pages 182–191, 1994. W.W. Hager et al. (eds).
- [4] J.-L. GOFFIN, ZQ LUO, Y.YE. Complexity analysis of an interior cutting plane method for convex feasibility problems. *SIAM optimization*, 6(3) :638–652, August 1996.
- [5] J.-P. VIAL. A generic path-following algorithm with a sliding constraint and its application to linear programming and the computation of analytic centers. Technical report, LOGILAB-HEC, Management Studies, February 1996.
- [6] O.DU MERLE. *Points intérieurs et plans coupants : mise en oeuvre et développement d'une méthode pour l'optimisation convexe et la programmation linéaire structurée de grande taille*. PhD thesis, Genève, 1995.
- [7] S. BOYD, L. VANDENBERGHE. Convex optimization. Course reader, Stanford University 1996.
- [8] Y.YE. An $o(n^3)$ potential reduction algorithm for linear programming. *Math.Programming*, 50 :239–258, 1991.
- [9] Y.YE. A potential reduction algorithm allowing column generation. *SIAM optimization*, 2(1) :7–20, February 1992.
- [10] Y.YE. Complexity analysis of the analytic center cutting plane method that uses multiple cuts. *Math.Programming*, 78 :85–104, 1997.
- [11] Y.YE. *Interior point algorithms Theory and Analysis*. J.Wiley and Sons, New York, 1997.