

# Chapitre XIII

## LES MÉTHODES DE POINT INTÉRIEUR <sup>1</sup>

### XIII.1 Introduction

Nous débutons par un rappel de la formulation standard d'un problème d'optimisation<sup>2</sup> linéaire et donnons un bref aperçu des différences principales entre l'algorithme simplexe, l'approche traditionnelle pour résoudre un problème linéaire et les méthodes de point intérieur, une famille d'algorithmes développés plus récemment à laquelle ce chapitre est consacré.

#### XIII.1.1 Programmation linéaire

Pour rappel (cf. chapitre I), un problème de programmation linéaire consiste à optimiser une fonction objectif linéaire  $z$  dépendant de  $n$  variables de décision tout en vérifiant un ensemble de contraintes linéaires (égalités et/ou inégalités), ce que l'on peut écrire mathématiquement comme (en utilisant des notations matricielles)

$$\min_{x \in \mathbb{R}^n} z = cx \quad \text{tel que} \quad \begin{cases} T^E x = d^E \\ T^I x \geq d^I \end{cases}, \quad (\text{XIII.1})$$

où le vecteur colonne  $x$  contient les  $n$  variables de décision, le vecteur ligne  $c$  définit la fonction objectif et les paires  $(T^E, d^E)$  et  $(T^I, d^I)$  définissent les  $m^E$  contraintes égalités et les  $m^I$  contraintes d'inégalité. Le vecteur colonne  $x$  et le vecteur ligne  $c$  sont de taille  $n$ , les vecteurs colonnes  $d^E$  et  $d^I$  sont de taille  $m^E$  et  $m^I$  tandis que les matrices  $T^E$  et  $T^I$  sont de dimensions respectives  $m^E \times n$  et  $m^I \times n$ .

Cependant, de nombreux problèmes linéaires utilisent des contraintes d'inégalité plus simples, par exemple des contraintes de positivité ( $x \geq 0$ ) ou des bornes inférieures et supérieures ( $l \leq x \leq u$ ). La *forme standard* pour la programmation linéaire est une forme particulière de problème d'optimisation linéaire que nous utiliserons pour le développement théorique des méthodes de point intérieur :

$$\min_{x \in \mathbb{R}^n} cx \quad \text{tel que} \quad \begin{cases} Tx = d \\ x \geq 0 \end{cases}. \quad (\text{XIII.2})$$

Les seules contraintes d'inégalités dans ce problème sont des contraintes de positivité, et celles-ci concernent la *totalité* des variables, ce qui signifie qu'aucune variable n'est libre (on a donc que  $m^I$  est égal à  $n$ ,  $T^I$  est la matrice identité et  $d^I$  est le vecteur nul). Rappelons que n'importe quel

<sup>1</sup>Ce chapitre a été rédigé par François Glineur, Dr. en Sciences Appliquées, Chercheur FNRS durant cinq ans dans le Service de Mathématique et de Recherche Opérationnelle de la Faculté Polytechnique de Mons et Chargé de Cours à l'Université Catholique de Louvain depuis le 1er octobre 2002.

<sup>2</sup>Nous utiliserons indistinctement les expressions *programmation* linéaire et *optimisation* linéaire.

problème d'optimisation linéaire exprimé sous la forme générale (XIII.1) admet un programme équivalent sous forme standard, que l'on peut obtenir en ajoutant et/ou supprimant des variables et/ou des contraintes (le fait pour un problème d'être *équivalent* à un problème donné signifie que la résolution du problème transformé permet de trouver la solution du problème original).

### XIII.1.2 L'algorithme simplexe

Pour rappel (cf. chapitre II), l'ensemble de tous les vecteurs  $x$  satisfaisant les contraintes de (XIII.2) forme un polyèdre dans  $\mathbb{R}^n$ . Puisque la fonction objectif est linéaire, chaque hyperplan de  $\mathbb{R}^n$  orthogonal au vecteur  $c$  constitue un ensemble de solutions possédant le même coût, ce qui entraîne que la solution optimale doit se trouver sur l'un des sommets du polyèdre (il est également possible qu'une face entière du polyèdre soit optimale, ou qu'il n'existe aucune solution optimale parce que les contraintes du problème sont contradictoires ou encore qu'il n'existe pas de solution optimale finie parce que le polyèdre n'est pas borné dans la direction de la fonction objectif).

L'idée principale de l'algorithme simplexe consiste à explorer de façon itérative les sommets de ce polyèdre, en se déplaçant de sommet en sommet tout en améliorant la valeur de la fonction objectif (en utilisant une caractérisation algébrique des sommets appelée *base*). Quand il n'est plus possible d'effectuer un tel déplacement sans détériorer la valeur de la fonction objectif, on arrête l'algorithme : cela se produit nécessairement après un nombre fini de déplacements et le sommet résultant est bien optimal [4].

### XIII.1.3 Un premier coup d'œil aux méthodes de point intérieur

Nous sommes à présent en mesure de fournir une première description des méthodes de point intérieur. Contrairement à l'algorithme simplexe qui utilise les sommets, ces méthodes sont initialisées avec un point de départ situé à *l'intérieur* de l'ensemble des solutions admissibles. En nous basant sur la formulation standard (XIII.2), nous appellerons domaine admissible  $\mathcal{P}^3$  la l'ensemble des vecteurs  $x$  qui satisfont les contraintes, c'est-à-dire

$$\mathcal{P} = \{x \in \mathbb{R}^n \mid Tx = d \text{ et } x \geq 0\} ,$$

tandis que l'ensemble associé  $\mathcal{P}^+$  sera le sous-ensemble des points  $\mathcal{P}$  qui satisfont de manière stricte les contraintes de positivité

$$\mathcal{P}^+ = \{x \in \mathbb{R}^n \mid Tx = d \text{ et } x > 0\} .$$

$\mathcal{P}^+$  sera dénommé *domaine strictement admissible* et ses éléments seront appelés *points* (ou *solutions*) *strictement admissibles*.

Les méthodes de point intérieur sont des méthodes *itératives* qui calculent une suite d'itérés appartenant à  $\mathcal{P}^+$  et convergeant vers une solution optimale, à l'opposé de l'algorithme simplexe qui obtient une solution optimale *exacte* après un nombre fini d'itérations. Les itérés des méthodes de point intérieur tendent donc vers une solution optimale sans jamais l'atteindre (puisque les solutions optimales n'appartiennent pas à  $\mathcal{P}^+$  mais bien à  $\mathcal{P} \setminus \mathcal{P}^+$ , la frontière du domaine admissible). Ceci n'est un inconvénient qu'en apparence, puisque

- la plupart du temps, une solution approchée (avec par exemple une précision relative de  $10^{-8}$ ) se révèle tout à fait suffisante,
- il existe une procédure d'*arrondi* capable de convertir une solution intérieure quasi-optimale en un sommet optimal exact (voir par exemple [23, chapitre 7]).

Une autre différence dans le comportement de ces méthodes se produit lorsqu'une face entière de  $\mathcal{P}$  est optimale : les méthodes de point intérieur convergent alors vers un point situé à l'intérieur de cette face, tandis que l'algorithme simplexe aboutira sur l'un de ses sommets.

<sup>3</sup>Dans ce chapitre, la notation  $\mathcal{P}$  remplace la notation  $\mathcal{D}$  utilisée précédemment.

La dernière différence que nous mentionnerons à ce stade concerne la complexité algorithmique de ces méthodes. Tandis que l'algorithme simplexe peut nécessiter un nombre d'itérations qui croît exponentiellement avec la taille du problème à résoudre [11], les méthodes de point intérieur ne nécessitent qu'un nombre d'itérations borné de façon polynomiale par la taille du problème pour atteindre une précision donnée. Cette propriété fondamentale est sans nul doute principalement responsable de l'énorme quantité de recherches effectuées dans le domaine des méthodes de point intérieur pour la programmation linéaire.

### XIII.1.4 Un bref compte-rendu historique

En complément au tableau figurant à la fin de l'avant-propos, soulignons—sans être exhaustif—les étapes importantes qui ont jalonné le développement des méthodes de point intérieur.

#### Premier pas des méthodes de point intérieur et complexité polynomiale.

- 1955. K. R. Frisch propose une méthode *barrière* pour résoudre des problèmes non linéaires [7].
- 1967. P. Huard introduit la *méthode des centres* pour résoudre des problèmes possédant des contraintes non linéaires [8].
- 1968. A. V. Fiacco et G. P. McCormick développent la méthode barrière pour la programmation non linéaire convexe [6].
- 1978. L. G. Khachiyan applique la méthode de l'*ellipsoïde* (initialement introduite par N. Shor en 1970 [17]) à la programmation linéaire et prouve que sa complexité algorithmique est de type polynomial [10], et donc que le problème (LP) est de classe  $\mathcal{P}$  (cf. chapitre **\*\*XII\*\***).

Il est important de réaliser que la méthode barrière fut développée pour résoudre des problèmes non linéaires. Bien qu'en principe également applicable à la programmation linéaire, ces auteurs ne la considéraient pas comme une alternative viable à l'algorithme simplexe. Il faut également signaler que la meilleure complexité algorithmique de la méthode de l'ellipsoïde comparée à l'algorithme simplexe ne présente qu'un intérêt théorique, car la méthode de l'ellipsoïde s'avérera être très lente en pratique (l'algorithme simplexe n'exhibe un nombre exponentiel d'itérations que sur quelques problèmes spécifiquement créés à cet effet et résout les problèmes réels beaucoup plus rapidement, tandis que la méthode de l'ellipsoïde nécessite un nombre d'itérations pratiquement toujours égal à sa borne polynomiale de pire cas, qui s'avère être nettement plus coûteuse en temps de calculs que le comportement typique de l'algorithme simplexe).

#### La révolution des méthodes de point intérieur.

- 1984. N. Karmarkar découvre une méthode de point intérieur de complexité polynomiale plus efficace en pratique que celle de Khachiyan. Il annonce également des performances supérieures à celles de l'algorithme simplexe [9].
- 1994. Y. Nesterov et A. Nemirovski publient une étude sur les méthodes de point intérieur polynomiales appliquées à la programmation convexe [15].
- 2000. Depuis l'annonce de Karmarkar, plus de 3000 articles de recherche portant sur les méthodes de point intérieur ont été publiés par la communauté scientifique, ainsi que quelques ouvrages de références (voir par exemple [23, 16, 24]). Les recherches se dirigent à présent vers la programmation non linéaire, plus particulièrement dans le domaine de l'optimisation convexe.

Avec le recul, il faut modérer l'affirmation de Karmarkar : sa méthode n'était en définitive pas véritablement supérieure aux meilleures implémentations de l'algorithme simplexe disponibles à l'époque, surtout pour la résolution de problèmes de petite taille. Néanmoins, elle a eu le mérite de susciter de nombreuses recherches dans ce domaine. Signalons également pour l'anecdote que

la méthode de Khachiyan n'est pas à proprement parler la première méthode de résolution polynomiale pour la programmation linéaire. En effet, il a été montré a posteriori [2] que la méthode barrière de Fiacco et McCormick jouissait d'une complexité algorithmique de type polynomial lorsqu'elle était appliquée à la programmation linéaire. Ainsi, on disposait dès 1968 - mais sans en être conscient - d'une méthode polynomiale pour la programmation linéaire.

## XIII.2 Concepts de base

Passons à présent en revue les différents concepts nécessaires à une compréhension correcte des méthodes de point intérieur. Nous débutons avec la notion très utile de dualité pour la programmation linéaire, introduite au chapitre **\*\*CHAPDUAL\*\***.

### XIII.2.1 Dualité

Voici à nouveau le problème d'optimisation linéaire dans sa forme standard

$$\min_{x \in \mathbb{R}^n} z = cx \quad \text{tel que} \quad \begin{cases} Tx = d \\ x \geq 0 \end{cases} . \quad (\text{LP})$$

A l'aide des mêmes données (à savoir la matrice  $T$  et les vecteurs  $d$  et  $c$ ), il est possible de décrire un autre programme linéaire, en introduisant un vecteur ligne  $u$  de  $m$  variables duales :

$$\max_{u \in \mathbb{R}^m} w = ud \quad \text{tel que} \quad \begin{cases} uT \leq c \\ u \text{ est libre} \end{cases} . \quad (\text{LD}')$$

Ce problème est fortement lié à (LP) et sera pour cette raison appelé le problème *dual* de (LP) (qui sera quant à lui baptisé problème *primal*). On vérifie aisément que ce problème peut également être formulé comme

$$\max_{u \in \mathbb{R}^m, s \in \mathbb{R}^n} w = ud \quad \text{tel que} \quad \begin{cases} uT + s = c \\ s \geq 0 \text{ et } u \text{ libre} \end{cases} . \quad (\text{LD})$$

L'introduction de ce vecteur ligne supplémentaire  $s$ , contenant les  $n$  variables d'écart du dual, permettra de simplifier grandement nos notations, et nous utiliserons donc principalement cette seconde formulation du dual. Définissons également les domaines admissible  $\mathcal{D}^4$  et strictement admissible  $\mathcal{D}^+$  pour le problème dual, de façon tout à fait similaire aux domaines  $\mathcal{P}$  et  $\mathcal{P}^+$  pour le primal

$$\begin{aligned} \mathcal{D} &= \{(u, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid uT + s = c \text{ et } s \geq 0\} , \\ \mathcal{D}^+ &= \{(u, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid uT + s = c \text{ et } s > 0\} . \end{aligned}$$

Dans ce qui suit, nous faisons l'hypothèse que la matrice  $T$  est de rang maximum, et donc que ses lignes sont linéairement indépendantes. Cette hypothèse peut être faite sans aucune perte de généralité : si une des lignes de  $T$  dépendait linéairement d'un ensemble d'autres lignes, la contrainte associée serait soit redondante (et peut donc être ignorée sans conséquence), soit impossible à satisfaire (conduisant alors à un problème insoluble), en fonction des composantes du vecteur  $d$  rassemblant les membres de droite des contraintes.

L'équation  $uT + s = c$  induit alors une correspondance bijective entre les variables  $u$  et  $s$  sur le domaine admissible du problème dual. Dans la suite, nous emploierons donc indistinctement  $(u, s)$ ,  $u$  ou  $s$  en tant que variables duales.

Rappelons à présent les propriétés fondamentales de dualité (cf. chapitre IV.2) :

<sup>4</sup>Dans ce chapitre, la notation  $\mathcal{D}$  représente le domaine admissible du dual et non plus du primal comme dans les chapitres précédents.

- Si  $x$  est admissible pour (LP) et  $(u, s)$  est admissible pour (LD), nous avons l'inégalité  $ud \leq cx$  entre les fonctions objectifs. En d'autres termes, toute solution admissible pour (LD) fournit une borne inférieure au problème (LP) et toute solution admissible pour (LP) fournit une borne supérieure pour (LD). C'est la propriété de dualité *faible*. La quantité  $cx - ud$ , toujours positive ou nulle, sera dénommée *saut de dualité*, on montre qu'elle est en fait égale à  $sx$  (en effet, en utilisant le fait que  $d = Tx$  et  $c = uT + s$ , on trouve  $cx - ud = uTx + sx - uTx = sx$ ).
- Les solutions  $x$  et  $(u, s)$  sont optimales pour les problèmes (LP) et (LD) si et seulement si le saut de dualité correspondant est égal à zéro. C'est la propriété de dualité *forte*. Ceci implique que dans la situation où les deux problèmes admettent une solution optimale, les valeurs optimales des deux fonctions objectifs sont égales. Dans ce cas, puisque  $sx = 0$  et  $x \geq 0, s \geq 0$ , chacun des produits  $x_i s_i$  doit être égal à zéro, ce qui entraîne qu'au moins une des deux variables de chaque paire  $\{x_i, s_i\}$  est égal à zéro pour chaque  $i$  (c'est le *théorème des écarts complémentaires*).
- Des deux théorèmes de dualité, on déduit aisément que les deux problèmes (LP) et (LD) doivent obligatoirement se trouver dans l'une des trois situations suivantes :
  1. Les deux problèmes admettent une solution optimale finie (et la dualité forte garantit qu'ils partagent alors la même valeur optimale de la fonction objectif).
  2. Un des problèmes n'est pas borné (sa valeur optimale est infinie) tandis que l'autre ne possède aucune solution admissible (son domaine admissible est vide). Ceci est une conséquence de la dualité faible.
  3. Aucun des problèmes n'admet de solution admissible.

C'est le *théorème fondamental de dualité* pour la programmation linéaire.

### XIII.2.2 Conditions d'optimalité

La théorie de Karush-Kuhn-Tucker (KKT) permet d'écrire des conditions nécessaires d'optimalité pour tout problème d'optimisation contraint possédant une fonction objectif différentiable (voir par exemple [12, chapitre 7]). De plus, lorsque le problème est convexe, ces conditions sont également suffisantes, ce qui est le cas pour l'optimisation linéaire. En considérant le problème (LP), l'écriture de ces conditions mène au système (celui-ci peut également aisément être obtenu à l'aide de la théorie des multiplicateurs de Lagrange)

$$x \text{ est optimal pour (LP)} \Leftrightarrow \exists (y, t) \quad \text{tel que} \quad \begin{cases} Tx = d \\ T^T y + t = c^T \\ x_i t_i = 0 \quad \forall i \\ x \text{ et } t \geq 0 \end{cases} \quad \text{(KKT)}$$

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

$$x \text{ est optimal pour (LP)} \Leftrightarrow \exists (u, s) \quad \text{tel que} \quad \begin{cases} Tx = d \\ uT + s = c \\ x_i s_i = 0 \quad \forall i \\ x \text{ et } s \geq 0 \end{cases} \quad \text{(KKT)}$$

Finalement, en utilisant les définitions de  $\mathcal{P}$  et  $\mathcal{D}$  et le fait que lorsqu'un vecteur ligne  $a$  et un vecteur colonne  $b$  de même longueur sont positifs,

$$a_i b_i = 0 \quad \forall i \Leftrightarrow \sum_i a_i b_i = 0 \Leftrightarrow ab = 0,$$

on retrouve le théorème des écarts complémentaires

$$x \text{ est optimal pour (LP)} \Leftrightarrow \exists (u, s) \text{ tel que } \begin{cases} x \in \mathcal{P} \\ (u, s) \in \mathcal{D} \\ sx = 0 \end{cases} .$$

Comme nous l'avons indiqué au chapitre \*\*IV.2\*\*, ceci n'est en fait rien d'autre qu'une reformulation de la propriété de dualité forte, révélant à nouveau les liens étroits qui unissent les problèmes primal et dual : une condition à la fois nécessaire et suffisante pour l'optimalité d'une solution admissible pour le problème primal est l'existence d'une solution admissible pour le problème dual avec un saut de dualité nul (c'est-à-dire la même valeur de la fonction objectif).

De façon tout à fait similaire, l'écriture des conditions KKT pour le problème dual mènerait au même système d'équations, imposant l'existence d'une solution primale admissible avec un saut de dualité nul.

### XIII.2.3 La méthode de Newton

Une des idées maîtresses dans la conception des méthodes de point intérieur consiste à remplacer la résolution des problèmes d'optimisation (LP) et (LD) par celle du système d'équations non linéaires (KKT). Il est dès lors naturel de lui appliquer une des méthodes les plus populaires pour la résolution de systèmes d'équations non-linéaires : la *méthode de Newton*, dont nous décrivons à présent brièvement le principe (à proprement parler, les deux premières conditions sont linéaires tandis que seules les équations  $x_i s_i = 0$  sont non linéaires ; les contraintes de positivité ne sont pas des équations et ne peuvent par conséquent pas être traitées par une telle méthode).

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

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

et le pas de Newton  $\Delta x^{(k)}$  est choisi de telle façon que cette approximation linéaire est égale à zéro : on pose donc  $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$  avec  $\Delta x^{(k)} = -J(x^{(k)})^{-1}F(x^{(k)})$ . Le calcul de  $\Delta x^{(k)}$  est généralement effectué en pratique via la résolution du système linéaire  $J(x^{(k)})\Delta x^{(k)} = -F(x^{(k)})$  plutôt qu'en évaluant explicitement l'inverse de  $J(x^{(k)})$ . La convergence vers une solution est garantie à partir du moment où l'itéré initial  $x^{(0)}$  se trouve dans un voisinage suffisamment proche d'un des zéros de  $F$ .

La méthode de Newton est également applicable à des problèmes de minimisation, en procédant comme suit : soit  $g : \mathbb{R}^n \mapsto \mathbb{R}$  une fonction à minimiser. Nous formons l'approximation du second ordre de  $g(x)$  autour de  $x^{(k)}$ , soit

$$g(x^{(k)} + \Delta x^{(k)}) \approx g(x^{(k)}) + \nabla g(x^{(k)})^T \Delta x^{(k)} + \frac{1}{2} \Delta x^{(k)T} \nabla^2 g(x^{(k)}) \Delta x^{(k)} .$$

Si la matrice hessienne  $\nabla^2 g(x^{(k)})$  est définie positive, ce qui se produit lorsque  $g$  est strictement convexe, cette approximation possède un minimum unique, qui sera choisi comme itéré suivant. Il est donc défini par  $\Delta x^{(k)} = -\nabla^2 g(x^{(k)})^{-1} \nabla g(x^{(k)})$ , ce qui décrit en fait une méthode équivalente à l'application de la méthode de Newton à la condition d'optimalité  $\nabla g(x) = 0$  basée sur le gradient de  $g$ .

Les méthodes de point intérieur vont donc tenter de résoudre le systèmes d'équations (KKT) à l'aide de la méthode de Newton. Cette approche présente néanmoins deux inconvénients majeurs :

- D'une part, les contraintes de positivité sur  $x$  et  $s$  ne peuvent être directement prises en compte par la fonction  $F$ , et sont donc totalement ignorées par la méthode de Newton. Rien ne garantit donc que les itérés fournis seront positifs, et donc admissibles.

- D'autre part, la convergence de la méthode de Newton n'est garantie que si l'on démarre d'un point suffisamment proche de la solution recherchée, condition qui peut difficilement être vérifiée dans le cas général.

Une modification de la longueur du pas fourni par la méthode de Newton permet de résoudre ces deux difficultés de façon relativement satisfaisante (voir le paragraphe \*\*XIII.3.2\*\* sur les méthodes de mise à l'échelle affine). Toutefois, l'introduction des concepts de *fonction barrière* et de *chemin central* apporte une réponse bien plus efficace et convaincante à ces deux préoccupations.

### XIII.2.4 La fonction barrière

Une fonction barrière  $\phi : \mathbb{R}^+ \mapsto \mathbb{R}$  est simplement une fonction différentiable telle que  $\lim_{x \rightarrow 0^+} \phi(x) = +\infty$ . À l'aide d'une telle barrière, il est possible de dériver à partir d'un problème possédant des contraintes d'inégalités toute une famille paramétrée de problèmes non contraints, de la façon suivante :

$$\begin{aligned} \min_{x \in \mathbb{R}^n} f(x) \quad \text{tel que} \quad g_i(x) \geq 0 \quad \forall i & \quad (G) \\ \rightarrow \min_{x \in \mathbb{R}^n} f(x) + \mu \sum_i \phi(g_i(x)) , & \quad (G^\mu) \end{aligned}$$

où le paramètre  $\mu \in \mathbb{R}^+$  est un réel positif. Le rôle du terme barrière ajouté consiste à tenir les itérés, générés par une méthode d'optimisation pour problèmes sans contraintes, éloignés de la zone non admissible (c'est-à-dire où un ou plusieurs des termes  $g_i$  sont négatifs). Bien sûr, on ne peut pas s'attendre à ce que les solutions optimales du problème  $(G^\mu)$  soit égales à celles du problème d'origine  $(G)$ . En fait, chaque valeur de  $\mu$  fournit un problème  $(G^\mu)$  différent avec ses propres solutions optimales.

Cependant, si on résout une série de problèmes  $(G^\mu)$  en faisant décroître le paramètre  $\mu$  vers zéro, on peut s'attendre à ce que la suite des solutions optimales obtenues converge vers la solution optimale du problème d'origine  $(G)$ , puisque l'impact du terme barrière devient de moins en moins prononcé en comparaison avec la fonction objectif réelle. L'avantage de cette procédure réside dans le fait que chaque solution optimale obtenue au cours de la résolution de la série de problèmes paramétrés satisfera strictement les contraintes d'inégalité  $g_i(x) > 0$ , conduisant ainsi à la limite à une solution admissible et optimale pour le problème  $(G)$  (la notion de fonction barrière fut introduite à l'origine par Frisch puis Fiacco et McCormick pour traiter les problèmes non linéaires [7, 6]).

L'application de cette technique à la programmation linéaire va nous mener au dernier concept fondamental pour les méthodes de point intérieur : le *chemin central*.

### XIII.2.5 Le chemin central

Les méthodes de point intérieur sont basées sur l'utilisation d'une fonction barrière bien spécifique, appelée *barrière logarithmique* :

$$\phi(x) = -\log(x) .$$

À l'aide de cette fonction  $\phi$ , appliquons un terme barrière au problème d'optimisation linéaire (LP)

$$\min_{x \in \mathbb{R}^n} cx - \mu \sum_i \log(x_i) \quad \text{tel que} \quad \begin{cases} Tx = d \\ x > 0 \end{cases} \quad (LP^\mu)$$

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

$$\max_{u \in \mathbb{R}^m} ud + \mu \sum_i \log(s_i) \quad \text{tel que} \quad \begin{cases} uT + s = c \\ s > 0 \text{ et } u \text{ libre} \end{cases} . \quad (LD^\mu)$$

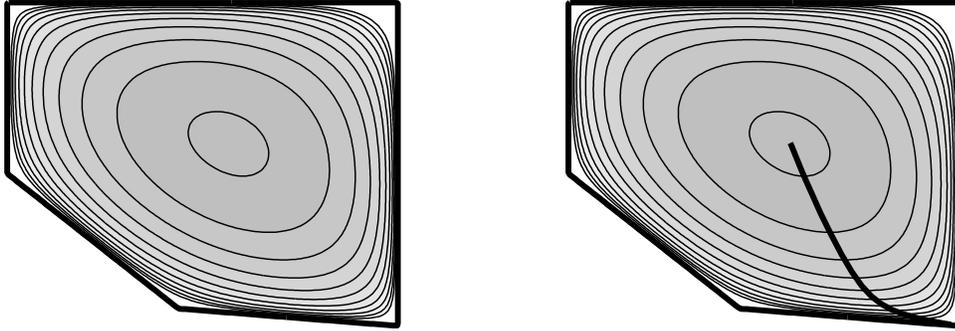


Figure XIII.1: Courbes de niveau du problème  $(LP^\mu)$  (à gauche) et chemin central (à droite).

Un exemple de problème linéaire (format dual  $(LD')$ , avec inégalités), où l'on a représenté les courbes de niveau de l'objectif perturbé par la fonction barrière du problème  $(LP^\mu)$  est représenté sur la figure XIII.1.

On peut montrer (voir par exemple [16, théorème II.4]) que ces deux problèmes ont chacun une solution optimale unique  $x^{(\mu)}$  et  $(u^{(\mu)}, s^{(\mu)})$  pour tout  $\mu > 0$  à condition de faire l'hypothèse que les domaines strictement admissibles  $\mathcal{P}^+$  et  $\mathcal{D}^+$  sont tous deux non vides (cette hypothèse, connue sous le nom de *condition de point intérieur*, est en fait nécessaire et suffisante). Nous admettrons donc dans la suite de ce chapitre qu'il existe au moins une solution strictement admissible pour le problème primal et pour le problème dual (nous verrons dans le paragraphe XIII.4.1 que cette condition n'est pas véritablement restrictive et comment il est possible de s'en affranchir en pratique).

Lorsque  $\mu$  varie sur l'intervalle  $]0 + \infty[$ , on appellera ces deux ensembles de solutions optimales  $\{x^{(\mu)} \mid \mu > 0\} \subset \mathcal{P}^+$  et  $\{(u^{(\mu)}, s^{(\mu)}) \mid \mu > 0\} \subset \mathcal{D}^+$  respectivement *chemin central* primal et dual (voir un exemple à droite de la figure XIII.1). Ces courbes paramétrées possèdent les propriétés suivantes :

- La valeur de la fonction objectif primale (resp. duale)  $cx$  (resp.  $ud$ ) décroît (resp. croît) de façon monotone le long du chemin central primal (resp. dual) lorsque  $\mu$  tend en décroissant vers zéro.
- Le saut de dualité  $cx^{(\mu)} - ud^{(\mu)}$  évalué pour la solution primale-duale  $(x^{(\mu)}, u^{(\mu)}, s^{(\mu)})$  est précisément égal à  $n\mu$ , raison pour laquelle on baptisera  $\mu$  *mesure de dualité*. Lorsqu'un point  $(x, u, s)$  ne se trouve pas exactement sur le chemin central, on pourra calculer une mesure de dualité *estimée* à l'aide de la formule  $\mu = (cx - ud)/n$ .
- Les points limites  $x^* = \lim_{\mu \rightarrow 0} x^{(\mu)}$  et  $(u^*, s^*) = \lim_{\mu \rightarrow 0} (u^{(\mu)}, s^{(\mu)})$  existent et sont par conséquent des solutions optimales pour les problèmes  $(LP)$  et  $(LD)$  (en vertu du fait qu'on a par continuité  $cx^* - u^*d = 0$  et de la propriété de dualité forte). De plus, il est possible de montrer que ces solutions vérifient l'inégalité stricte  $x^* + s^* > 0$ , et on appellera un tel couple de solutions optimales *strictement complémentaires* (pour toute solution optimale  $(x, s)$ , on sait que  $x_i s_i = 0$ , et donc qu'au moins une des deux variables  $x_i$  et  $s_i$  est nulle, ce qui justifie le nom de solution complémentaire ; dans le cas d'une solution strictement complémentaire, on a *exactement* une des deux variables  $x_i$  et  $s_i$  égale à zéro (cf. théorème fort des écarts complémentaires au paragraphe IV.2.3).

### XIII.2.6 Lien entre le chemin central et les conditions d'optimalité KKT

Pour conclure cette section, nous mettons en évidence un lien intéressant entre le chemin central et les conditions d'optimalité (KKT) : si on tente d'écrire les conditions d'optimalité (KKT) pour

le problème primal perturbé ( $LP^\mu$ ) ou pour le problème dual perturbé ( $LD^\mu$ ), on trouve dans les deux cas le même système de conditions nécessaires et suffisantes :

$$\left\{ \begin{array}{l} Tx = d \\ uT + s = c \\ x_i s_i = \mu \forall i \\ x \text{ et } s > 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} x \in \mathcal{P}^+ \\ (u, s) \in \mathcal{D}^+ \\ x_i s_i = \mu \forall i \end{array} \right. . \quad (\text{KKT}^\mu)$$

Ce système est très similaire au système (KKT) d'origine, les seules différences consistant en une modification du membre de droite de la troisième condition et l'emploi d'inégalités strictes. Cela signifie donc que les points du chemin central satisfont une version légèrement perturbée des conditions d'optimalité (KKT) pour les problèmes (LP) et (LD).

Nous sommes à présent en possession de tous les outils nécessaires à la description des méthodes de point intérieur pour la programmation linéaire.

### XIII.3 Méthodes de point intérieur

Depuis la découverte de Karmarkar, de nombreuses variantes de méthodes de point intérieur ont été développées. Il est important de réaliser qu'il existe un véritable arsenal de méthodes, basées sur les mêmes principes fondamentaux mais dont les caractéristiques individuelles peuvent varier assez fortement.

Parmi les critères couramment utilisés pour différencier les méthodes, mentionnons

- **Espace des itérés.** Une méthode est dite *primale*, *duale* ou *primale-duale* lorsque ses itérés appartiennent respectivement à l'espace des variables primales, duales ou le produit cartésien de ces deux espaces.
- **Type d'itérés.** Une méthode est dite *admissible* lorsque ses itérés sont admissibles, c'est-à-dire lorsqu'ils satisfont à la fois les contraintes d'égalité et de positivité. Dans le cas d'une méthode *non admissible*, on autorise les itérés à ne plus vérifier les contraintes d'égalité, tout en s'assurant qu'ils satisfont toujours les contraintes de positivité.
- **Type d'algorithme.** C'est le point de différenciation principal entre les méthodes. Bien que les dénominations ne soient pas à l'heure actuelle totalement standardisées, on distinguera les méthodes de *suitte de chemin* (*path-following algorithms*), les méthodes dites de *mise à l'échelle affine* (*affine-scaling algorithms*) et les méthodes de *réduction de potentiel* (*potential reduction algorithms*). Les trois paragraphes qui suivent décriront ces trois types de méthodes de manière plus détaillée.
- **Type de pas.** Afin de garantir une complexité algorithmique de type polynomial, certains algorithmes sont forcés de prendre de très petits pas à chaque itération, ce qui conduit à un nombre d'itérations assez élevé lorsqu'on les applique à des problèmes pratiques (ceci n'entre pas en contradiction avec le fait que le nombre d'itérations est borné de façon polynomiale par la taille du problème ; cela peut simplement signifier que les coefficients du polynôme sont élevés). Ces méthodes sont appelées méthodes à *pas courts* et présentent surtout un intérêt théorique. Par conséquent, des méthodes à *pas longs* ont été développées, permettant à chaque itération une mise à jour bien plus importante des variables, et constituent les seules méthodes véritablement utilisées en pratique.

Notre objectif n'est pas de fournir une liste exhaustive de toutes les méthodes qui ont été proposées à ce jour, mais plutôt de présenter quelques algorithmes représentatifs, en mettant en évidence les idées sous-jacentes.

### XIII.3.1 Méthodes de suivi de chemin

#### a) Principe

Débutons avec la classe de méthodes la plus élégante : les méthodes de suivi de chemin. Comme le suggère leur dénomination, elles consistent à suivre le chemin central jusqu'à sa limite optimale. Il est assez naturel de proposer l'algorithme conceptuel suivant (nous préférons rester encore assez général à ce stade et ne précisons pas si la méthode est primale, duale ou primale-duale)

**Soit** un itéré initial  $v^{(0)}$  et une suite de mesures de dualité décroissant de façon monotone vers zéro :  $\mu_1 > \mu_2 > \mu_3 > \dots > 0$  et  $\lim_{k \rightarrow 0} \mu_k = 0$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

En prenant  $v^{(k)}$  comme point de départ, calculer  $v^{(k+1)}$ , le point du chemin central possédant une mesure de dualité égale à  $\mu_{k+1}$ .

**Fin**

Il est clair que cette procédure conduit l'itéré  $v^{(k)}$  à tendre vers le point limite du chemin central, qui est une solution optimale du problème.

Cependant, la détermination à chaque itération d'un point du chemin central via la méthode de Newton requiert en principe le calcul de la solution d'un problème de minimisation du type  $(LP^\mu)$  ou la résolution du système de conditions  $(KKT^\mu)$ , ce qui demande potentiellement beaucoup de calculs. En effet, si on admet qu'une minimisation du type  $(LP^\mu)$  n'est pas fondamentalement différente de celle du problème de départ  $(LP)$ , et n'est par conséquent pas beaucoup plus facile à effectuer, on conçoit aisément que la résolution à chaque itération d'un problème presque aussi compliqué que celui qu'on cherche à résoudre à l'origine n'est pas une solution viable.

C'est pourquoi les méthodes de suivi de chemin calculent en fait des itérés se situant *approximativement* sur le chemin central, économisant ainsi de nombreux calculs, et ne suivent donc que grossièrement le chemin central. L'algorithme conceptuel devient alors :

**Soit** un itéré initial  $v^{(0)}$  et une suite de mesures de dualité décroissant de façon monotone vers zéro :  $\mu_1 > \mu_2 > \mu_3 > \dots > 0$  et  $\lim_{k \rightarrow 0} \mu_k = 0$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

En prenant  $v^{(k)}$  comme point de départ, calculer  $v^{(k+1)}$ , une approximation du point du chemin central possédant une mesure de dualité égale à  $\mu_{k+1}$ .

**Fin**

Il est clair que la tâche principale dans l'analyse de la convergence et de la complexité algorithmique de ces méthodes consiste à évaluer avec quelle précision on approche les cibles sur le chemin central (et donc à quelle distance on reste du chemin central).

#### b) Méthode de suivi de chemin primale-duale à pas courts

Cet algorithme spécifique est une méthode primale-duale admissible, ce qui signifie que ses itérés appartiennent à l'ensemble  $\mathcal{P}^+ \times \mathcal{D}^+$ . Soit  $(x^{(k)}, u^{(k)}, s^{(k)})$  l'itéré de l'itération courante, avec sa mesure de dualité  $\mu_k$ . Nous supposons également que cet itéré est *proche* du point  $(x^{(\mu_k)}, u^{(\mu_k)}, s^{(\mu_k)})$  du chemin central. Pour calculer l'itéré suivant, on vise  $(x^{(\mu_{k+1})}, u^{(\mu_{k+1})}, s^{(\mu_{k+1})})$ , un point du chemin central avec une mesure de dualité plus petite  $\mu_{k+1}$  (et donc plus proche de la limite optimale du chemin central). Les deux caractéristiques principales de la méthode à pas courts sont

- La mesure de dualité du point visé pour l'itéré suivant est définie par  $\mu_{k+1} = \sigma \mu_k$  où  $\sigma$  est une constante strictement comprise entre 0 et 1.

- L'itéré suivant sera calculé en appliquant *une seule* itération de la méthode de Newton aux conditions primales-duales perturbées (KKT $^\mu$ ) avec la valeur  $\mu = \sigma\mu_k$  qui définit une cible sur le chemin central (notons bien que nous ignorons pour le moment les contraintes de positivité)

$$\begin{cases} Tx = d \\ uT + s = c \\ x_i s_i = \sigma\mu_k \quad \forall i \end{cases} . \quad (\text{XIII.3})$$

Formellement, nous avons introduit la méthode de Newton en tant qu'outil permettant de calculer les zéros d'une fonction  $F$  et non comme façon de résoudre un système d'équations. Définissons tout d'abord une fonction dont les zéros seront les solutions du système (XIII.3). En effet, en considérant

$$F^{(k)} : \mathbb{R}^{2n+m} \mapsto \mathbb{R}^{2n+m} : \begin{pmatrix} x \\ u^T \\ s^T \end{pmatrix} \mapsto \begin{pmatrix} Tx - d \\ (uT + s - c)^T \\ X^{(k)}S^{(k)}e - \sigma\mu_k e \end{pmatrix} ,$$

où  $e$  représente un vecteur dont toutes les composantes sont égales à 1, tandis que  $X^{(k)}$  et  $S^{(k)}$  sont des matrices carrées reprenant respectivement sur leur diagonale les vecteurs  $x^{(k)}$  et  $s^{(k)}$  (ces notations sont couramment utilisées dans le domaine des méthodes de point intérieur), on trouve alors que le pas préconisé par la méthode de Newton (cf. paragraphe XIII.2.3) est solution du système d'équations linéaires suivant

$$\begin{pmatrix} 0 & T^T & I \\ T & 0 & 0 \\ S^{(k)} & 0 & X^{(k)} \end{pmatrix} \begin{pmatrix} \Delta x^{(k)} \\ \Delta u^{(k)T} \\ \Delta s^{(k)T} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^{(k)}S^{(k)}e + \sigma\mu_k e \end{pmatrix} . \quad (\text{XIII.4})$$

On obtient donc l'algorithme suivant

**Soit** un itéré initial  $(x^{(0)}, u^{(0)}, s^{(0)}) \in \mathcal{P}^+ \times \mathcal{D}^+$  possédant une mesure de dualité  $\mu_0$  et une constante  $0 < \sigma < 1$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas de Newton  $(\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)})$  à l'aide du système d'équations linéaires (XIII.4).

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

**Fin**

Esquissons à présent une preuve de la correction de cet algorithme. Afin que notre stratégie de suivi de chemin fonctionne, nous devons garantir que nos itérés  $(x^{(k)}, u^{(k)}, s^{(k)})$  restent suffisamment proches des points  $(x^{(\mu_k)}, u^{(\mu_k)}, s^{(\mu_k)})$  situés sur le chemin central qui nous guide vers une solution optimale. A cet effet, définissons une quantité mesurant la proximité entre un itéré strictement admissible  $(x, u, s) \in \mathcal{P}^+ \times \mathcal{D}^+$  et le point du chemin central  $(x^{(\mu)}, u^{(\mu)}, s^{(\mu)})$ . Puisque la propriété principale de ce point central (hormis son admissibilité) est  $x_i s_i = \mu \forall i$ , ou de manière équivalente  $x \circ s^T = \mu e$  (on note ici  $x \circ s^T$  le vecteur colonne reprenant le produit composante par composante des vecteurs  $x$  et  $s^T$ ), la quantité suivante (voir par exemple [23])

$$\delta(x, s, \mu) = \frac{1}{\mu} \|x \circ s^T - \mu e\| = \left\| \frac{x \circ s^T}{\mu} - e \right\|$$

semble adéquate : elle est égale à zéro si et seulement si  $(x, u, s)$  est égal à  $(x^{(\mu)}, u^{(\mu)}, s^{(\mu)})$  et augmente au fur et à mesure que l'on s'éloigne de ce point central. Il est également intéressant de constater que la taille du voisinage défini par  $\delta(x, s, \mu) < R$  décroît avec  $\mu$  en raison du facteur de tête  $\frac{1}{\mu}$ .

Une autre possibilité pour mesurer la proximité consiste à prendre (voir [16])

$$\delta(x, s, \mu) = \frac{1}{2} \left\| \sqrt{\frac{x \circ s^T}{\mu}} - \left( \sqrt{\frac{x \circ s^T}{\mu}} \right)^{-1} \right\|$$

où les racines carrées et l'inversion agissent sur les vecteurs composante par composante. A l'aide d'une telle mesure de proximité, l'analyse de l'algorithme repose sur les étapes suivantes :

1. **Admissibilité stricte.** Prouver que l'admissibilité stricte est préservée par le pas de Newton : si  $(x^{(k)}, u^{(k)}, s^{(k)}) \in \mathcal{P}^+ \times \mathcal{D}^+$ , alors  $(x^{(k+1)}, u^{(k+1)}, s^{(k+1)}) \in \mathcal{P}^+ \times \mathcal{D}^+$ . Il faudra être particulièrement attentif aux contraintes de positivité, puisqu'elles ne sont a priori pas prises en compte par la méthode de Newton.
2. **Mesure de dualité.** Prouver que la mesure de dualité visée est atteinte après le pas de Newton : si  $(x^{(k)}, u^{(k)}, s^{(k)})$  possède une mesure de dualité égale à  $\mu_k$ , l'itéré suivant  $(x^{(k+1)}, u^{(k+1)}, s^{(k+1)})$  a une mesure de dualité égale à  $\sigma \mu_k$ .
3. **Proximité.** Prouver que la proximité au chemin central est préservée : il existe une constante  $\tau$  telle que si  $\delta(x^{(k)}, s^{(k)}, \mu_k) < \tau$ , on a  $\delta(x^{(k+1)}, s^{(k+1)}, \mu_{k+1}) < \tau$  après le pas de Newton.

En ajoutant une hypothèse initiale stipulant que  $\delta(x^{(0)}, s^{(0)}, \mu_0) < \tau$ , on peut alors démontrer que la suite des itérés produit par l'algorithme restera confinée dans le voisinage imposé du chemin central et convergera donc (approximativement) vers son point limite, qui est une solution optimale strictement complémentaire.

La dernière question délicate consiste à choisir une combinaison adéquate des constantes  $\sigma$  et  $\tau$  permettant de démontrer les trois étapes ci-dessus. Dans le cas de la première mesure de proximité, on peut choisir (voir [23, chapitre 5])

$$\sigma = 1 - \frac{0.4}{\sqrt{n}} \text{ et } \tau = 0.4 ,$$

où  $n$  dénote la taille des vecteurs  $x$  et  $s$ , tandis que pour la seconde mesure de dualité le choix suivant est acceptable (voir [16, chapitre III.11])

$$\sigma = 1 - \frac{1}{\sqrt{n}} \text{ et } \tau = \frac{1}{\sqrt{2}} .$$

Pour terminer cette description, nous spécifions le critère d'arrêt de la méthode. Étant donné un paramètre  $\varepsilon$  déterminant la précision requise, nous arrêtons l'algorithme lorsque le saut de dualité devient inférieur à  $\varepsilon$ , ce qui se produit dès que  $n\mu_k < \varepsilon$ . Ceci garantit que  $cx$  et  $ud$  approchent la véritable valeur optimale de la fonction objectif avec une erreur inférieure à  $\varepsilon$ . Nous sommes à présent en mesure de formuler l'algorithme sous sa forme finale :

**Soit** un itéré initial  $(x^{(0)}, u^{(0)}, s^{(0)}) \in \mathcal{P}^+ \times \mathcal{D}^+$  possédant une mesure de dualité  $\mu_0$ , la précision requise  $\varepsilon$  et des constantes appropriées  $0 < \sigma < 1$  et  $\tau$  telles que  $\delta(x^{(0)}, u^{(0)}, s^{(0)}) < \tau$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas de Newton  $(\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)})$  à l'aide du système d'équations linéaires (XIII.4).

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

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

En outre, il est possible de prouver qu'une solution de précision  $\varepsilon$  est atteinte après un nombre d'itérations  $N$  vérifiant

$$N = O\left(\sqrt{n} \log \frac{n\mu_0}{\varepsilon}\right). \quad (\text{XIII.5})$$

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

Toutefois, il est important de réaliser que les valeurs de  $\sigma$  préconisées ci-dessus seront en pratique presque égales à un, ce qui conduira à une décroissance très lente des mesures de dualité et par conséquent à des pas de Newton relativement courts (d'où la dénomination de la méthode). Dès lors, bien que de complexité algorithmique polynomiale, cette méthode nécessite un grand nombre d'itérations et n'est pas la plus efficace d'un point de vue pratique.

### c) Méthode de suivi de chemin duale à pas courts

Cette seconde méthode de suivi de chemin est très similaire à la précédente, mis à part le fait que ses itérés appartiennent cette fois uniquement au domaine admissible dual  $\mathcal{D}^+$ . En gardant le principe général qui consiste à suivre le chemin central dual et à y viser les points  $(u^{(\mu_k)}, s^{(\mu_k)})$ , on effectue les adaptations suivantes (il est bien sûr également possible de concevoir une méthode de suivi de chemin primale de façon complètement similaire) :

- Il n'est plus possible de déduire le pas de la méthode de Newton des conditions (KKT $^\mu$ ), puisqu'elles font apparaître à la fois les variables primales et les variables duales. On utilise en remplacement un pas de la méthode de Newton appliquée à la minimisation du problème dual (LD $^\mu$ ) perturbé par la fonction barrière, ce qui conduit au système d'équations linéaires suivant, de taille  $(n + m) \times (n + m)$

$$\begin{pmatrix} \Delta u^{(k)} & \Delta s^{(k)} \end{pmatrix} \begin{pmatrix} T & TS^{(k)-2}T^T \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\sigma\mu_k}d^T - e^T S^{(k)-1}T^T \end{pmatrix}. \quad (\text{XIII.6})$$

- Il est également nécessaire de modifier la mesure de proximité employée : on définit à présent  $\delta(s, \mu)$  selon [16, chapitre III.12]

$$\delta(s, \mu) = \min_{x \in \mathbb{R}^n} \{\delta(x, s, \mu) \mid Tx = d\} = \frac{1}{\mu} \min_{x \in \mathbb{R}^n} \{\|x \circ s^T - \mu e\| \mid Tx = d\}$$

(cette mesure est nulle si et seulement si  $s = s^{(\mu)}$ ).

L'algorithme devient alors

**Soit** un itéré initial  $(u^{(0)}, s^{(0)}) \in \mathcal{D}^+$  possédant une mesure de dualité égale à  $\mu_0$ , la précision requise  $\varepsilon$  et des constantes appropriées  $0 < \sigma < 1$  et  $\tau$  telles que  $\delta(u^{(0)}, s^{(0)}) < \tau$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

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

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

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

Dans ce cas, on peut par exemple prendre

$$\sigma = 1 - \frac{1}{3\sqrt{n}} \text{ et } \tau = \frac{1}{\sqrt{2}},$$

ce qui conduit à la même borne (XIII.5) sur le nombre total d'itérations.

#### d) Méthode de suivi de chemin primale-duale à pas longs

La méthode de suivi de chemin que nous sommes sur le point de décrire tente de contourner la principale limitation des deux méthodes précédentes : la très petite taille de leurs pas. Comme on l'a fait remarquer à du point b) ci-dessus, la raison fondamentale de leur progression très lente en pratique est la valeur de  $\sigma$ , qu'il faut choisir presque égale à un afin d'être capable de prouver la complexité algorithmique polynomiale de la méthode.

Une façon d'accélérer la méthode consisterait simplement à décroître la mesure de dualité de façon plus agressive, c'est-à-dire en utilisant  $\mu_{k+1} = \sigma\mu_k$  avec une valeur de  $\sigma$  plus petite. Cependant, cette modification d'apparence anodine détruit les bonnes propriétés des algorithmes à pas courts. En effet, si la cible sur le chemin central est trop éloignée de l'itéré courant, on constate que

- Le pas de Newton fourni par la résolution de (XIII.4) n'est plus nécessairement admissible, ce qui est facile à comprendre. On demande à la méthode de Newton de résoudre le système  $(\text{KKT}^\mu)$ , composé de deux ensembles d'équations linéaires et d'un ensemble d'équations légèrement non linéaires. En raison de cette non linéarité, le système linéaire que nous résolvons n'est qu'une approximation locale du véritable système d'équations et par conséquent plus on s'éloigne de l'itéré courant, moins l'approximation est précise. Quant le point visé se trouve suffisamment loin, l'approximation linéaire devient si mauvaise que le terme barrière (non linéaire) ne joue plus son rôle et le pas de Newton quitte le domaine admissible en cessant de satisfaire les contraintes de positivité  $x > 0$  et  $s > 0$ . Notons toutefois que puisque les deux premières séries d'équations  $Tx = d$  et  $uT + s = c$  sont linéaires, elles sont automatiquement satisfaites après le pas de Newton complet.

Puisque le respect des contraintes de positivité stricte est un des principes fondamentaux des méthodes de point intérieur, nous n'avons dans ce cas d'autre recours que d'appliquer un pas de Newton *partiel*, c'est-à-dire réduit via un facteur  $\alpha_k < 1$  de façon à rester dans le domaine strictement admissible  $\mathcal{P}^+ \times \mathcal{D}^+$ :

$$(x^{(k+1)}, u^{(k+1)}, s^{(k+1)}) = (x^{(k)}, u^{(k)}, s^{(k)}) + \alpha_k(\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)}).$$

- Cette réduction du pas de Newton détruit la propriété qui stipulait que la mesure de dualité de la cible visée sur le chemin central était toujours atteinte. En fait, on montre aisément que la mesure de dualité après un pas de Newton partiel est égale à  $(1 - \alpha_k(1 - \sigma))\mu_k$ , qui varie de façon linéaire entre  $\mu_k$  et  $\sigma\mu_k$  lorsque  $\alpha$  décroît depuis 1 jusqu'à 0.

Il n'existe malheureusement pas de moyen de contourner cette difficulté, et il faudra accepter le fait que les itérés n'atteignent pas la mesure de dualité visée, à moins que l'on puisse prendre un pas de Newton complet.

- Il n'est plus possible de garantir qu'un unique pas de Newton restaure la proximité au chemin central au sens de l'inégalité  $\delta(x, s, \mu) < \tau$ , pour les mêmes raisons que ci-dessus (non linéarité). Dans la stratégie à pas longs, on appliquera dès lors plusieurs pas de Newton visant la même mesure de dualité jusqu'à ce que la proximité au chemin central soit rétablie. Ce n'est qu'à ce moment qu'on peut choisir une autre cible et décroître  $\mu$ .

Cette méthode à pas longs peut être décrite comme suit :

**Soit** un itéré initial  $(x^{(0)}, u^{(0)}, s^{(0)}) \in \mathcal{P}^+ \times \mathcal{D}^+$ , une mesure de dualité initiale  $\mu_0$ , la précision requise  $\varepsilon$  et des constantes appropriées  $0 < \sigma < 1$  et  $\tau$  telles que  $\delta(x^{(0)}, u^{(0)}, s^{(0)}) < \tau$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas de Newton  $(\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)})$  à l'aide du système d'équations linéaires (XIII.4).

Poser  $(x^{(k+1)}, u^{(k+1)}, s^{(k+1)}) = (x^{(k)}, u^{(k)}, s^{(k)}) + \alpha_k(\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)})$  avec une longueur de pas  $\alpha_k$  choisie de façon à ce que  $(x^{(k+1)}, u^{(k+1)}, s^{(k+1)}) \in \mathcal{P}^+ \times \mathcal{D}^+$ .

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

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

Contrairement à ce qui était imposé par l'analyse de la complexité algorithmique des méthodes à pas courts, on peut ici choisir n'importe quelle valeur pour la constante  $\sigma$ , en particulier des valeurs très inférieures à 1. Ce sont les choix de  $\tau$  et des  $\alpha_k$  qui rendent la méthode polynomiale. La difficulté principale de l'analyse consiste ici à évaluer le nombre d'itérations visant la même valeur de la mesure de dualité nécessaires au rétablissement de la proximité au chemin central. En prenant pour  $\sigma$  une constante indépendante de  $n$  (telle que .5, .1 ou .01), on peut montrer qu'un choix approprié des constantes  $\tau$  et  $\alpha_k$  conduit à un nombre total d'itérations égal à (voir [23, chapitre 5])

$$N = O\left(n \log \frac{n\mu_0}{\varepsilon}\right).$$

Ce résultat est passablement paradoxal : bien que cette méthode effectue des pas plus longs et soit plus efficace en pratique que les méthodes à pas courts, sa complexité algorithmique de pire cas est moins bonne que celle des méthodes à pas courts (XIII.5).

### XIII.3.2 Méthodes de mise à l'échelle affine

#### a) Principe

La quantité impressionnante de recherches effectuées sur le thème des méthodes de point intérieur pour la programmation linéaire fut déclenchée par l'article pionnier de Karmarkar [9]. Sa méthode utilisait des transformations projectives et n'était pas décrite en termes de chemin central ou de méthode de Newton. Les méthodes de mise à l'échelle affine se basent sur des principes similaires, sans toutefois recourir à l'usage de ces transformations projectives (on a découvert, bien après l'explosion des recherches dans ce domaine, que ces méthodes avaient été proposées initialement par Dikin, un chercheur russe, près de 17 ans avant l'article de Karmarkar [5]).

Les méthodes de mise à l'échelle affine ne suivent pas explicitement le chemin central, et ne s'y réfèrent même pas. Le principe de base de ces méthodes est le suivant : considérons par exemple le problème primal (LP)

$$\min_{x \in \mathbb{R}^n} cx \quad \text{tel que} \quad \begin{cases} Tx = d, \\ x \geq 0. \end{cases} \quad (\text{LP})$$

Ce problème est difficile à résoudre à cause des contraintes de positivité, qui donnent au domaine admissible une forme polyédrale. Considérons l'itéré courant  $x^{(k)}$  et remplaçons le domaine admissible polyédral par un ellipsoïde centré en  $x^{(k)}$  qui lui soit inscrit. On va à présent minimiser la fonction objectif sur cet ellipsoïde, ce qui devrait être plus aisé que sur un polyèdre, et prendre le minimum obtenu comme itéré suivant.

Comment construire un ellipsoïde centré en  $x^{(k)}$  et inscrit à l'intérieur du domaine admissible ? Soit une matrice diagonale strictement positive  $D$ . On montre aisément que le problème (P<sub>D</sub>)

$$\min_{\bar{x} \in \mathbb{R}^n} (cD)\bar{x} \quad \text{tel que} \quad \begin{cases} TD\bar{x} = d, \\ \bar{x} \geq 0 \end{cases} \quad (\text{P}_D)$$

est équivalent à (LP), les variables  $x$  y étant simplement mises à l'échelle via l'équation  $x = D\bar{x}$  (d'où la dénomination de la méthode). En choisissant la matrice diagonale particulière  $D = X^{(k)}$ , qui fait correspondre à l'itéré courant  $x^{(k)}$  le vecteur  $\bar{x} = e$ , on obtient le problème suivant

$$\min_{\bar{x} \in \mathbb{R}^n} (cX^{(k)})\bar{x} \quad \text{tel que} \quad \begin{cases} TX^{(k)}\bar{x} = d, \\ \bar{x} \geq 0. \end{cases}$$

On peut alors restreindre le domaine admissible défini par  $\bar{x} \geq 0$  à une boule de rayon 1 centrée en  $e$ , puisqu'on a l'inclusion  $\{\bar{x} \mid \|\bar{x} - e\| \leq 1\} \subset \{\bar{x} \mid \bar{x} \geq 0\}$ . Notre problème devient

$$\min_{\bar{x} \in \mathbb{R}^n} (cX^{(k)})\bar{x} \quad \text{tel que} \quad \begin{cases} TX^{(k)}\bar{x} = d, \\ \|\bar{x} - e\| \leq 1, \end{cases}$$

c'est-à-dire la minimisation d'une fonction objectif linéaire sur l'intersection d'une boule unité et d'un sous-espace défini par des contraintes linéaires. On peut montrer que la solution de ce problème peut se calculer aisément de façon analytique, via la résolution d'un système d'équations linéaires. Exprimé en fonction des variables d'origine  $x$ , ce problème est équivalent à

$$\min_{x \in \mathbb{R}^n} cx \quad \text{tel que} \quad \begin{cases} Tx = d, \\ \|X^{(k)-1}x - e\| \leq 1, \end{cases}$$

dont le domaine admissible est un ellipsoïde centré en  $x^{(k)}$ . Cet ellipsoïde, appelé *ellipsoïde de Dikin*, est entièrement inclus à l'intérieur de  $\mathcal{P}$ . Le minimum sur cet ellipsoïde est donné par  $x^{(k)} + \Delta x^{(k)}$ , avec

$$\Delta x^{(k)} = -\frac{X^{(k)}P_{TX^{(k)}}X^{(k)}c^T}{\|P_{TX^{(k)}}X^{(k)}c^T\|}. \quad (\text{XIII.7})$$

où  $P_Q$  est la matrice de projection sur  $\text{Ker } Q = \{x \mid Qx = 0\}$ , qui peut s'écrire  $P_Q = I - Q^T(QQ^T)^{-1}Q$  lorsque la matrice  $Q$  est de rang maximum.

Puisque l'ellipsoïde appartient entièrement au domaine admissible, le pas  $\Delta x^{(k)}$  est admissible et on peut raisonnablement supposer que l'itéré suivant  $x^{(k)} + \Delta x^{(k)}$  sera plus proche de la solution optimale que  $x^{(k)}$ .

### b) Méthodes de mise à l'échelle affine à pas courts et à pas longs

En introduisant une constante  $\rho$  destinée à réduire la taille du pas effectué, on peut décrire l'algorithme selon

**Soit** un itéré initial  $x^{(0)} \in \mathcal{P}^+$  et une constante  $0 < \rho < 1$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas issu de la mise à l'échelle affine  $\Delta x^{(k)}$  à l'aide du système d'équations linéaires (XIII.7) et poser  $x^{(k+1)} = x^{(k)} + \rho \Delta x^{(k)}$ .

**Fin**

Cet algorithme est en fait la méthode de mise à l'échelle affine à pas court. La convergence vers une solution primale optimale a été prouvée lorsque  $\rho = \frac{1}{8}$ , mais on ne sait toujours pas à ce jour si sa complexité est de type polynomial (si le problème donné vérifie certaines conditions de non-dégénérescence, on peut également prouver la convergence pour toutes les valeurs de  $\rho$  satisfaisant  $0 < \rho < 1$ , voir [21]). Il est bien évidemment possible de concevoir une variante duale et même une variante primale-duale de cette méthode (il suffit de définir les ellipsoïdes de Dikin correspondants).

On peut également tenter de rendre l'algorithme plus efficace en prenant des pas plus longs, y compris en s'autorisant à sortir de l'ellipsoïde de Dikin. En gardant la même direction que dans le cas de la méthode à pas courts, le pas le plus long que l'on peut prendre sans quitter le domaine primal admissible est donné par

$$\Delta x^{(k)} = -\frac{X^{(k)}P_{TX^{(k)}}X^{(k)}c^T}{\max [P_{TX^{(k)}}X^{(k)}c^T]}, \quad (\text{XIII.8})$$

où  $\max[v]$  dénote la composante maximale du vecteur  $v$ , ce qui conduit à formuler l'algorithme suivant :

**Soit** un itéré initial  $x^{(0)}$  et une constante  $0 < \lambda < 1$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas issu de la mise à l'échelle affine  $\Delta x^{(k)}$  à l'aide du système d'équations linéaires (XIII.8) et poser  $x^{(k+1)} = x^{(k)} + \lambda \Delta x^{(k)}$ .

**Fin**

La constante  $\lambda$  détermine quelle fraction de la distance à la frontière du domaine admissible doit être parcourue (cette constante doit être strictement inférieure à 1 puisqu'on désire rester à l'intérieur du domaine admissible). La convergence vers une solution optimale a été prouvée lorsque  $0 < \lambda \leq 2/3$  mais, de façon assez surprenante, un contre-exemple a été découvert pour le cas  $\lambda = 0.999$  (voir [13]). En définitive, comme pour la méthode de mise à l'échelle à pas courts, on ne sait pas à l'heure actuelle si cette méthode jouit d'une complexité algorithmique de type polynomial.

### c) Lien avec les méthodes de suivi de chemin

Il existe un lien intéressant et peut-être inattendu entre les méthodes de mise à l'échelle affine et les algorithmes à suivi de chemin. Si on prend par exemple la définition (XIII.6) du pas de Newton pour la méthode de suivi de chemin duale à pas courts, et que l'on fait tendre  $\sigma$  vers zéro (ce qui revient à faire tendre la mesure de dualité visée vers zéro), la direction limite qui en résulte est précisément égale à celle préconisée par la méthode de mise à l'échelle affine duale ! Cette propriété surprenante, également valide pour les méthodes primales, nous permet de donner un autre éclairage à ces deux méthodes :

- Les méthodes de mise à l'échelle affine peuvent être vues comme une application de la méthode de Newton visant directement le point limite du chemin central, c'est-à-dire tentant d'atteindre la solution optimale directement sans suivre le chemin central. Cependant, tout comme pour les méthodes à pas longs, il est nécessaire de réduire le pas préconisé par la méthode de Newton afin de rester à l'intérieur du domaine admissible.
- En regardant (XIII.6), on voit qu'il est possible de décomposer le pas de Newton dual en deux parties:

$$(\Delta u^{(k)}, \Delta s^{(k)}) = \frac{1}{\sigma \mu_k} (\Delta^a u^{(k)}, \Delta^a s^{(k)}) + (\Delta^c u^{(k)}, \Delta^c s^{(k)}),$$

où

$$(\Delta^a u^{(k)}, \Delta^a s^{(k)}) \begin{pmatrix} T & TS^{(k)-2}T^T \\ I & 0 \end{pmatrix} = (0, d^T)$$

et

$$(\Delta^c u^{(k)}, \Delta^c s^{(k)}) \begin{pmatrix} T & TS^{(k)-2}T^T \\ I & 0 \end{pmatrix} = (0, -e^T S^{(k)-1} T^T).$$

- $(\Delta^a u^{(k)}, \Delta^a s^{(k)})$  est la composante de mise à l'échelle affine. C'est la direction préconisée par les méthodes de mise à l'échelle affine, avec pour seul l'objectif l'optimalité de l'itéré suivant.
- $(\Delta^c u^{(k)}, \Delta^c s^{(k)})$  est la composante de centrage. On peut montrer qu'elle vise un point du chemin central possédant la même mesure de dualité que l'itéré courant, c'est-à-dire qu'elle tente uniquement d'améliorer la proximité vis-à-vis du chemin central.

En fait, on peut prouver a posteriori que la plupart des méthodes de point intérieur préconisent des pas résultant de la combinaison de ces deux directions de base.

## XIII.3.3 Méthodes de réduction de potentiel

### a) Principe

Plutôt que de viser une suite décroissante de mesures de dualité, la méthode de Karmarkar utilisait une *fonction potentiel* pour jauger la progression de ses itérés vers la solution optimale. L'objectif d'une fonction potentiel est d'évaluer le mérite d'un itéré, selon les critères suivants :

- Elle doit tendre vers  $-\infty$  si et seulement si les itérés tendent vers l'optimalité.

- Elle doit tendre vers  $+\infty$  lorsque les itérés tendent vers la frontière du domaine admissible sans simultanément s'approcher d'une solution optimale (on ne peut bien sûr pas simplement empêcher la méthode d'approcher la frontière du domaine admissible, puisque la solution optimale recherchée s'y trouve forcément).

L'objectif principal d'une méthode de réduction de potentiel est simplement de réduire la fonction potentiel d'une quantité fixée  $\delta$  lors de chaque itération (d'où son nom). La convergence découle directement de la première propriété ci-dessus.

### b) Méthode de réduction de potentiel primale-duale

Nous décrivons l'application de cette stratégie dans le cas primal-dual. On définit la fonction potentiel de Tanabe-Todd-Ye (voir [19, 20]) sur le domaine strictement admissible primal-dual  $\mathcal{P}^+ \times \mathcal{D}^+$  par

$$\Phi_\rho(x, s) = \rho \log sx - \sum_i \log x_i s_i ,$$

où  $\rho$  est une constante strictement supérieure à  $n$ . On peut également l'écrire comme

$$\Phi_\rho(x, s) = (\rho - n) \log sx - \sum_i \log \frac{x_i s_i}{sx/n} + n \log n$$

et effectuer les observations suivantes

- Le premier terme fait tendre la fonction potentiel vers  $-\infty$  lorsque  $(x, s)$  tendent vers l'optimalité, puisque dans ce cas le saut de dualité  $sx$  tend vers 0.
- Le second terme mesure la *centralité* de l'itéré. Un itéré parfaitement centré verra tous ses produits  $x_i s_i$  égaux à leur valeur moyenne  $sx/n$ , ce qui annulera ce second terme. Dès que ces produits deviennent différents (et s'écartent de leur valeur moyenne), ce terme augmente et peut même tendre vers  $+\infty$  si l'un des produits  $x_i s_i$  tend vers zéro sans que  $sx$  tende également vers zéro (ce qui signifie exactement que l'on s'approche de la frontière du domaine admissible sans tendre vers une solution optimale).

La direction des pas que prend cette méthode n'est pas nouvelle : il s'agit de la même que celle de la méthode de suivi de chemin primale-duale, en visant une mesure de dualité égale à  $n\mu_k/\rho$  (c'est-à-dire avec  $\sigma = n/\rho$ ). Toutefois, dans le cas d'une méthode de réduction de potentiel,  $\mu_k$  ne suivra pas une suite décroissante de valeurs prédéterminées, mais sera recalculé à chaque itération (puisque cet algorithme ne peut garantir que la mesure de dualité visée par le pas de Newton sera bien atteinte). On procède comme suit :

**Soit** un itéré initial  $(x^{(0)}, u^{(0)}, s^{(0)}) \in \mathcal{P}^+ \times \mathcal{D}^+$  possédant une mesure de dualité  $\mu_0$  et une constante  $\rho > n$ . Poser  $\sigma = n/\rho$ .

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas de Newton  $(\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)})$  à l'aide du système d'équations linéaires (XIII.4).

Poser  $(x^{(k+1)}, u^{(k+1)}, s^{(k+1)}) = (x^{(k)}, u^{(k)}, s^{(k)}) + \alpha_k (\Delta x^{(k)}, \Delta u^{(k)}, \Delta s^{(k)})$  où  $\alpha_k$  est défini par

$$\begin{aligned} \alpha_k &= \arg \min_{\alpha} \Phi_\rho(x^{(k)} + \alpha \Delta x^{(k)}, s^{(k)} + \alpha \Delta s^{(k)}) \\ &\text{tel que } (x^{(k)} + \alpha \Delta x^{(k)}, u^{(k)} + \alpha \Delta u^{(k)}, s^{(k)} + \alpha \Delta s^{(k)}) \in \mathcal{P}^+ \times \mathcal{D}^+ . \end{aligned}$$

Évaluer  $\mu_{k+1}$  via  $(s^{(k+1)} x^{(k+1)})/n$ .

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

Le principe de cette méthode consiste donc à minimiser à chaque itération la fonction potentiel le long de la direction préconisée par la méthode de Newton. Le point clé dans l'analyse de sa complexité algorithmique réside dans la preuve que ce pas occasionnera à chaque itération une réduction de la fonction potentiel  $\Phi_\rho$  au moins égale à une quantité fixée  $\delta$ . En posant  $\rho = n + \sqrt{n}$ , on peut montrer que

$$\Phi_\rho(x^{(k+1)}, s^{(k+1)}) \leq \Phi_\rho(x^{(k)}, s^{(k)}) - \delta$$

avec  $\delta = 0.16$  (voir par exemple [3]), ce qui conduit à un nombre total d'itérations égal à

$$N = O\left(\sqrt{n} \log \frac{n\mu_0}{\varepsilon}\right),$$

faisant jeu égal avec les meilleures méthodes de suivi de chemin.

Il est en général beaucoup trop coûteux pour un algorithme pratique de minimiser exactement la fonction potentiel le long de la direction de recherche, car  $\Phi_\rho$  est une fonction hautement non linéaire. On utilise par conséquent l'une des stratégies suivantes :

- Définir une approximation quadratique de  $\Phi_\rho$  le long de la direction de recherche et prendre le point qui atteint son minimum comme itéré suivant.
- Parcourir un pourcentage fixé (par exemple 95%) du plus grand pas le long de la direction de recherche qui reste à l'intérieur du domaine admissible.

Notons toutefois qu'on ne peut dans ce cas continuer à garantir une complexité algorithmique polynomiale que si on peut s'assurer que la fonction potentiel est réduite d'une quantité constante à chaque itération.

## XIII.4 Améliorations

Les méthodes que nous avons décrites jusqu'à présent souffrent de certaines limitations (nécessité de connaître un point de départ admissible, nombre d'itérations potentiellement élevé en pratique) qui restreignent essentiellement leur utilisation à un contexte assez théorique. Nous présentons dans ce qui suit différentes améliorations facilitant leur implémentation et leur utilisation en pratique.

### XIII.4.1 Méthodes non-admissibles

Les algorithmes décrits jusqu'à présent sont des méthodes admissibles, ce qui signifie qu'ils nécessitent un itéré de départ strictement admissible. Un tel point n'est malheureusement pas toujours aisément disponible :

- La formulation de certains problèmes ne permet pas d'obtenir directement une solution (strictement) admissible, et chercher une telle solution peut s'avérer aussi difficile que de résoudre le problème d'optimisation en lui-même.
- Certains problèmes ne possèdent aucun point strictement admissible, bien qu'ils soient tout à fait valides et admettent des solutions optimales finies. Cette situation se produit en fait lorsque l'ensemble des solutions optimales n'est pas borné (ce qui est le cas par exemple lorsqu'une variable non bornée par les contraintes est absente de la fonction objectif).

Il existe deux stratégies permettant de traiter ces situations :

- Transformer le problème (généralement via l'ajout de contraintes et/ou de variables) de manière à lui faire admettre un point de départ strictement admissible aisément identifiable (ce qui sera présenté dans le paragraphe suivant).
- Modifier l'algorithme de façon à lui permettre de travailler avec des itérés n'appartenant pas au domaine admissible, ce qui permet d'obtenir une méthode non admissible.

Donnons à présent un aperçu de cette seconde option.

Rappelons que les itérés d'une méthode non admissible ne sont pas obligés de vérifier les contraintes d'égalité linéaires  $Tx = d$  et  $uT + s = c$  mais doivent toutefois rester positifs ( $x > 0$  et  $s > 0$ ). L'idée maîtresse pour convertir un algorithme admissible en méthode non admissible consiste à utiliser la méthode de Newton pour rendre les itérés progressivement admissibles au cours de la résolution. Cela revient en fait à modifier légèrement le système d'équations linéaires (XIII.4), qui devient

$$\begin{pmatrix} 0 & T^T & I \\ T & 0 & 0 \\ S^{(k)} & 0 & X^{(k)} \end{pmatrix} \begin{pmatrix} \Delta x^{(k)} \\ \Delta u^{(k)T} \\ \Delta s^{(k)T} \end{pmatrix} = \begin{pmatrix} (c - u^{(k)}T - s^{(k)})^T \\ d - Tx^{(k)} \\ -X^{(k)}S^{(k)}e + \sigma\mu_k e \end{pmatrix}. \quad (\text{XIII.9})$$

La seule différence avec le système dans le cas admissible se trouve dans le vecteur du membre de droite, qui incorpore à présent les *résidus* primal  $d - Tx^{(k)}$  et dual  $c - u^{(k)}T - s^{(k)}$ . Les pas de la méthode de Newton tenteront alors de réduire simultanément la mesure de dualité de l'itéré courant et son écart par rapport à l'admissibilité.

Des variantes non admissibles ont été développées à la fois pour les méthodes de suivi de chemin et les méthodes de réduction de potentiel. Sans rentrer dans plus de détails, mentionnons qu'il est nécessaire d'inclure une contrainte supplémentaire sur le pas préconisé par la méthode afin de s'assurer que l'écart par rapport à l'admissibilité est réduit au moins au même rythme que la mesure de dualité (cela permet d'éviter que l'algorithme se termine avec une solution "optimale" du point de vue de la mesure de dualité mais non admissible). La complexité algorithmique de ces méthodes est généralement identique à celle de leurs contreparties admissibles, bien que leur analyse soit en général beaucoup plus ardue.

### XIII.4.2 Technique du problème homogène auto-dual

Une autre façon de pallier l'indisponibilité d'un itéré de départ strictement admissible consiste à modifier le problème de façon à lui faire admettre un point de départ strictement admissible aisément identifiable. La méthode des variables artificielles présentée au paragraphe II.5 est un exemple d'application de cette technique. Nous présentons ici une seconde stratégie plus spécifique aux méthodes de point intérieur. Choisissons un itéré de départ arbitraire  $(x^{(0)}, u^{(0)}, s^{(0)})$  tel que  $x^{(0)} > 0$  et  $s^{(0)} > 0$  et définissons les quantités suivantes

$$\begin{aligned} \hat{d} &= d - Tx^{(0)} \\ \hat{c} &= c - u^{(0)}T - s^{(0)} \\ \hat{g} &= u^{(0)}d - cx^{(0)} - 1 \\ \hat{h} &= s^{(0)}x^{(0)} + 1. \end{aligned}$$

On considère à présent le problème suivant, initialement présenté dans [25]

$$\begin{array}{rcllcl} \min & & & \hat{h}\theta & & \\ \text{tel que} & Tx & -d\tau & +\hat{d}\theta & & = 0 \\ & -uT & & +c\tau & -s & = 0 \\ & ud & -cx & -\hat{g}\theta & & -\kappa = 0 \\ & -u\hat{d} & +\hat{c}x & +\hat{g}\tau & & = -\hat{h} \\ & x \geq 0 & \tau \geq 0 & & s \geq 0 & \kappa \geq 0 \end{array} \quad (\text{HSD})$$

Il n'est pas difficile de trouver un point de départ strictement admissible pour ce problème. En fait, on vérifie aisément que le point  $(x, u, s, \tau, \kappa, \theta) = (x^{(0)}, u^{(0)}, s^{(0)}, 1, 1, 1)$  est un choix possible. Sans rentrer dans les détails, on peut donner une brève description des nouvelles variables introduites dans le problème (HSD) :  $\tau$  est une variable d'homogénéisation,  $\theta$  mesure l'écart par rapport à l'admissibilité et  $\kappa$  renvoie au saut de dualité du problème d'origine. Signalons encore que les deux premières conditions correspondent aux contraintes linéaires  $Tx = d$  et  $uT + s = c$ . Ce problème possède les caractéristiques suivantes :

- Ce problème est homogène, c'est-à-dire que son membre de droite est égale au vecteur nul (exception faite de sa dernière composante, nécessaire à la dernière égalité qui est une contrainte d'homogénéisation).
- Ce problème est auto-dual, ce qui signifie que son dual lui est identique (c'est dû au fait que la matrice des coefficients est antisymétrique).
- La valeur optimale du problème (HSD) est égale à 0 (ce qui signifie  $\theta^* = 0$ ).
- Étant donné une solution optimale au problème (HSD) strictement complémentaire  $(x^*, u^*, s^*, \tau^*, \kappa^*, 0)$ , on a soit  $\tau^* > 0$ , soit  $\kappa^* > 0$ .
  - Si  $\tau^* > 0$ , alors  $(x^*/\tau^*, u^*/\tau^*, s^*/\tau^*)$  est une solution optimale du problème d'origine.
  - Si  $\kappa^* > 0$ , alors le problème d'origine ne possède pas de solution optimale finie. De plus, on a dans ce cas  $u^*d - cx^* > 0$  et
    - \* Lorsque  $u^*d > 0$ , le problème primal (LP) n'admet aucune solution admissible.
    - \* Lorsque  $-cx^* > 0$ , le problème dual (LD) n'admet aucune solution admissible.

Puisque nous connaissons un point de départ strictement admissible pour ce problème, nous pouvons lui appliquer une méthode de suivi de chemin admissible qui convergera vers une solution optimale strictement complémentaire. A l'aide des propriétés mentionnées ci-dessus, il est alors toujours possible de calculer une solution optimale du problème d'origine ou de détecter l'absence de solution admissible.

Les dimensions du problème homogène auto-dual sont approximativement le double de celles du problème d'origine, ce qui peut être vu comme un inconvénient en pratique. Cependant, il est possible de tirer parti de la propriété d'auto-dualité et d'utiliser certaines techniques d'algèbre linéaire afin de résoudre ce problème à un coût presque identique à celui du problème d'origine (voir par exemple [1, section 6.3]).

### XIII.4.3 Différences entre algorithme théorique et implémentation

Nous avons déjà observé qu'une complexité algorithmique de type polynomial n'est pas nécessairement une garantie d'un bon comportement en pratique. Les méthodes à pas courts sont irrémédiablement trop lentes en raison de l'infime réduction de la mesure de dualité qu'elles préconisent à chaque itération. Les méthodes à pas longs sont plus performantes mais restent relativement lentes. C'est la raison pour laquelle les praticiens ont examiné différentes adaptations destinées à accélérer le comportement pratique de ces algorithmes. Il est important de réaliser que les résultats de complexité algorithmique que nous avons mentionnés jusqu'à présent ne s'appliquent pas à ces méthodes modifiées, puisqu'elles ne suivent plus fidèlement la théorie.

La méthode de suivi de chemin primale-duale non admissible est de loin l'algorithme de point intérieur le plus couramment implémenté. On lui apporte généralement les modifications suivantes :

- La méthode théorique à pas longs nécessite plusieurs pas de Newton visant la même mesure de dualité pour rétablir la proximité au chemin central. Les algorithmes pratiques ignorent cette considération et, à la manière des méthodes à pas courts, n'effectuent qu'un seul pas de Newton.
- Au lieu de choisir la longueur de pas recommandée par la théorie, les algorithmes pratiques considèrent généralement une large fraction du plus grand pas restant à l'intérieur du domaine admissible (on utilise couramment des valeurs telles que 99.5% ou 99.9% du pas maximal). Cette modification est particulièrement efficace dans le cas des méthodes primales-duales.

- On utilise des longueurs de pas différentes pour les itérés primaux et duaux, c'est-à-dire que l'on prend

$$x^{(k+1)} = x^{(k)} + \alpha^P \Delta x^{(k)} \text{ et } (u^{(k+1)}, s^{(k+1)}) = (u^{(k)}, s^{(k)}) + \alpha^D (\Delta u^{(k)}, \Delta s^{(k)}) .$$

Ces pas sont choisis conformément à la remarque précédente, par exemple selon

$$(\alpha^P, \alpha^D) = 0.995 (\alpha_{\max}^P, \alpha_{\max}^D)$$

. Cette modification est souvent responsable à elle seule d'une réduction substantielle du nombre total d'itérations, sans que ce comportement soit à l'heure actuelle justifié par la théorie.

#### XIII.4.4 L'algorithme prédicteur-correcteur de Mehrotra

La description des méthodes qui a été faite dans la section précédente a souligné le fait que la constante  $\sigma$ , responsable du choix de la mesure de dualité  $\sigma \mu_k$  visée pour l'itéré suivant, joue un rôle très important dans la détermination de l'efficacité de l'algorithme :

- Choisir  $\sigma$  proche de 1 permet d'employer un pas de Newton complet, mais ce pas est souvent très court et ne réalise que peu de progrès vers la solution optimale. Cependant, il présente l'avantage d'augmenter la proximité vis-à-vis du chemin central.
- Choisir une valeur plus faible de  $\sigma$  résulte souvent dans un pas de Newton plus grand autorisant une progression plus conséquente vers la solution optimale, mais ce pas conduit généralement hors du domaine admissible et doit être réduit. De plus, ce type de pas tend généralement à éloigner les itérés du chemin central.

On comprend alors que le choix de la meilleure valeur possible de  $\sigma$  puisse dépendre de l'itéré courant : petite si une cible éloignée est facile à atteindre, grande dans le cas contraire. En se basant sur ces considérations, Mehrotra a conçu un choix heuristique de  $\sigma$  très efficace : l'algorithme prédicteur-correcteur [14].

Cet algorithme commence par calculer un pas *prédicteur*  $(\Delta x^{(k)a}, \Delta u^{(k)a}, \Delta s^{(k)a})$  obtenu en résolvant le système (XIII.9) pour  $\sigma = 0$ , en visant donc directement la limite optimale du chemin central. On calcule alors séparément les longueurs de pas maximales pour le primal et le dual, selon

$$\alpha_k^{a,P} = \arg \max \left\{ \alpha \in [0, 1] \mid x^{(k)} + \alpha \Delta x^{(k)a} \geq 0 \right\} , \quad (\text{XIII.10})$$

$$\alpha_k^{a,D} = \arg \max \left\{ \alpha \in [0, 1] \mid s^{(k)} + \alpha \Delta s^{(k)a} \geq 0 \right\} . \quad (\text{XIII.11})$$

Enfin, on évalue la mesure de dualité de l'itéré ainsi obtenu à l'aide de

$$\mu_{k+1}^a = \frac{1}{n} (s^{(k)} + \alpha_k^{a,D} \Delta s^{(k)a}) (x^{(k)} + \alpha_k^{a,P} \Delta x^{(k)a}) . \quad (\text{XIII.12})$$

Cette quantité mesure la facilité avec laquelle on peut progresser vers l'optimalité : si elle est très inférieure à la mesure de dualité actuelle  $\mu_k$ , on pourra choisir une petite valeur de  $\sigma$  et espérer progresser de façon notable vers l'optimum, tandis que si elle n'est que légèrement inférieure à  $\mu_k$ , il faut être plus prudent et choisir une valeur de  $\sigma$  plus proche de 1, de façon à se rapprocher du chemin central avec l'espoir de se retrouver en meilleure posture pour une grande réduction de la mesure de dualité lors de l'itération *suivante*. Mehrotra suggère l'heuristique suivante, qui s'est révélée très efficace en pratique

$$\sigma = \left( \frac{\mu_{k+1}^a}{\mu_k} \right)^3 .$$

On peut alors calculer le pas *correcteur*  $(\Delta x^{(k)c}, \Delta u^{(k)c}, \Delta s^{(k)c})$  calculé avec cette valeur de  $\sigma$  et prendre séparément pour le primal et le dual les pas maximaux restant à l'intérieur du domaine admissible.

On peut encore légèrement améliorer cet algorithme en faisant l'observation suivante : après un pas prédicteur complet, les produits  $x_i s_i$  deviennent  $(x_i + \Delta x_i^a)(s_i + \Delta s_i^a)$ , quantités qui sont en fait égales à  $\Delta x_i^a \Delta s_i^a$ . Puisque ce pas de Newton tentait de rendre le produit  $x_i s_i$  égal à zéro, le produit réellement obtenu  $\Delta x_i^a \Delta s_i^a$  mesure en quelque sorte l'erreur commise par la méthode de Newton en raison de la non linéarité des équations que l'on essaie de résoudre. On peut dès lors incorporer ce terme d'erreur dans le calcul du pas correcteur, à l'aide de la modification suivante du membre de droite dans (XIII.9)

$$\begin{pmatrix} 0 & T^T & I \\ T & 0 & 0 \\ S^{(k)} & 0 & X^{(k)} \end{pmatrix} \begin{pmatrix} \Delta x^{(k)} \\ \Delta u^{(k)T} \\ \Delta s^{(k)T} \end{pmatrix} = \begin{pmatrix} (c - u^{(k)T} - s^{(k)})^T \\ d - T x^{(k)} \\ -X^{(k)} S^{(k)} e - \Delta X_k^a \Delta S_k^a e + \sigma \mu_k e \end{pmatrix}. \quad (\text{XIII.13})$$

Cette stratégie consistant à calculer un pas en tenant compte des résultats d'une prédiction du premier ordre résulte en une méthode du second ordre. L'algorithme complet s'écrit alors :

**Soit** un itéré initial  $(x^{(0)}, u^{(0)}, s^{(0)})$  possédant une mesure de dualité  $\mu_0$  telle que  $x^{(0)} > 0$  et  $s^{(0)} > 0$ , la précision requise  $\varepsilon$  et une constante  $\rho < 1$  (par exemple 0.995 ou 0.999).

**Répéter** pour  $k = 0, 1, 2, \dots$

Calculer le pas de Newton prédicteur  $(\Delta x^{(k)a}, \Delta u^{(k)a}, \Delta s^{(k)a})$  à l'aide du système d'équations linéaires (XIII.9) et  $\sigma = 0$ .

Calculer les longueurs de pas maximales et la mesure de dualité qui en résulte selon les équations (XIII.10), (XIII.11) et (XIII.12).

Calculer le pas de Newton correcteur  $(\Delta x^{(k)c}, \Delta u^{(k)c}, \Delta s^{(k)c})$  à l'aide du système d'équations linéaires modifié (XIII.13) et  $\sigma = (\mu_{k+1}^a / \mu_k)^3$ .

Calculer les longueurs de pas maximales selon

$$\begin{aligned} \alpha_k^P &= \arg \max \left\{ \alpha \in [0, 1] \mid x^{(k)} + \alpha \Delta x^{(k)c} \geq 0 \right\}, \\ \alpha_k^D &= \arg \max \left\{ \alpha \in [0, 1] \mid s^{(k)} + \alpha \Delta s^{(k)c} \geq 0 \right\}. \end{aligned}$$

Poser  $x^{(k+1)} = x^{(k)} + \rho \alpha_k^P \Delta x^{(k)c}$  et  $(u^{(k+1)}, s^{(k+1)}) = (u^{(k)}, s^{(k)}) + \rho \alpha_k^D (\Delta u^{(k)c}, \Delta s^{(k)c})$ .

Évaluer  $\mu_{k+1}$  à l'aide de  $(s^{(k+1)} x^{(k+1)})/n$ .

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

Il est crucial de réaliser que le pas prédicteur n'est utilisé que pour choisir  $\sigma$  et déterminer le membre de droite dans (XIII.13) et n'est donc pas appliqué à l'itéré courant. Ceci a une conséquence importante sur les calculs effectués par l'algorithme, car la détermination du pas prédicteur *et* du pas correcteur se fait à partir du même itéré courant, ce qui implique que les matrices des systèmes linéaires (XIII.9) et (XIII.13) sont identiques, seuls les vecteurs des membres de droite différant. Comme on le verra dans le paragraphe XIII.5.1, la résolution du second système pourra réutiliser la factorisation de la matrice utilisée pour le pas prédicteur et ne nécessitera qu'une opération de substitution peu coûteuse en temps de calcul. Cette caractéristique est en partie responsable de la grande efficacité de l'algorithme de Mehrotra : une heuristique ingénieuse pour réduire la mesure de dualité en augmentant très peu le temps de calcul.

## XIII.5 Implémentation

Nous présentons ici certains faits importants concernant l'implémentation des méthodes de point intérieur.

### XIII.5.1 Algèbre linéaire

Il est essentiel de réaliser que la résolution du système d'équations linéaires définissant le pas de Newton consomme la majeure partie de temps de calcul des méthodes de point intérieur (certains auteurs rapportent de 80 à 90% du temps de calcul total). Son implémentation doit par conséquent être réalisée avec beaucoup de soin. Les équations (XIII.9) ne sont généralement pas résolues telles quelles : on effectue successivement l'élimination par substitution des variables  $\Delta s$  (où on a défini  $D^{(k)2} = S^{(k)-1}X^{(k)}$ )

$$\begin{pmatrix} -D^{(k)-2} & T^T \\ T & 0 \end{pmatrix} \begin{pmatrix} \Delta x^{(k)} \\ \Delta u^{(k)T} \end{pmatrix} = \begin{pmatrix} (c - u^{(k)T} - \sigma\mu_k e^T X^{(k)-1})^T \\ d - Tx^{(k)} \end{pmatrix} \quad (\text{XIII.14})$$

$$\Delta s^{(k)} = -s^{(k)} + \sigma\mu_k e^T X^{(k)-1} - \Delta x^{(k)T} D^{(k)-2}, \quad (\text{XIII.15})$$

puis celle des variables  $\Delta x$ , pour obtenir le système

$$\Delta u^{(k)T} D^{(k)2} T^T = d^T - (x^{(k)T} - c D^{(k)2} + u^{(k)T} D^{(k)2} + \sigma\mu_k e^T S^{(k)-1}) T^T \quad (\text{XIII.16})$$

$$\Delta s^{(k)} = c - u^{(k)T} - s^{(k)} - \Delta u^{(k)T} \quad (\text{XIII.17})$$

$$\Delta x^{(k)} = -x^{(k)} + \sigma\mu_k S^{(k)-1} e - D^{(k)2} \Delta s^{(k)T}. \quad (\text{XIII.18})$$

Le système d'équations linéaires (XIII.14) est appelé système *augmenté* : il est symétrique et peut se résoudre à l'aide d'une factorisation de Bunch-Partlett. Toutefois, la façon la plus courante de calculer le pas de Newton consiste à résoudre le système (XIII.16) (aussi connu sous le nom d'équations *normales*) à l'aide d'une factorisation de Cholevsky, en tirant parti du fait que la matrice  $TD^{(k)2}T^T$  est définie positive (voir la discussion dans [1]). A ce stade, il est important de noter que la plupart des problèmes rencontrés en pratique ne comportent que peu d'éléments non nuls dans la matrice  $T$ . Il est par conséquent crucial d'exploiter le caractère *creux* de cette matrice afin de réduire à la fois les temps de calcul et la capacité de stockage mémoire requise. De façon plus spécifique, il s'agit d'identifier une permutation des lignes et des colonnes de la matrice  $TD^{(k)2}T^T$  menant à un facteur de Cholevsky le plus creux possible (malheureusement, le problème consistant à trouver la permutation optimale est  $\mathcal{NP}$ -difficile, ce qui a conduit au développement de nombreuses heuristiques, telle que celle du degré minimum (*minimum degree*) ou du remplissage local minimum (*minimum local fill-in*), voir par exemple [?, section 6.4.1] AndersenGondzioMeszarosXu96. Pour un problème donné, cette permutation ne doit cependant être calculée qu'une fois pour toutes, puisque la disposition des éléments non nuls de la matrice  $TD^{(k)2}T^T$  est la même pour chaque itération.

Mentionnons également que la complexité algorithmique associée à la résolution de ce système linéaire est de  $\mathcal{O}(n^3)$  opérations arithmétiques élémentaires, ce qui donne aux meilleurs méthodes de point intérieur une complexité algorithmique totale égale à

$$\mathcal{O}\left(n^{3.5} \log \frac{n\mu_0}{\varepsilon}\right)$$

opérations arithmétiques élémentaires. Pour être complets, mentionnons qu'une technique relativement sophistiquée basée sur la mise à jour partielle de la matrice  $TD^{(k)2}T^T$  des équations normales permet de réduire encore cette complexité totale à  $\mathcal{O}\left(n^3 \log \frac{n\mu_0}{\varepsilon}\right)$  opérations (voir par exemple [16, chapitre 17]).

### XIII.5.2 Prétraitement

Dans la plupart des cas, les problèmes de programmation linéaire que l'on cherche à résoudre ne sont pas formulés sous la forme standard. La première tâche d'un logiciel de résolution consiste donc à convertir le problème qui lui est fourni, en lui ajoutant des variables et des contraintes :

- Les contraintes d'inégalité peuvent être transformées en contraintes d'égalité à l'aide d'une variable d'écart (cf. paragraphe I.1).

- Une variable libre peut être décomposée en deux variables positives :  $x_i = x_i^+ - x_i^-$  avec  $x_i^+ \geq 0$  et  $x_i^- \geq 0$  (cf. paragraphe I.1). Toutefois, cette procédure présente certains inconvénients dans le cas de la résolution par une méthode de point intérieur (cette transformation rend l'ensemble des solutions optimales non borné et entraîne l'absence de solutions primales-duales strictement admissibles) de telle manière qu'en pratique, les logiciels de résolution utilisent plutôt une modification de leur algorithme permettant de traiter directement les variables libres.
- Les bornes inférieures  $l_i \leq x_i$  (et supérieures  $x_i \leq u_i$ ) sont traitées à l'aide d'une translation  $x_i = l_i + x'_i$  (et  $x_i = u_i - x'_i$ ) avec  $x'_i \geq 0$  (cf. chapitre VII).
- Une variable présentant à la fois une borne inférieure et une borne supérieure  $l_i \leq x_i \leq u_i$  pourrait être traitée à l'aide d'une variable d'écart, mais les logiciels de résolution sont souvent basés sur une variante de la forme standard qui tient directement compte de ce genre de contraintes.

Après cette conversion initiale, il n'est pas rare qu'une série de transformations simples permette de réduire fortement la taille du problème (cela peut être également le cas si le programme linéaire considéré a été obtenu automatiquement via l'utilisation d'un langage de modélisation)

- Une ligne ou une colonne entièrement nulle dans la matrice  $T$  des coefficients est soit redondante, soit rend le problème insoluble.
- Une contrainte d'égalité ne faisant intervenir qu'une seule variable peut être ôtée de la formulation et utilisée pour fixer la valeur de cette variable.
- Une contrainte d'égalité faisant intervenir exactement deux variables peut être utilisée pour éliminer une de ces deux variables par substitution.
- Deux lignes identiques de la matrice  $T$  des coefficients sont soit redondantes (l'une d'entre elles peut être retirée), soit inconsistantes (et rendent le problème insoluble).
- Certaines contraintes peuvent permettre le calcul de bornes inférieures ou supérieures sur certaines variables. Ces bornes peuvent améliorer des bornes existantes, détecter des contraintes redondantes ou diagnostiquer un problème insoluble.

Tous les logiciels de résolution appliquent ces règles (et d'autres) de façon répétée avant de commencer à résoudre un problème.

### XIII.5.3 Itéré de départ et critère d'arrêt

La question de l'identification d'un point de départ approprié a été traitée via la technique du problème homogène auto-dual (cf. paragraphe XIII.4.2) et les méthodes non admissibles (cf. paragraphe XIII.4.1). Dans ces deux cas, n'importe quel itéré vérifiant  $x^{(0)} > 0$  et  $s^{(0)} > 0$  peut être choisi comme point de départ. Cependant, l'efficacité des algorithmes peut être grandement influencée par ce choix.

Bien qu'elle ne repose pas sur une analyse théorique, l'heuristique suivante est souvent utilisée pour fournir un point de départ. On résout d'abord

$$\min_{x \in \mathbb{R}^n} cx + \frac{\omega}{2} \|x\|^2 \quad \text{tel que} \quad Tx = d \quad \text{et} \quad \min_{(u,s) \in \mathbb{R}^m \times \mathbb{R}^n} ud + \frac{\omega}{2} \|s\|^2 \quad \text{tel que} \quad uT + s = c .$$

Ces problèmes quadratiques convexes peuvent être résolus analytiquement en un temps de calcul comparable à celui d'une itération d'une méthode de point intérieur. Les éléments négatifs des vecteurs  $x$  et  $s$  obtenus sont ensuite remplacés par une petite constante positive pour fournir  $x^{(0)}$  et  $(u^{(0)}, s^{(0)})$ .

Une petite valeur prédéterminée  $\varepsilon_g$  du saut de dualité constitue habituellement le critère d'arrêt pour les méthodes de point intérieur. Dans le cas d'une méthode non admissible, les écarts primal

et dual par rapport à l'admissibilité doivent également être pris en compte, et on exige également qu'ils descendent sous une valeur prédéfinie  $\varepsilon_i$ . On peut par exemple utiliser les conditions suivantes

$$\frac{\|Tx - d\|}{\|d\| + 1} < \varepsilon_i, \quad \frac{\|uT + s - c\|}{\|c\| + 1} < \varepsilon_i, \quad \frac{\|cx - ud\|}{\|cx\| + 1} < \varepsilon_g .$$

Le rôle des dénominateurs est de permettre une mesure relative de la précision obtenue, tandis que l'ajout de la constante +1 permet d'éviter une éventuelle division par zéro. Enfin, signalons que lorsqu'on applique une méthode non admissible à un problème ne possédant pas de solution admissible, on observe une divergence des itérés (leur norme tend vers l'infini). Les logiciels de résolution sont capables de détecter ce comportement, ce qui leur permet de diagnostiquer un problème insoluble.

## XIII.6 Remarques finales

La théorie des méthodes de point intérieur est à présent bien établie ; plusieurs ouvrages de référence ont été publiés sur le sujet (voir par exemple [23, 16, 24]). D'un point de vue pratique, ces méthodes sont compétitives avec les meilleures implémentations de l'algorithme simplexe, en particulier pour les problèmes de grande taille (citons entre autres les logiciels commerciaux CPLEX, Xpress-MP et MOSEK).

Toutefois, certains questions n'ont pas encore trouvé de réponse satisfaisante, comme la justification des nombreuses différences existant entre les algorithmes théoriques et ceux qui sont réellement implémentés, ou le fait que le nombre d'itérations que l'on observe en pratique lors de la résolution de problèmes semble presque indépendant de la taille du problème considéré, ce qui est en contradiction avec la borne théorique en  $\sqrt{n}$ .

Pour conclure, mentionnons le fait que, contrairement à l'algorithme simplexe, la plupart des méthodes de point intérieur peuvent être adaptées à la résolution de certains problèmes de programmation non linéaire. Mentionnons les deux domaines suivants :

- La *programmation semidéfinie* est une généralisation prometteuse de la programmation linéaire où la contrainte de positivité sur le vecteur des variables  $x \geq 0$  est remplacée par une contrainte stipulant qu'une matrice symétrique de variables  $X$  est semidéfinie positive. Ce type de problème admet de nombreuses applications dans des domaines tels que l'automatique et le contrôle de systèmes, l'optimisation des structures, l'optimisation combinatoire, etc. (voir [22] pour un aperçu). Les méthodes que nous avons présentées peuvent être adaptées au cas de la programmation semidéfinie avec relativement peu d'efforts et plusieurs logiciels de résolution permettent de résoudre ces problèmes efficacement en pratique (voir par exemple [18]).
- Nesterov et Nemirovski présentent dans la monographie [15] une théorie complète décrivant une classe de méthodes de point intérieur applicables à l'ensemble des problèmes d'*optimisation convexe*. Ils démontrent le caractère polynomial de la complexité algorithmique de ces algorithmes et font le lien entre leur efficacité et l'existence d'un certain type de fonction barrière appelé fonction barrière *auto-concordante*.

# Bibliographie

- [1] E. D. Andersen, J. Gondzio, Cs. Mészáros, and X. Xu. Implementation of interior-point methods for large scale linear programs. In T. Terlaky, editor, *Interior Point Methods of Mathematical Programming*, volume 5 of *Applied Optimization*, pages 189–252. Kluwer Academic Publishers, 1996.
- [2] K. M. Anstreicher. On long step path following and SUMT for linear and quadratic programming. Technical report, Yale School of Management, Yale University, New Haven, CT, 1990.
- [3] K. M. Anstreicher. Potential reduction algorithms. In T. Terlaky, editor, *Interior Point Methods of Mathematical Programming*, volume 5 of *Applied Optimization*, pages 125–158. Kluwer Academic Publishers, 1996.
- [4] G. B. Dantzig. *Linear programming and extensions*. Princeton University Press, Princeton, N.J., 1963.
- [5] I. I. Dikin. Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk SSSR*, 174:747–748, 1967.
- [6] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York, 1968. Reprinted in *SIAM Classics in Applied Mathematics*, SIAM Publications, 1990.
- [7] K. R. Frisch. The logarithmic potential method of convex programming. Technical report, University Institute of Economics, Oslo, Norway, 1955.
- [8] P. Huard. Resolution of mathematical programming with nonlinear constraints by the method of centers. In J. Abadie, editor, *Nonlinear Programming*, pages 207–219. North Holland, Amsterdam, The Netherlands, 1967.
- [9] N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [10] L. G. Khachiyan. A polynomial algorithm in linear programming. *Soviet Mathematics Doklady*, 20:191–194, 1979.
- [11] V. Klee and G. J. Minty. *How good is the simplex algorithm ?*, pages 159–175. Inequalities, O. Shisha ed. Academic Press, New York, 1972.
- [12] O. L. Mangasarian. *Nonlinear Programming*. McGraw-Hill, New York, 1969.
- [13] W. F. Mascarenhas. The affine scaling algorithm fails for  $\lambda = 0.999$ . Technical report, Universidade Estadual de Campinas, Campinas S. P., Brazil, October 1993.
- [14] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2:575–601, 1992.

- [15] Y. E. Nesterov and A. S. Nemirovski. *Interior-point polynomial methods in convex programming*. SIAM Studies in Applied Mathematics. SIAM Publications, Philadelphia, 1994.
- [16] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization. An Interior Point Approach*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Chichester, UK, 1997.
- [17] N. Z. Shor. Utilization of the operation of space dilatation in the minimization of convex functions. *Kibernetika*, 1:6–12, 1970.
- [18] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11-12:625–653, 1999. Special issue on Interior Point Methods (CD supplement with software).
- [19] K. Tanabe. Centered newton method for mathematical programming. In M. Iri and K. Yajima, editors, *System Modeling and Optimization*, volume 113 of *Lecture Notes in Control and Information Sciences*, pages 197–206. Springer, New York, 1988.
- [20] M. J. Todd and Y. Ye. A centered projective algorithm for linear programming. *Mathematics of Operations Research*, 15:508–529, 1990.
- [21] T. Tsuchiya. Affine scaling algorithm. In T. Terlaky, editor, *Interior Point Methods of Mathematical Programming*, volume 5 of *Applied Optimization*, pages 35–82. Kluwer Academic Publishers, 1996.
- [22] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.
- [23] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, 1997.
- [24] Y. Ye. *Interior Point Algorithms, Theory and Analysis*. John Wiley & Sons, Chichester, UK, 1997.
- [25] Y. Ye, M. J. Todd, and S. Mizuno. An  $O(\sqrt{n}L)$ -iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19:53–67, 1994.