

LES EFFETS DE L'EXPOSANT DE LA FONCTION BARRIÈRE MULTIPLICATIVE DANS LES MÉTHODES DE POINTS INTÉRIEURS

ADAMA COULIBALY¹ ET JEAN-PIERRE CROUZEIX²

Abstract. Potential functions in interior point methods are used to determine descent directions and to prove the convergence. They depend on a parameter which is usually taken equal to or greater than the size of the problem. Actually, smaller values give a better conditioning of the method near an optimal solution. This assertion is illustrated by a few numerical experiments.

Résumé. Les méthodes de points intérieurs en programmation linéaire connaissent un grand succès depuis l'introduction de l'algorithme de Karmarkar. La convergence de l'algorithme repose sur une fonction potentielle qui, sous sa forme multiplicative, fait apparaître un exposant p . Cet exposant est, de façon générale, choisi supérieur au nombre de variables n du problème. Nous montrons dans cet article que l'on peut utiliser des valeurs de p plus petites que n . Ceci permet d'améliorer le conditionnement de la méthode au voisinage de la solution optimale.

Mots Clés. Interior point methods, Karmarkar algorithm, multiplicative and additive potential functions, barrier function.

Classification Mathématique. 49M35, 90C05, 26B25.

Reçu en avril 2003.

¹ UFR Math.-Info., Université de COCODY, BP 582, Abidjan 22, Côte d'Ivoire.

² CUST et LIMOS, Université Blaise Pascal, Campus des Cézaux, 63174 Aubière, France.

1. INTRODUCTION ET NOTATIONS

On s'intéresse au problème de programmation linéaire

$$\alpha = \min_x [\langle c, x \rangle : x \in \mathbb{R}^n, x \geq 0, Ax = a] \quad (LP)$$

où $a \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ et A est une matrice réelle $m \times n$ de rang m . On pose

$$\begin{aligned} S &= \{x \in \mathbb{R}^n : x \geq 0, Ax = a\}, \\ \tilde{S} &= \{x \in \mathbb{R}^n : x > 0, Ax = a\}, \\ S_{\text{opt}} &= \{x \in \mathbb{R}^n : x \geq 0, Ax = a, \langle c, x \rangle = \alpha\}. \end{aligned}$$

La méthode du simplexe permet de résoudre ce problème en un nombre fini d'opérations élémentaires, mais ce nombre d'opérations augmente de façon exponentielle avec la taille du problème. Une alternative consiste à utiliser une méthode de points intérieurs. Dans ce travail, on considère le problème

$$\inf [f_p(x) : x \in S] \quad (P_m)$$

où la fonction f_p , $p > 0$, appelée *fonction barrière multiplicative*, est définie sur \tilde{S} par

$$f_p(x) = (\langle c, x \rangle - \alpha)^p \prod_{i=1}^n x_i^{-1}$$

et est prolongée par semi-continuité inférieure sur S . On peut tout autant, et de façon totalement équivalente, considérer la *fonction barrière logarithmique*

$$g_p(x) = \ln(f_p(x)) = p \ln(\langle c, x \rangle - \alpha) - \sum_{i=1}^n \ln(x_i^{-1})$$

et le problème associé

$$\inf [g_p(x) : x \in \tilde{S}]. \quad (P_l)$$

On sait que les fonctions f_p et g_p sont pseudoconvexes sur S dès que $p \geq n$ dans le cas général et dès que $p \geq n - 1$ lorsque S est compact. En outre f_p est convexe sur S dès que $p \geq n + 1$ dans le cas général et dès que $p \geq n$ lorsque S est compact [3, 4, 10, 12].

En raison du comportement de leurs fonctions objectives sur la frontière relative de S , les problèmes (P_m) et (P_l) peuvent être considérés comme des problèmes d'optimisation sans contraintes; on peut alors utiliser une méthode de descente, par exemple la méthode de Newton. Au voisinage d'une solution optimale on se trouve confronté à un problème de conditionnement, de sorte qu'il est difficile d'approcher cette solution optimale. Il ressort des expériences numériques que nous avons effectuées que le conditionnement s'améliore lorsque p diminue, d'où

l'intérêt de choisir p le plus petit possible sans pourtant perdre les propriétés de convexité des fonctions f_p et g_p . C'est l'objet de cet article.

Dans les parties 2 et 3, nous nous intéressons tout d'abord aux valeurs de p pour lesquelles les problèmes (LP) , (P_m) et (P_l) admettent les mêmes solutions optimales, ensuite aux valeurs pour lesquelles les seuls points stationnaires de (P_m) ou (P_l) sont les solutions optimales des problèmes, puis aux valeurs de p pour lesquelles les fonctions f_p et g_p sont convexes ou pseudoconvexes.

Nous étudions ensuite, dans la partie 4, les effets du paramètre p sur la direction de Newton, puis, dans la partie 5, le comportement de la méthode de Newton au voisinage de la solution optimale lorsque celle-ci est unique. Enfin, dans la dernière partie, nous présentons des expériences numériques.

Les notations utilisées dans cet article sont les suivantes :

$\langle \cdot, \cdot \rangle$, $\|\cdot\|$, $\|\cdot\|_\infty$ désignent respectivement le produit scalaire usuel, la norme euclidienne et la norme du maximum de \mathbb{R}^n . Étant donnés $x, y \in \mathbb{R}^n$, on dira que $x \leq y$ (respectivement $x < y$) si et seulement si pour tout $i \in \{1, \dots, n\}$, $x_i \leq y_i$ (respectivement $x_i < y_i$). Étant donné $x \in \mathbb{R}^n$, on désigne par $X = \text{diag}(x)$ la matrice diagonale d'ordre n définie par $X_{ii} = x_i$ pour tout i . Pour $x \in S$, on note $I(x)$ l'ensemble

$$I(x) = \{i \in \{1, \dots, n\} : x_i = 0\}.$$

I désigne la matrice identité et e le vecteur dont toutes les composantes sont égales à 1. S'il est besoin de faire apparaître la dimension k , on utilisera les notations I_k et e_k .

Étant donnée une matrice carrée symétrique réelle M d'ordre n , l'inertie de M notée $\text{Inertie}(M)$ est le triplet

$$\text{Inertie}(M) = (\mu_+(M), \mu_-(M), \mu_0(M)) \quad (1)$$

où $\mu_+(M)$, $\mu_-(M)$, et $\mu_0(M)$ désignent respectivement les nombres de valeurs propres de M strictement positives, strictement négatives et nulles; les valeurs propres sont comptées avec leur ordre de multiplicité et donc $\mu_+(M) + \mu_-(M) + \mu_0(M) = n$.

Soit M une matrice carrée décomposée en blocs de la façon suivante

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

les matrices A et D étant carrées et A non singulière. On appelle *complément de Schur* de A par M la matrice notée M/A définie par

$$M/A = D - CA^{-1}B. \quad (2)$$

On a alors

$$\det(M) = \det(M/A) \det(A) \quad (3)$$

et lorsque la matrice M est symétrique,

$$\text{Inertie}(M) = \text{Inertie}(M/A) + \text{Inertie}(A). \quad (4)$$

Étant donnés $x \in \mathbb{R}^n$ et A matrice $m \times n$ de rang m , on dénote par $\text{proj}(x|A) = (I - A^t(AA^t)^{-1}A)x$ la projection euclidienne de x sur le noyau de A . Pour calculer numériquement $y = \text{proj}(x|A)$ on résout le système linéaire d'ordre $n + m$

$$\begin{pmatrix} I & A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} y \\ u \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}.$$

2. LES PROBLÈMES (LP) , (P_m) ET (P_l)

Dans tout l'article, on suppose que les hypothèses suivantes sont vérifiées :

- (H0) : la matrice A est de rang m et $a \neq 0$;
- (H1) : $\tilde{S} \neq \emptyset$;
- (H2) : $S_{\text{opt}} \neq \emptyset$ et $S_{\text{opt}} \neq S$;
- (H3) : la valeur optimale α est connue.

L'hypothèse (H0) signifie que le système $Ax = a$ n'est pas redondant et que le problème n'est pas dégénéré (si $a = 0$ alors ou bien 0 est solution optimale du problème ou bien $\alpha = -\infty$).

L'hypothèse (H1) implique que \tilde{S} est l'intérieur relatif de S et que S est la fermeture de \tilde{S} . L'hypothèse (H2) implique que $S_{\text{opt}} \cap \tilde{S} = \emptyset$ et donc que S_{opt} est contenu dans la frontière relative de S . Ce sont des hypothèses classiques dans les méthodes de points intérieurs.

L'hypothèse (H3) est plus contraignante. On sait cependant qu'étant donné un problème de programmation linéaire on peut, par adjonction du problème dual, se ramener à ce cas (on a alors $\alpha = 0$). Cette opération a l'inconvénient d'augmenter la taille du problème.

Remarque 2.1. On note que, en raison des hypothèses (H0), (H1) et (H2), la $n \times (m + 1)$ matrice (A^t, c) est de rang $m + 1$.

Rappelons que nous voulons résoudre le problème (LP) au moyen du problème (P_m) . La première condition à imposer au paramètre p est d'être tel que les problèmes (LP) et (P_m) partagent les mêmes solutions optimales. Pour cela, définissons

$$\hat{p}_0 = \max [\text{card}(I(x)) : x \in S_{\text{opt}}]. \quad (5)$$

En raison des conditions d'optimalité appliquées à (LP) , on a l'inégalité $\hat{p}_0 \geq n - m + d$ dans laquelle d est la dimension du sous espace affine engendré par S_{opt} . Dans le cas particulier où S_{opt} est réduit à un point on a $\hat{p}_0 = n - m$.

Proposition 2.1. *Supposons $p > \hat{p}_0$. Alors,*

$$f_p(x) = \begin{cases} (\langle c, x \rangle - \alpha)^p \prod_{i=1}^n x_i^{-1} & \text{si } x \in \tilde{S}, \\ 0 & \text{si } x \in S_{\text{opt}}, \\ +\infty & \text{ailleurs.} \end{cases} \quad (6)$$

Il s'ensuit que les problèmes (LP) et (P_m) ont les mêmes solutions optimales.

Démonstration. On note tout d'abord que $f_p(x) > 0$ pour tout x appartenant à \tilde{S} et donc par semi-continuité inférieure $f_p(x) \geq 0$ pour tout x dans S . Il est clair que $f_p(x) = +\infty$ si x appartient à la frontière relative de S sans être solution optimale. Il reste le cas où x est solution optimale : prenons \tilde{x} arbitraire dans \tilde{S} et, pour $t \in]0, 1[$, $x_t = x + t(\tilde{x} - x)$. Alors $x_t \in \tilde{S}$ et

$$\begin{aligned} f_p(x_t) &= t^p \langle c, \tilde{x} - x \rangle^p \prod_{i=1}^n [x_i + t(\tilde{x}_i - x_i)]^{-1} \\ &= t^{p - \text{card}I(x)} \langle c, \tilde{x} - x \rangle^p \prod_{i \in I(x)} (\tilde{x}_i)^{-1} \prod_{i \notin I(x)} [x_i + t(\tilde{x}_i - x_i)]^{-1}. \end{aligned}$$

Passer à la limite lorsque $t \rightarrow 0_+$. On en déduit que $f_p(x) = 0$. \square

Si on travaille avec la fonction barrière logarithmique g_p , $f_p(x) = 0$ correspond à $g_p(x) = -\infty$.

On supposera désormais $p > \hat{p}_0$. Les solutions des problèmes (LP) , (P_m) et (P_l) étant les mêmes et les problèmes (P_m) et (P_l) pouvant être considérés sans contraintes, on pourra résoudre ces problèmes à l'aide d'une méthode de descente. De façon générale, une méthode de descente génère une suite d'itérés qui converge vers un point stationnaire. On demandera donc aux problèmes (P_m) et (P_l) de n'admettre comme points stationnaires que des solutions optimales. Rappelons qu'un point stationnaire est un point vérifiant les conditions d'optimalité du premier ordre. Les points stationnaires de (P_m) et (P_l) sont les mêmes. Introduisons les fonctions suivantes

$$\tilde{f}_p(x, t) = \frac{t^p}{\prod_{i=1}^n x_i} \quad \text{et} \quad \tilde{g}_p(x, t) = p \ln t - \sum_{i=1}^n \ln x_i \quad (7)$$

définies sur

$$\tilde{T} = \{(x, t) : x > 0, Ax = a, \langle c, x \rangle - t = \alpha\} \quad (8)$$

et prolongées par semi-continuité inférieure sur

$$T = \{(x, t) : x \geq 0, Ax = a, \langle c, x \rangle - t = \alpha\}$$

et les problèmes associés

$$\inf [f_p(x, t) : (x, t) \in T], \quad (P_{ma})$$

$$\inf [g_p(x, t) : (x, t) \in T]. \quad (P_{la})$$

Il est facile de voir que $x \in \tilde{S}$ est un point stationnaire de (P_m) et de (P_l) si et seulement $(x, t = \langle c, x \rangle - \alpha)$ est un point stationnaire de (P_{ma}) et (P_{la}) . Par définition, (x, t) est stationnaire pour ces problèmes lorsque le vecteur $\nabla g_p(x, t)$

est combinaison linéaire des $m + 1$ colonnes de la matrice

$$\begin{pmatrix} A^t & c \\ 0 & -1 \end{pmatrix}.$$

Nous avons vu dans la remarque 2.1 que ces colonnes sont linéairement indépendantes. On en déduit le résultat suivant :

Remarque 2.2. Le point $(x, t) \in \tilde{T}$ n'est pas stationnaire si et seulement si la matrice

$$\begin{pmatrix} A^t & c & -X^{-1}e \\ 0 & -1 & \frac{p}{t} \end{pmatrix}$$

est de rang $m + 2$.

On choisira p de façon à ce que cette condition soit satisfaite.

Les méthodes de descente sont particulièrement efficaces lorsque la fonction est convexe ou pseudoconvexe. Rappelons qu'une fonction différentiable f est dite *pseudoconvexe* sur le convexe C si

$$x, y \in C \text{ et } f(y) < f(x) \implies \langle \nabla f(x), y - x \rangle < 0$$

et *strictement pseudoconvexe* si

$$x, y \in C, x \neq y \text{ et } f(y) \leq f(x) \implies \langle \nabla f(x), y - x \rangle < 0.$$

Si f est pseudoconvexe et $f(y) < f(x)$, alors $y - x$ est direction de descente, en outre le long d'une direction de descente tout minimum local est global. Il est clair qu'une fonction convexe est pseudoconvexe. Nous utiliserons les caractérisations du second ordre de convexité et pseudoconvexité sur un sous-espace affine suivantes.

Proposition 2.2 [3]. Soient C un convexe ouvert de \mathbb{R}^n , $b \in \mathbb{R}^m$, B une matrice $m \times n$ de rang m , $D = \{x \in C : Bx = b\}$ et $f : C \rightarrow \mathbb{R}$ deux fois différentiable sur D n'admettant aucun point stationnaire sur D . Alors,

a) f est convexe sur D si et seulement si

$$x \in D \text{ et } Bh = 0 \implies \langle \nabla^2 f(x)h, h \rangle \geq 0 ;$$

b) si $\langle \nabla^2 f(x)h, h \rangle > 0$ pour tout $x \in D$ et $h \neq 0$ tels que $Bh = 0$ alors f est strictement convexe sur D . Si en outre $\nabla^2 f$ est continue en x , f est fortement convexe sur D dans un voisinage de x ;

c) f est pseudoconvexe sur D si et seulement si

$$x \in D, \langle \nabla f(x), h \rangle = 0 \text{ et } Bh = 0 \implies \langle \nabla^2 f(x)h, h \rangle \geq 0 ;$$

d) si $\langle \nabla^2 f(x)h, h \rangle > 0$ pour tout $x \in D$ et $h \neq 0$ tels que $\langle \nabla f(x), h \rangle = 0$ et $Bh = 0$ alors f est strictement pseudoconvexe sur D .

On est ainsi amené à étudier la semi-définie positivité d'une forme quadratique sur un sous-espace affine, on utilise pour cela la proposition suivante.

Proposition 2.3 [1]. *Soit la matrice suivante*

$$M = \begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix}$$

où A est une matrice symétrique $n \times n$ et B est une matrice $m \times n$ de rang m . On a les résultats suivants :

- $\mu_+(M) \geq m$ et $\mu_-(M) \geq m$;
- $(Bh = 0 \Rightarrow \langle Ah, h \rangle \geq 0) \iff \mu_-(M) = m$;
- $(Bh = 0, h \neq 0 \Rightarrow \langle Ah, h \rangle > 0) \iff \mu_+(M) = n$.

Il est clair que f_p , g_p et \tilde{f}_p sont pseudoconvexes si et seulement si \tilde{g}_p l'est. En raison des deux propositions ci-dessus on est amené à étudier, pour tout $(x, t) \in \tilde{T}$, l'inertie de la matrice suivante

$$N(p, x, t) = \begin{pmatrix} X^{-2} & 0 & A^t & c & -X^{-1}e \\ 0 & -\frac{p}{t^2} & 0 & -1 & \frac{p}{t} \\ A & 0 & 0 & 0 & 0 \\ c^t & -1 & 0 & 0 & 0 \\ -e^t X^{-1} & \frac{p}{t} & 0 & 0 & 0 \end{pmatrix}.$$

Notons que, en raison de la remarque 2.2, si (x, t) est un point stationnaire alors $\det(N(p, x, t)) = 0$ et si (x, t) n'est pas stationnaire alors $\mu_-(N(p, x, t)) \geq m + 2$. La matrice $N(p, x, t)$ a la même inertie que la matrice suivante

$$\begin{pmatrix} I & 0 & XA^t & Xc & -e \\ 0 & -\frac{1}{p} & 0 & -\frac{t}{p} & 1 \\ AX & 0 & 0 & 0 & 0 \\ c^t X & -\frac{t}{p} & 0 & 0 & 0 \\ -e^t & 1 & 0 & 0 & 0 \end{pmatrix}.$$

En prenant le complément de Schur de cette matrice par la matrice $\begin{pmatrix} -\frac{1}{p} & 1 \\ 1 & 0 \end{pmatrix}$ on voit que la condition suffisante de stricte pseudoconvexité en (x, t) tient si et seulement si la matrice

$$\hat{N}(p, x, t) = \begin{pmatrix} I - \frac{1}{p}ee^t & XA^t & Xc - \frac{t}{p}e \\ AX & 0 & 0 \\ c^t X - \frac{t}{p}e^t & 0 & 0 \end{pmatrix}$$

a exactement n valeurs propres strictement positives, cette condition est vérifiée dès que $p > n$ car la matrice $I - \frac{1}{p}ee^t$ est alors définie positive. Posons $P_{x,t}(p) = (-1)^{m+1} \det(\hat{N}(p, x, t))$, $P_{x,t}$ est un polynôme de degré 2 en $\frac{1}{p}$, la condition de stricte pseudoconvexité de la proposition 2.2 en (x, t) est $P_{x,t}(p) > 0$.

La fonction f_p est convexe sur S si et seulement si \tilde{f}_p l'est sur T . La condition suffisante de forte convexité de la fonction \tilde{f}_p en (x, t) est reliée à l'inertie de la matrice suivante

$$M(p, x, t) = \begin{pmatrix} X^{-1}(I + ee^t)X^{-1} & -\frac{p}{t}X^{-1}e & A^t & c \\ -\frac{p}{t}e^tX^{-1} & -\frac{p^2}{t^2} + \frac{p}{t} & 0 & -1 \\ A & 0 & 0 & 0 \\ c^t & -1 & 0 & 0 \end{pmatrix}$$

qui a même inertie que la matrice suivante

$$\tilde{M}(p, x, t) = \begin{pmatrix} I + ee^t & -e & XA^t & Xc \\ -e^t & 1 - \frac{1}{p} & 0 & -\frac{t}{p} \\ AX & 0 & 0 & 0 \\ c^tX & -\frac{t}{p} & 0 & 0 \end{pmatrix}.$$

Posons $Q_{x,t}(p) = (-1)^{m+1} \det(\tilde{M}(p, x, t))$, $Q_{x,t}$ est un polynôme de degré 2 en $\frac{1}{p}$. Puisque la matrice $\begin{pmatrix} A^t & c \\ 0 & -1 \end{pmatrix}$ est de rang $m+1$, la condition de forte convexité de \tilde{f}_p en (x, t) est donc équivalente à $Q_{x,t}(p) > 0$.

En prenant le complément de Schur de la matrice $\tilde{M}(p+1, x, t)$ par la matrice d'ordre 1 $\frac{p}{p+1}$, on voit que la condition de forte convexité en (x, t) pour \tilde{f}_{p+1} tient si et seulement si la matrice

$$\hat{M}(p, x, t) = \begin{pmatrix} I - \frac{1}{p}ee^t & XA^t & Xc - \frac{t}{p}e \\ AX & 0 & 0 \\ c^tX - \frac{t}{p}e^t & 0 & -\frac{t^2}{p(p+1)} \end{pmatrix}$$

a exactement n valeurs propres strictement positives. On en déduit la proposition suivante :

Proposition 2.4. a) Si $p \geq n$ alors la fonction f_{p+1} est convexe sur S .

b) Si la fonction f_{p+1} est convexe sur S , alors les fonctions f_p et g_p sont pseudoconvexes sur S .

Démonstration. a) Si $k > n$ alors la matrice $I - \frac{1}{k}ee^t$ est définie positive, la fonction \tilde{f}_{k+1} est strictement convexe sur \tilde{T} , passer à la limite lorsque k tend vers p .

b) Il est clair que $\mu_+(\hat{N}(p, x, t)) \geq \mu_+(\hat{M}(p, x, t))$. \square

Nous avons vu que les polynômes $P_{x,t}$ et $Q_{x,t}$ sont de degré 2 en $\frac{1}{p}$ et sont strictement positifs dès que $p > n$ pour le premier et $p > n + 1$ pour le second. Il est donc possible de déterminer, en fonction des racines éventuelles de ces polynômes, les quantités minimales $p_1(x, t)$ et $p_2(x, t)$ telles que les conditions de stricte pseudoconvexité et stricte convexité tiennent pour $p > p_1(x, t)$ et $p > p_2(x, t)$ respectivement. On pose ensuite

$$\hat{p}_1 = \sup \left[p_1(x, t) : (x, t) \in \tilde{T} \right] \text{ et } \hat{p}_2 = \sup \left[p_2(x, t) : (x, t) \in \tilde{T} \right].$$

Proposition 2.5. *Si $p > \hat{p}_1$, les fonctions f_p et g_p sont strictement pseudoconvexes sur \tilde{S} , si $p > \hat{p}_2$, la fonction f_p est strictement convexe sur \tilde{S} .*

S'il est possible de calculer les coefficients des polynômes $P_{x,t}$ et $Q_{x,t}$ et par là les valeurs $p_1(x, t)$ et $p_2(x, t)$, on peut chercher des majorations plus commodes de $p_1(x, t)$ et $p_2(x, t)$. Pour cela considérons la restriction de la fonction

$$h_p(x, t) = t^p \Pi x_i^{-1} \quad t > 0, \quad x > 0$$

au sous-espace affine $\{(x, t) : Ax = a\}$. La convexité de cette fonction est liée à l'inertie de la matrice

$$Q(p, x, t) = \begin{pmatrix} X^{-1}(I + ee^t)X^{-1} & pt^{-1}X^{-1}e & A^t \\ pt^{-1}e^tX^{-1} & (p^2 - p)t^{-2} & 0 \\ A & 0 & 0 \end{pmatrix}$$

qui a même inertie que la matrice

$$\tilde{Q}(p, x, t) = \begin{pmatrix} I + ee^t & e & XA^t \\ e^t & 1 - \frac{1}{p} & 0 \\ AX & 0 & 0 \end{pmatrix}.$$

La condition de forte convexité en (x, t) tient si $\mu_+(\tilde{Q}(p, x, t)) = n + 1$ soit encore, puisque la matrice AX est de rang m , si

$$\delta = (-1)^m \det(\tilde{Q}(p, x, t)) > 0.$$

Il est facile de voir que

$$\begin{aligned} (-1)^m \delta &= \det \begin{pmatrix} I + ee^t & 0 & XA^t \\ 0 & 1 & 0 \\ AX & 0 & 0 \end{pmatrix} - \frac{1}{p} \det \begin{pmatrix} I + ee^t & 0 & XA^t \\ 0 & 1 & 0 \\ AX & 0 & 0 \end{pmatrix} \\ &= \det \begin{pmatrix} I & XA^t \\ AX & 0 \end{pmatrix} - \frac{1}{p} \det \begin{pmatrix} I + ee^t & XA^t \\ AX & 0 \end{pmatrix}. \end{aligned}$$

On en déduit que la condition tient si et seulement si

$$p > (n+1) \det \left(AX^2 A^t - \frac{1}{n+1} aa^t \right) \det [(AX^2 A^t)^{-1}].$$

Soit encore, en considérant la matrice symétrique définie positive B telle que $B^2 AX^2 A^t = I$, si

$$p > (n+1) \det \left(I - \frac{1}{n+1} B^{-1} aa^t B^{-1} \right) = n+1 - \langle (AX^2 A^t)^{-1} a, a \rangle.$$

On pose pour tout $x > 0$

$$p_3(x) = n+1 - \langle (AX^2 A^t)^{-1} a, a \rangle \quad (9)$$

et

$$\hat{p}_3 = \sup \left[p_3(x) : x \in \tilde{S} \right].$$

La fonction f_p est alors strictement convexe sur \tilde{S} dès que $p > \hat{p}_3$. En raison de la proposition 2.4, la fonction f_p est strictement pseudoconvexe sur \tilde{S} dès que $p > \hat{p}_3 - 1$.

La détermination de $p_3(x)$ est donc beaucoup plus facile que celles de $p_1(x, t)$ et $p_2(x, t)$. Nous allons étudier cette fonction p_3 dans la section suivante.

3. QUELQUES PROPRIÉTÉS DE LA FONCTION p_3

La fonction $r(x) = \langle (AX^2 A^t)^{-1} a, a \rangle$ a été étudiée par Crouzeix–Kebbour [4]. Nous complétons cette étude ci-dessous.

Soit $x \in \tilde{S}$, on sait que la matrice AX est une matrice $m \times n$ de rang m . On posera

$$P_x = I - XA^t(AX^2 A^t)^{-1}AX$$

P_x est la matrice projection sur le noyau de la matrice AX . On a les relations suivantes :

$$p_3(x) = 1 + \|P_x e\|^2 \quad (10)$$

$$= n+1 - \inf_y [\|y\|^2 : AX(y - e) = 0] \quad (11)$$

$$= n+1 + \inf_v [\|XA^t v\|^2 - 2\langle a, v \rangle] \quad (12)$$

$$= 1 + \inf_u [\|XA^t u - e\|^2] \quad (13)$$

$$= n+1 - \sup_u \left[\frac{\langle XA^t u, e \rangle^2}{\|XA^t u\|^2} : u \neq 0 \right]. \quad (14)$$

L'équivalence entre (9) et (10) vient de la relation $AXe = Ax = a$, en résolvant les problèmes d'optimisation on obtient l'équivalence entre (9, 11) et (12), l'équivalence

entre (12) et (14) est obtenue en faisant $v = tu$, $t \in \mathbb{R}$ dans (12) et en prenant l'infimum par rapport à t , enfin (13) est une réécriture de (12).

À partir de (12) on voit que la fonction p_3 est définie non seulement sur \tilde{S} mais aussi sur tout l'espace \mathbb{R}^n . En outre, on a les propriétés suivantes :

- Proposition 3.1.** a) p_3 est semi-continue supérieurement sur \mathbb{R}^n ;
 b) $1 < p_3(x) < n + 1$ pour tout $x \in \mathbb{R}^n$;
 c) si S est compact, alors $p_3(x) < n$ pour tout $x \in \tilde{S}$;
 d) la fonction p_3 est différentiable sur \tilde{S} et on a

$$\nabla p_3(x) = -2X^{-1}\text{diag}(P_x e)(e - P_x e).$$

Démonstration. a) Pour tout u , la fonction $x \rightarrow \|XA^t u - e\|^2$ est continue en x . En raison de (12), la fonction p_3 est semi-continue supérieurement puisque elle est l'infimum de fonctions semi-continues supérieurement. b) On ne peut avoir $P_x e = 0$ car on aurait alors $0 = AXe = Ax = a \neq 0$ et donc $1 < p_3(x)$. De même, on ne peut avoir $y = 0$ dans (11) et donc $p_3(x) < n + 1$. Fixons $x \in \tilde{S}$, $\sup\{\langle X^{-1}e, y \rangle : y \in S\}$ est fini car S est compact. Par dualité en programmation linéaire, il existe $u \in \mathbb{R}^m$ tel que $A^t u \geq X^{-1}e > 0$. Posons $w = XA^t u$, alors $w \geq e > 0$. Puisque $n^{-1} < \|h\|^2 \leq 1$ pour tout $h > 0$ tel que $\langle h, e \rangle = 1$, (14) implique $p_3(x) < n$. La fonction $\varphi(x, u) = 2^{-1}\|XA^t u - e\|^2 = 2^{-1}\|\text{diag}(A^t u)x - e\|^2$ est différentiable à tout ordre sur $(0, \infty)^n \times \mathbb{R}^m$. $x \in \tilde{S}$ étant fixé, elle atteint son minimum au point $u(x) = (AX^2 A^t)^{-1}AX$. Il s'ensuit que

$$\begin{aligned} 2^{-1}\nabla p_3(x) &= \nabla \varphi_x(x, u(x)) \\ &= \text{diag}(A^t u(x))[\text{diag}(A^t u(x))x - e] \\ &= X^{-1}\text{diag}(XA^t u(x))[XA^t u(x) - e] \\ &= -X^{-1}\text{diag}(e - P_x e)P_x e \\ &= -X^{-1}\text{diag}(P_x e)(e - P_x e). \end{aligned}$$

En effet $XA^t u(x) = e - P_x e$. □

Remarque 3.1. Soit $x \in \tilde{S}$, la restriction de p_3 au sous espace $\{y : Ay = a\}$ est stationnaire en x si et seulement si $\langle \nabla p_3(x), Xh \rangle = 0$ pour tout h tel que $AXh = 0$. Cela signifie que $P_x(X\nabla p_3(x)) = 0$, c'est-à-dire $P_x(\text{diag}(P_x e)(e - P_x e)) = 0$.

Remarque 3.2. On peut voir à partir de la démonstration qu'en fait p_3 est différentiable en tout $x \in \mathbb{R}^n$ tel que AX est de rang m .

Intéressons nous maintenant au comportement de la fonction p_3 aux points de la frontière relative de S .

Proposition 3.2. Soit $x \in S/\tilde{S}$. Posons $q(x) = n - \text{card}(I(x))$. Alors $1 < p_3(x) \leq q(x) + 1$. Si en outre S est compact, $p_3(x) \leq q(x)$.

Démonstration. Posons $J(x) = \{i = 1, 2, \dots, n : i \notin J\}$, B et C les matrices obtenues à partir de A en ne retenant respectivement que les colonnes correspondant aux indices appartenant à $I(x)$ et $J(x)$, de la même façon, les vecteurs y et z sont obtenus à partir du vecteur x . On a alors $x = (y, z)$ et $Ax = By + Cz$. La fonction $h_p(z) = t^p \Pi z_k^{-1}$ est convexe dans un voisinage de x sur le sous espace $\{z : z > 0, Cz = a\}$ dès que $p \geq q(x) + 1$ en raison du b) de la proposition précédente et $p \geq q(x)$ lorsque $\{z : z \geq 0, Cz = a\}$ est compact, ce qui est le cas lorsque S est compact. \square

Corollaire 3.1. *Si S est compact alors $\hat{p}_3 < n$.*

Démonstration. Puisque la fonction p_3 est semi-continue supérieurement sur S , son maximum sur S est atteint. On a vu que sur \tilde{S} et S/\tilde{S} on a $p_3(x) < n$. \square

La fonction p_3 dépend de l'ensemble S et non du vecteur c . On sait que S est compact si et seulement si

$$d \geq 0 \text{ et } Ad = 0 \implies d = 0.$$

Lorsque S est non compact, on a le résultat suivant.

Proposition 3.3. *Si S est non compact, alors $n \leq \hat{p}_3 \leq n + 1$.*

Démonstration. Prenons $c = e$ et supposons que la fonction f_p est convexe avec, pour contradiction, $p < n$. On note que S_{opt} est alors non vide et compact puisque

$$d \geq 0, \langle c, d \rangle \leq 0 \text{ et } Ad = 0 \implies d = 0.$$

Il s'ensuit que la fonction f_p est inf-compacte. D'autre part, il existe $\hat{d} \geq 0$ tel que $A\hat{d} = 0$ et $\langle e, \hat{d} \rangle > 0$. Soit $\bar{x} \in \tilde{S}$, on a pour tout $t > 0$

$$f_p(\bar{x} + t\hat{d}) = \left(\langle c, \bar{x} + t\hat{d} \rangle - \alpha \right)^p \Pi(\bar{x}_i + t\hat{d}_i)^{-1}.$$

Lorsque p tend vers l'infini, $f_p(\bar{x} + t\hat{d})$ tend vers 0 et puisque f_p est convexe, $f_p(\bar{x} + t\hat{d}) \leq f_p(\bar{x})$ pour tout $t \geq 0$ et ainsi f_p est non inf-compacte. De là, $\hat{p}_3 \geq n$ puisque f_p est convexe pour tout $p \geq \hat{p}_3$. \square

4. INFLUENCE DU FACTEUR p SUR LA DIRECTION DE NEWTON

Une des méthodes les plus efficaces en optimisation sans contraintes pour minimiser une fonction f deux fois différentiable est la méthode de Newton. Elle s'applique également lorsque l'ensemble réalisable est un sous-espace affine. Dans le cas présent du problème (P_m) , elle consiste, l'itéré $x_k > 0$ tel que $Ax_k = a$ ayant

été obtenu lors de l'itération précédente, à calculer la direction d_k en résolvant le problème d'optimisation quadratique

$$\min_d \left[\frac{1}{2} \langle \nabla^2 f_p(x_k) d, d \rangle + \langle \nabla f_p(x_k), d \rangle : Ad = a \right]$$

puis à déterminer t_k solution du problème

$$\min_t [f_p(x_k + t \|d_k\|^{-1} d_k) : t \in \mathbb{R}],$$

on prend ensuite $x_{k+1} = x_k + t_k d_k$.

Pour obtenir d_k , on résout le système linéaire

$$\begin{pmatrix} \nabla^2 f_p(x_k) & A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} d_k \\ v_k \end{pmatrix} = \begin{pmatrix} -\nabla f_p(x_k) \\ 0 \end{pmatrix}. \quad (15)$$

Ce système a une solution et une seule dès lors que la restriction de f_p au sous espace affine est fortement convexe dans un voisinage de x_k ce qui est le cas lorsque $p > p_3(x)$. Rappelons que $f_p(x) = \exp(g_p(x))$ et donc

$$\frac{1}{f_p(x)} \nabla f_p(x) = \nabla g_p(x) \quad \text{et} \quad \frac{1}{f_p(x)} \nabla^2 f_p(x) = \nabla^2 g_p(x) + \nabla^t g_p(x) \nabla g_p(x).$$

Considérons le système

$$\begin{pmatrix} \nabla^2 g_p(x_k) & A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} \delta_k \\ u_k \end{pmatrix} = \begin{pmatrix} -\nabla g_p(x_k) \\ 0 \end{pmatrix}. \quad (16)$$

Les solutions des systèmes (15) et (16) sont clairement reliées. Il est en effet facile de voir que $(1 - \langle \nabla g_p(x_k), \delta_k \rangle) d_k = \delta_k$. Le deuxième système est plus facile à résoudre que le premier. Puisque d_k et δ_k sont colinéaires, t_k est aussi solution du problème

$$\min_t [g_p(x_k + t \|\delta_k\|^{-1} \delta_k) : t \in \mathbb{R}].$$

Il est à noter que d_k et δ_k peuvent être de signe contraire et donc la direction δ_k peut être une direction de montée et non de descente en x_k pour les fonctions f_p et g_p , ce sera en fait le cas ici. Notons que l'on a

$$\begin{aligned} \langle \nabla g_p(x_k), \delta_k \rangle &= \left\langle \begin{pmatrix} \nabla g_p(x_k) \\ 0 \end{pmatrix}, \begin{pmatrix} \delta_k \\ u_k \end{pmatrix} \right\rangle \\ &= - \left\langle \begin{pmatrix} \nabla^2 g_p(x_k) & A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} \delta_k \\ u_k \end{pmatrix}, \begin{pmatrix} \delta_k \\ u_k \end{pmatrix} \right\rangle \\ &= - \langle \nabla^2 g_p(x_k) \delta_k, \delta_k \rangle \end{aligned}$$

de même, on a $\langle \nabla f_p(x_k), d_k \rangle = - \langle \nabla^2 f_p(x_k) d_k, d_k \rangle$ ce qui est conforme au fait que d_k correspond à la direction de Newton en x_k pour f_p .

En remplaçant dans (16) $\nabla g_p(x_k)$ et $\nabla^2 g_p(x_k)$ par leurs valeurs on obtient le système

$$\begin{pmatrix} I - p\hat{c}_k\hat{c}_k^t & X_k A^t \\ AX_k & 0 \end{pmatrix} \begin{pmatrix} X_k^{-1}\delta_k \\ u_k \end{pmatrix} = \begin{pmatrix} e - p\hat{c}_k \\ 0 \end{pmatrix} \quad (17)$$

où $\hat{c}_k = (\langle c, x_k \rangle - \alpha)^{-1} X_k c$.

Il en résulte qu'il existe λ tel que

$$X_k^{-1}\delta_k = P_k e + \lambda P_k \hat{c}_k$$

où $P_k = I - X_k A^t (A X_k^2 A^t)^{-1} A X_k$ désigne la matrice projection sur le noyau de la matrice $A X_k$. Finalement, on obtient

$$\begin{aligned} X_k^{-1}\delta_k &= P_k e + p \frac{1 - \langle P_k \hat{c}_k, e \rangle}{p \|P_k \hat{c}_k\|^2 - 1} P_k \hat{c}_k \\ &= \left[P_k e + \frac{1 - \langle P_k \hat{c}_k, e \rangle}{\|P_k \hat{c}_k\|^2} P_k \hat{c}_k \right] + \frac{1 - \langle P_k \hat{c}_k, e \rangle}{\|P_k \hat{c}_k\|^2 (p \|P_k \hat{c}_k\|^2 - 1)} P_k \hat{c}_k. \end{aligned}$$

Il est intéressant de comparer cette direction δ_k avec la direction du gradient projeté

$$\hat{\delta}_k = P X_k^{-1} [p\hat{c}_k - e]$$

où $P = I - A^t (A A^t)^{-1} A$ désigne la matrice projection sur le noyau de la matrice A . On voit que p a un effet multiplicatif sur le vecteur \hat{c}_k pour $\hat{\delta}_k$ et un effet contraire pour δ_k .

5. LE COMPORTEMENT PRÈS D'UNE SOLUTION OPTIMALE

Dans cette partie, nous supposons que \bar{x} est la solution unique du problème (LP) et que la matrice $A\bar{X}$ est de rang m . Nous allons montrer que la direction de Newton en un point $x = \bar{x} + h \in \tilde{S}$ proche de \bar{x} est très proche de h . Ce sera donc une très bonne direction.

Puisque $A\bar{X}$ est de rang m et \bar{x} est solution optimale, il résulte des conditions de complémentarité que \bar{x} a exactement m composantes strictement positives, en outre les m colonnes de A correspondantes sont linéairement indépendantes. Nous supposons que l'on a $p > n - m$, nous savons alors qu'en raison des propositions 2.4 et 3.2, les fonctions f_p et g_p sont strictement pseudoconvexes dans un voisinage de \bar{x} .

Sans perte de généralité, on suppose que les $n - m$ premières composantes de \bar{x} sont nulles. Il existe une matrice $R = (I_{n-m}, M)$ avec M matrice $(n - m) \times m$ telle que

$$A(x - \bar{x}) = 0 \iff \exists y \in \mathbb{R}^{n-m} \text{ tel que } x = \bar{x} + R^t y.$$

Lorsque $x = \bar{x} + R^t y \in \tilde{S}$ (on a alors $y > 0$), on pose

$$h_p(y) = g_p(\bar{x} + R^t y).$$

On a

$$\nabla h_p(y) = R\nabla g_p(x) \quad \text{et} \quad \nabla^2 h_p(y) = R\nabla^2 g_p(x)R^t,$$

la direction de Newton Δ en y pour la fonction h_p est reliée à la direction de Newton δ en x pour la fonction g_p par la relation $\Delta = R^t\delta$. Notons que

$$h_p(y) = p \ln \langle Rc, y \rangle - \sum_{i=1}^n \ln(\bar{x}_i + r_i^t y)$$

où r_i^t désigne la ligne i de R^t . Posons

$$\hat{c} = Rc \quad \text{et} \quad b_i = \frac{r_i}{\bar{x}_{i+n-m}} \quad \forall i = 1, \dots, m.$$

Puisque \bar{x} est solution optimale unique, on a $\hat{c} > 0$. L'équation $\nabla^2 h_p(y)\Delta = -\nabla h_p(y)$ s'écrit

$$\left(Y^{-2} + \sum \frac{b_i b_i^t}{(1 + \langle b_i, y \rangle)^2} - p \frac{\hat{c} \hat{c}^t}{\langle \hat{c}, y \rangle^2} \right) \Delta = Y^{-1}e + \sum \frac{b_i}{1 + \langle b_i, y \rangle} - p \frac{\hat{c}}{\langle \hat{c}, y \rangle}.$$

Posons k tel que $\Delta = Y(e + k)$. On a alors

$$Y^{-1}k = \sum \frac{b_i}{(1 + \langle b_i, y \rangle)^2} + p \frac{\langle Y^2 \hat{c}, Y^{-1}k \rangle}{\langle \hat{c}, y \rangle^2} \hat{c} - \sum \frac{\langle Y^2 b_i, Y^{-1}k \rangle}{(1 + \langle b_i, y \rangle)^2} b_i.$$

On note que le vecteur $Y^{-1}k \in \mathbb{R}^{n-m}$ est combinaison linéaire des $m+1$ vecteurs \hat{c} et b_i . Donc, il existe μ_i et λ tels que $Y^{-1}k = \sum \mu_i b_i - \lambda \hat{c}$, il n'y a pas unicité en général. On remplace $Y^{-1}k$ par sa valeur dans l'expression ci-dessus. On obtient un système linéaire de $n-m$ équations à $m+1$ inconnues. Notons que si on sait trouver μ_i et λ tels que

$$\begin{aligned} (1 + \langle b_i, y \rangle)^2 \mu_i &= 1 - \sum_j \mu_j \langle Y b_i, Y b_j \rangle + \lambda \langle Y b_i, Y \hat{c} \rangle \\ -\langle \hat{c}, y \rangle^2 \lambda &= p \sum_j \mu_j \langle Y \hat{c}, Y b_j \rangle - p \lambda \|Y \hat{c}\|^2 \end{aligned}$$

notre problème est résolu. Nous sommes en présence d'un système linéaire de $m+1$ équations à $m+1$ inconnues. On obtient

$$\lambda = \sum \frac{p \langle Y \hat{c}, Y b_i \rangle}{p \|Y \hat{c}\|^2 - \langle Y \hat{c}, e \rangle^2} \mu_i.$$

Notons que l'on a

$$p \|Y \hat{c}\|^2 - \langle Y \hat{c}, e \rangle^2 = \langle (pI - ee^t)Y \hat{c}, Y \hat{c} \rangle.$$

Les valeurs propres de la matrice $(pI - ee^t)$ sont p et $p - n + m$ et sont donc strictement positives. Il s'ensuit que

$$(p - n + m)\|Yc\|^2 \leq p\|Y\hat{c}\|^2 - \langle Y\hat{c}, e \rangle^2.$$

Pour chaque i on a

$$|\langle Y\hat{c}, Yb_i \rangle| \leq \|Y\hat{c}\| \|Yb_i\| \leq \|Y\hat{c}\| \|y\| \|b_i\|_\infty.$$

D'autre part, puisque le vecteur \hat{c} est strictement positif, $\|Y\hat{c}\| = (\sum \hat{c}_j^2 y_j^2)^{\frac{1}{2}}$ est une norme pour le vecteur y et il existe donc une constante γ dépendant uniquement de \hat{c} tel que $\|Y\hat{c}\| \geq \gamma\|y\|$. Il s'ensuit que

$$\left| \frac{p\langle Y\hat{c}, Yb_i \rangle}{p\|Y\hat{c}\|^2 - \langle Y\hat{c}, e \rangle^2} \right| \leq \frac{p\|b_i\|_\infty}{(p - n + m)\gamma}. \quad (18)$$

On reporte la valeur de λ dans les m premières équations, on obtient un système de m équations à m inconnues. En négligeant les termes du premier ordre en y , on obtient les approximations suivantes :

$$\mu_i \sim 1 \text{ pour tout } i, \quad \text{puis } \lambda \sim \hat{\lambda} = p \frac{\langle Y\hat{c}, \sum Yb_i \rangle}{p\|Y\hat{c}\|^2 - \langle Y\hat{c}, e \rangle^2}.$$

On voit que λ est borné.

On a donc $\Delta \sim y + Y^2(\sum b_i - \lambda\hat{c})$. La droite d'origine $x = \bar{x} + R^t y$ et de direction $R\Delta$ passe donc très près de la solution optimale \bar{x} et est donc une très bonne direction de descente, on voit d'après l'équation (18) que la distortion due à c diminue lorsque p diminue. Par ailleurs, dans la recherche linéaire, le numérateur $(\langle c, x \rangle - \alpha)^p$ est moins proche de 0 si on travaille avec f_p , la quantité $p \log(\langle c, x \rangle - \alpha)$ est moins proche de $-\infty$ si on travaille avec g_p , la détermination du pas est donc moins sujette à erreurs.

6. EXPÉRIENCES NUMÉRIQUES

Afin de montrer que le choix de petites valeurs pour p permet effectivement d'améliorer la précision, nous avons comparé sur des problèmes tests un algorithme, appelé ici Alg I, basé sur les résultats que nous avons exposés ci-dessus avec deux algorithmes proposés par Gonzaga, appelés ici Alg II [6] et Alg III [8].

Les problèmes tests sont issus de la bibliothèque NETLIB, ils sont de la forme

$$\beta = \min_y [\langle d, y \rangle : y \geq 0, By = b] \quad (Pr)$$

où $d \in \mathbb{R}^{n_1}$, $b \in \mathbb{R}^{n_2}$ et B est une matrice $n_2 \times n_1$. La valeur β étant inconnue, on adjoint au problème son dual

$$\beta = \max_{z, w} [\langle b, z \rangle : w \geq 0, B^t z + w = d]. \quad (D)$$

On est donc conduit à trouver $y, w \in [0, \infty[^{n_1}$ et $u, v \in [0, \infty[^{n_2}$ tels que

$$By = b, \quad B^t(u - v) + w = d \quad \text{et} \quad d^t y - b^t(u - v) = 0$$

avec $z = u - v$. Finalement le problème est mis sous la forme

$$0 = \min [\langle c, x \rangle : x \geq 0, Ax = a], \quad (LP)$$

où

$$x = \begin{pmatrix} y \\ u \\ v \\ w \\ \lambda \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad a = \begin{pmatrix} b \\ d \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} B & 0 & 0 & 0 & b - Be_1 \\ 0 & B^t & -B^t & I & d - e_1 \\ d^t & -b^t & b^t & 0 & -d^t e_1 \end{pmatrix},$$

et e_1, e_2 sont les vecteurs respectivement de \mathbb{R}^{n_1} et \mathbb{R}^{n_2} dont toutes les composantes sont égales à 1. Ici $n = 2(n_1 + n_2) + 1$ et $m = n_2 + n_1 + 1$. On note que le vecteur x_0 , où $x_0^t = (e_1^t, e_2^t, e_2^t, e_1^t, 1)$, est une solution réalisable du problème.

Pour l'algorithme Alg I, nous nous sommes donnés 2 valeurs $\rho_1, \rho_2 > 1$ (dans les expériences nous avons pris $\rho_1 = 2$ et $\rho_0 = 1, 5$). A l'itération k , x_k ayant été calculé dans l'étape précédente, nous prenons

$$p_k = \max [n - m + \rho_1, p_3(x_k) + \rho_0].$$

On fait une recherche linéaire sur la direction de Newton en x_k pour la fonction g_{p_k} . Le paramètre p_k est donc ajusté à chaque itération.

L'algorithme Alg III est le même algorithme mais avec p fixe. Nous avons considéré le choix préconisé par Gonzaga dans son article [8] $p = n + \sqrt{n}$.

L'algorithme Alg II (Gonzaga [6]) nécessite une transformation préalable. (LP) est mis sous la forme

$$0 = \min [\langle \tilde{c}, \tilde{x} \rangle : \tilde{x} \geq 0, \tilde{A}\tilde{x} = 0, \sum \tilde{x}_i = 1],$$

où $\tilde{c}^t = (c, 0)^t$ est un vecteur de \mathbb{R}^{n+1} , $\tilde{A} = (A, -a)$ est une matrice $m \times (n + 1)$. L'algorithme minimise la fonction g_p avec, compte tenu de l'adjonction d'une variable supplémentaire, $p = n + 1$.

Dans les expériences numériques sur les trois algorithmes, les itérations ont été poursuivies jusqu'à ce que, compte-tenu des erreurs d'arrondis, elles n'apportent plus de modifications dans les itérés successifs.

Nous donnons dans le tableau suivant, pour chaque algorithme et pour chaque exemple, le nombre d'itérations it qui a été nécessaire, et, afin de vérifier la conformité des solutions optimales approchées primales et duales, les quantités $nfp = \|By - b\|$ et $nfd = \|B^t z + w - d\|$. Enfin nous donnons le saut de dualité $sd = |\langle d, y \rangle - \langle b, z \rangle|$.

Ce tableau montre que le choix de petites valeurs de p permet une nette amélioration de la précision des résultats, le conditionnement près de la solution est donc bien amélioré. En contre partie, le nombre d'itérations est augmenté mais on s'est approché plus près de la solution et d'autre part il est clair que pour les premiers itérés, lorsque l'on est loin de la solution, il est préférable de donner un poids plus important à c et donc choisir un p grand.

Comparaison des 3 algorithmes.

pb	Alg I				Alg II				Alg III			
	it	nfp	nfd	sd	it	nfp	nfd	sd	it	nfp	nfd	sd
Afiro	19	2,5d-12	8,7d-15	1,d-12	20	8,41d-07	9,12d-05	2,96d-02	18	7,9d-08	4,3d-05	4,2d-03
Sc50a	25	3,d-12	1,3d-14	9,4d-12	18	6,15d-07	3,38d-04	5d-03	17	3,6d-08	9,7d-05	1, 8d-03
Sc50b	24	4,8d-12	2,6d-14	6,1d-13	16	5,3d-08	2,5d-05	2,6d-03	18	1,1d-08	9,4d-05	7 ,9d-04
Adlittle	34	2,5d-08	2,5d-08	2,9d-07	21	7d-05	1d-02	0,4	19	5,3d-06	4,7d-02	0,4
Blend	26	7,4d-12	6,7d-12	1,9d-13	16	6,6d-07	5,69d-05	1,02d-05	23	4,1d-08	8,2d-05	8,9d-06
Share2b	38	1,2d-09	1,7d-10	1,5d-10	28	2,7d-05	1,9d-02	4,68d-06	36	1d-06	8,6d-03	1,2d-05
Scagr7	45	1,1d-09	4,2d-10	4d-09	19	1,3d-03	4,8d-03	6,19	33	7,53d-06	1,35d-03	0,1 5
Sc105	34	1,8d-10	3,3d-12	6,2d-13	18	2,9d-08	2,4d-04	3d-02	29	6,2d-10	4,59d-05	6, 96d-03
Sc205	46	2,65d-07	5d-09	4,76d-10	26	3,7d-08	2,8d-03	4,6d-02	25	2,24d-09	5,45d-04	1,34d-04
Beaconfd	37	5,1d-06	1,3d-07	1,4d-07	22	1,9d-05	6,17d-02	1d-02	18	0,52	9,41d-02	3 ,75d-02
Scorpion	47	1,6d-09	4,1d-08	2,6d-07	27	9,07d-07	5,13d-04	2,54d-04	36	2,08d-08	2, 65d-03	9,4d-04
Stocfor1	25	1,9d-08	3,5d-09	2,6d-10	22	8,4d-06	1,4	5d-05	38	2,43d-08	1,96d-02	2, 49d-06
E226	58	5,9d-05	6,4d-07	2,4d-07	28	1,91d-05	9,89d-03	1,93d-05	44	6,75d-07	1,36d- 02	1,42d-05
Scsd1	21	4,4d-12	1,9d-10	7,5d-09	19	2d-07	1,1d-05	1,2d-05	30	6,97d-08	1,24d-02	1 ,4d-04

RÉFÉRENCES

- [1] Y. Chabrillac et J.-P. Crouzeix, Definiteness and semi-definiteness of quadratic forms. *Linear Algebra Appl.* **63** (1984) 283-292.
- [2] A. Coulibaly, *Méthodes de points intérieurs en programmation linéaire*, Thèse de doctorat de l'Université Blaise Pascal. Clermont-Ferrand (1994).
- [3] J.-P. Crouzeix, J.A. Ferland et S. Schaible, Generalized convexity of functions on affine subspaces and applications. *Math. Programming* **56** (1992) 223-232.
- [4] J.-P. Crouzeix et R. Kebbou, Index multiplicatifs de convexité/concavité et applications. *Cahiers Centre Études Rech. Opér.* **34** (1992) 7-24.
- [5] G. De Ghelincq et J.-Ph. Vial, A polynomial Newton method for linear programming. *Algorithmica* **1** (1989) 425-453.
- [6] C.C. Gonzaga, *A Simple presentation of Karmarkar's algorithm*, Technical Report. Department of Mathematics, Federal University of Santa Catarina, Brasil (2002).
- [7] C.C. Gonzaga, Search direction for interior linear programming methods. *Algorithmica* **6** (1991) 153-181.
- [8] C.C. Gonzaga, On lower bounds updates in primal potential reduction methods for linear programming. *Math. Programming* **52** (1991).
- [9] E.V. Haynsworth, Determination of the inertia of a partitioned hermitian matrix. *Linear Algebra Appl.* **1** (1968) 73-81.
- [10] H. Imaï, On the convexity of the multiplicative version of Karmarkar's potential function. *Math. Programming* **40** (1988) 29-32.
- [11] M. Iri et H. Imaï, A mutiplicative barrier function method for linear programming. *Algorithmica* **1** (1986) 455-482.
- [12] N.K. Karmarkar, A new polynomial time algorithm for linear programming. *Combinatorica* **4** (1984) 373-395.
- [13] N.K. Karmarkar et K.G. Ramakrishnan, Computational results of an interior point algorithm for large scale linear program. *Math. Programming* **52** (1991) 55-586.
- [14] M. Kojima, S. Mizumo et A. Yoshise, *An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems*, Research Report No. B-217. Department of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan (1988).
- [15] P. Maréchal, On the convexity of the multiplicative potential and penalty function and related topics. *Math. Programming Ser. A* **89** (2001) 505-516.
- [16] M.J. Todd et Y. Yu, A Centered projective algorithm for linear programming. *Math. Oper. Res.* **15** (1990) 508-529.
- [17] T. Tsuchiya, Quadratic convergence of the Iri-Imaï algorithm for degenerate linear programming problems. *J. Optim. Theory Appl.* **87** (1995) 703-726.
- [18] Y. Ye, K.O. Kortanek et S. Huang, Near boundary behavior of primal-dual potential reduction algorithm for linear programming. *Math. Programming* **58** (1993) 243-255.
- [19] S. Zhang, Convergence property of the Iri-Imaï algorithm for some smooth convex programming problems. *J. Optim. Theory Appl.* **82** (1994) 121-137.