

CONTAINING AND SHRINKING ELLIPSOIDS IN THE PATH-FOLLOWING ALGORITHM

Yinyu YE*

Department of Management Sciences, The University of Iowa, Iowa City, IA, USA

Michael J. TODD**

Department of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, USA

Received 20 January 1988

We describe a new potential function and a sequence of ellipsoids in the path-following algorithm for convex quadratic programming. Each ellipsoid in the sequence contains all of the optimal primal and dual slack vectors. Furthermore, the volumes of the ellipsoids shrink at the ratio $2^{-\Omega(\sqrt{n})}$, in comparison to $2^{-\Omega(1)}$ in Karmarkar's algorithm and $2^{-\Omega(1/n)}$ in the ellipsoid method. We also show how to use these ellipsoids to identify the optimal basis in the course of the algorithm for linear programming.

Key words: Linear programming, convex quadratic programming, path-following algorithm, Karmarkar's algorithm, ellipsoid method.

1. Introduction

Since Karmarkar proposed the interior projective algorithm [7], another interior algorithm which avoids the projective transformations—the center path-following algorithm—has been developed by several authors. Sonnevend [15] proposed the concept of the 'analytic' center for polyhedra. Megiddo [11] and Bayer and Lagarias [1] studied the center pathway to the optimal set. Renegar [14] developed the center method with the first global convergence ratio $(1 - \Omega(1/\sqrt{n}))$ for linear programming (LP). Using the path-following idea and rank-one updating techniques, Gonzaga [5] and Vaidya [17] further reduced the solution time for LP to $O(n^3L)$. Recently, Kojima, Mizuno and Yoshise [9], Monteiro and Adler [12] and Ye [18] have analyzed the primal-dual center path-following algorithm for convex quadratic programming or convex linear complementarity problems.

On the other hand, there is an interesting relationship between the ellipsoid method and Karmarkar's algorithm. Todd [16] described a sequence of ellipsoids that contain all of the optimal dual solutions. Ye [20] found that Karmarkar's potential function represents the logarithmic volume of the ellipsoids containing all

* Research supported by The U.S. Army Research Office through The Mathematical Sciences Institute of Cornell University when the author was visiting at Cornell.

** Research supported in part by National Science Foundation Grant ECS-8602534 and Office of Naval Research Contract N00014-87-K-0212.

optimal dual slack vectors; therefore the containing ellipsoids shrink at the ratio $2^{-\Omega(1)}$ at each iteration as the potential function declines. Moreover, a criterion for checking the intersection of dual hyperplanes and the containing ellipsoid was developed to identify the optimal basis for LP [16, 19].

In this paper, we analyze a new sequence of ellipsoids in the primal-dual center path-following algorithm. We show that each ellipsoid in the sequence *contains* all of the optimal primal and dual (slack) solutions and these ellipsoids *shrink* at the global ratio $2^{-\Omega(\sqrt{n})}$, in comparison to $2^{-\Omega(1)}$ in Karmarkar's algorithm and $2^{-\Omega(1/n)}$ in the ellipsoid method [8, 10]. Similarly, we develop a criterion to identify the optimal basis for linear programming.

2. Path-following algorithm for convex quadratic programming

The convex quadratic program is usually stated in the following standard form:

$$\begin{aligned} \text{QP:} \quad & \text{minimize} \quad f(x) = \frac{1}{2}x^T Qx + c^T x \\ & \text{subject to} \quad Ax = b, \quad x \geq 0, \end{aligned}$$

where $Q \in \mathbb{R}^{n \times n}$, c and $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, Q is a positive semi-definite matrix, and superscript T denotes the transpose operation. The dual problem can be written as

$$\begin{aligned} \text{QD:} \quad & \text{maximize} \quad d(x, y) = b^T y - \frac{1}{2}x^T Qx \\ & \text{subject to} \quad Ax = b, \quad x \geq 0, \quad Qx + c - A^T y \geq 0, \end{aligned}$$

where $y \in \mathbb{R}^m$. For all x and y that are feasible for QD,

$$d(x, y) \leq z^* \leq f(x), \tag{1}$$

where z^* designates the optimal objective value of QP.

Based on the Kuhn-Tucker conditions, x^* is an optimal feasible solution for QP if and only if the following three optimality conditions hold [3]:

- (i) primal feasibility: x^* is feasible for QP;
- (ii) dual feasibility: there exists y^* such that x^* and y^* are feasible for QD;
- (iii) complementary slackness:

$$X^* s^* = 0 \quad \text{or} \quad f(x^*) = d(x^*, y^*) \tag{2}$$

where $s^* = Qx^* + c - A^T y^*$ is the corresponding dual slack vector.

In this paper, the upper-case letter X (S) designates the diagonal matrix whose entries are the components of the vector x (s), e is the vector of all ones, and $\|\cdot\|$ (without subscript) denotes the L_2 norm. These notations will be used throughout this paper. Note that

$$e^T Sx = e^T Xs = x^T s = f(x) - d(x, y)$$

whenever

$$s = Qx + c - A^T y \quad \text{and} \quad Ax = b.$$

The algorithm for solving QP and QD can be described as follows (see [9, 12, 18], which also describe how to obtain the initial x^0 and (y^0, s^0)).

Algorithm. Given x^0, y^0 and s^0 with

$$Ax^0 = b, x^0 > 0, \quad s^0 = Qx^0 + c - A^T y^0 > 0 \quad \text{and} \quad \|X^0 s^0 - z^0 e\| \leq \alpha z^0,$$

where $z^0 = (x^0)^T s^0 / n$ and $\alpha = \frac{1}{5}$; and given convergence criterion $\varepsilon > 0$;

set $k = 0$;

while $z^k \geq \varepsilon$ **do**

begin

let $\lambda = (1 - \alpha / \sqrt{n}) z^k$;

let Δx and Δy solve the linear equations

$$\begin{pmatrix} Q + (X^k)^{-1} S^k & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} \lambda (X^k)^{-1} e - s^k \\ 0 \end{pmatrix}; \quad (3)$$

let $x^{k+1} = x^k + \Delta x$, $y^{k+1} = y^k + \Delta y$, $\Delta s = Q\Delta x - A^T \Delta y$, $s^{k+1} = s^k + \Delta s$ and $z^{k+1} = (x^{k+1})^T s^{k+1} / n$;

$k = k + 1$;

end.

Instead of using the objective function as the merit function, let us define the following potential function associated with the algorithm

$$\phi(x^k, s^k) = 2n \ln((x^k)^T s^k) - \sum_{i=1}^n \ln(x_i^k s_i^k),$$

which is similar to Karmarkar's potential function [7]. Then we can establish the following result.

Theorem 1. Let $\alpha = \frac{1}{5}$. Then for all k , the sequences $\{x^k\}$, $\{y^k\}$ and $\{s^k\}$ generated by the algorithm satisfy

$$Ax^k = b, \quad x^k > 0, \quad s^k = Qx^k + c - A^T y^k > 0, \quad (4)$$

$$\|X^k s^k - z^k e\| \leq \alpha z^k \quad (5)$$

and

$$\phi(x^{k+1}, s^{k+1}) \leq \phi(x^k, s^k) - \frac{1}{10} \sqrt{n}. \quad (6)$$

Inequalities (4) ensure (strict) feasibility. We call inequality (5) the approximately centered condition. It ensures that x^k and (y^k, s^k) are close to the center pathways. In addition to (6), we have (see Lemma 2),

$$z^{k+1} \leq (1 - 1/(10\sqrt{n})) z^k.$$

This shows that the gap between the primal and dual objective values, or the mean value of the complementary slackness vector $X^k s^k$, is reduced at a fixed ratio $1 - 1/(10\sqrt{n}) < 1$.

We use induction to prove Theorem 1. We now state the following two lemmas, whose proofs follow arguments in [9, 12] (also see [18]) and are omitted here.

Lemma 1 below essentially says that x^{k+1} and s^{k+1} remain as interior solutions for QP and QD for sufficiently small α .

Lemma 1. *If (4) and (5) hold, then*

$$\|(X^k)^{-1}\Delta x\|^2 + \|(S^k)^{-1}\Delta s\|^2 \leq 2(\alpha/(1-\alpha))^2. \quad \square$$

The following lemma establishes a fixed convergence ratio for z^{k+1} and a bound for $\|X^{k+1}s^{k+1} - z^{k+1}e\|$.

Lemma 2. *If (4) and (5) hold, then*

$$\left(1 - \frac{\alpha}{\sqrt{n}}\right)z^k \leq z^{k+1} \leq \left(1 - \frac{\alpha}{\sqrt{n}} + \frac{\alpha^2}{(1-\alpha)n}\right)z^k$$

and

$$\|X^{k+1}s^{k+1} - z^{k+1}e\| \leq \alpha^2 z^k / (1-\alpha). \quad \square$$

Proof of Theorem 1. Assume (4) and (5) hold for k . Then $\alpha = \frac{1}{5}$ implies $2(\alpha/(1-\alpha))^2 < 1$ and so from Lemma 1, since $A\Delta x = 0$,

$$Ax^{k+1} = b, \quad x^{k+1} > 0, \quad \text{and} \quad s^{k+1} = Qx^{k+1} + c - A^T y^{k+1} > 0.$$

From Lemma 2,

$$\|X^{k+1}s^{k+1} - z^{k+1}e\| \leq \alpha^2 z^k / (1-\alpha) \leq \alpha(1 - \alpha/\sqrt{n})z^k \leq \alpha z^{k+1}$$

or

$$\|X^{k+1}s^{k+1}/z^{k+1} - e\| \leq \alpha. \quad (7)$$

Hence (4) and (5) hold for $k+1$. To prove (6), we note first that for any δ with $|\delta| \leq \beta < 1$,

$$\ln(1 + \delta) \geq \delta - \frac{1}{2}\delta^2/(1-\beta)$$

from [7] and [6]. Thus

$$\sum_{i=1}^n \ln(x_i^{k+1}s_i^{k+1}/z^{k+1}) \geq -\frac{\|X^{k+1}s^{k+1}/z^{k+1} - e\|^2}{2(1 - \|X^{k+1}s^{k+1}/z^{k+1} - e\|_\infty)}. \quad (8)$$

From (7) and (8),

$$\begin{aligned} & n \ln((x^{k+1})^T s^{k+1}) - \sum_{i=1}^n \ln(x_i^{k+1}s_i^{k+1}) \\ &= n \ln((x^{k+1})^T s^{k+1}/z^{k+1}) - \sum_{i=1}^n \ln(x_i^{k+1}s_i^{k+1}/z^{k+1}) \\ &\leq n \ln n + \frac{1}{2}\alpha^2/(1-\alpha) \\ &\leq n \ln((x^k)^T s^k) - \sum_{i=1}^n \ln(x_i^k s_i^k) + \frac{1}{2}\alpha^2/(1-\alpha), \end{aligned} \quad (9)$$

where the last relation holds due to the arithmetic-geometric mean inequality. Again from Lemma 2,

$$\begin{aligned} & n \ln((x^{k+1})^T s^{k+1}) - n \ln((x^k)^T s^k) \\ &= n \ln\left(\frac{z^{k+1}}{z^k}\right) \leq n \ln\left(1 - \frac{\alpha}{\sqrt{n}} + \frac{\alpha^2}{(1-\alpha)n}\right) \leq -\sqrt{n}\alpha + \frac{\alpha^2}{1-\alpha}. \end{aligned} \quad (10)$$

Adding (9) and (10), we have

$$\begin{aligned} \phi(x^{k+1}, s^{k+1}) - \phi(x^k, s^k) &\leq -\sqrt{n}\alpha + \frac{3}{2}\alpha^2/(1-\alpha) \\ &\leq -\frac{1}{5}\sqrt{n} + \frac{1}{10} \leq -\frac{1}{10}\sqrt{n}. \quad \square \end{aligned}$$

3. Containing ellipsoids in the path-following algorithm

Note that if $s^k = Qx^k + c - A^T y^k \geq 0$ and $s^* = Qx^* + c - A^T y^* \geq 0$, then

$$(x^k - x^*)^T (s^k - s^*) = (x^k - x^*)^T Q (x^k - x^*) \geq 0,$$

so that

$$(x^k)^T s^k + (x^*)^T s^* \geq (x^k)^T s^* + (x^*)^T s^k \geq 0.$$

Suppose x^* and s^* are optimal primal and dual (slack) vectors. Then

$$\begin{aligned} & \|X^k s^*\|^2 + \|S^k x^*\|^2 \\ &\leq ((x^k)^T s^* + (s^k)^T x^*)^2 \leq ((x^k)^T s^k + (x^*)^T s^*)^2 = ((x^k)^T s^k)^2. \end{aligned}$$

Now let us define a primal-dual ellipsoid

$$E^k = \{(x, s) \in \mathbb{R}^{2n} : \|X^k s\|^2 + \|S^k x\|^2 \leq ((x^k)^T s^k)^2\}.$$

Then, we must have

$$(x^*, s^*) \in E^k \quad \text{for all } k.$$

Furthermore, the volume of E^k is

$$V(E^k) = \frac{\gamma((x^k)^T s^k)^{2n}}{\det(X^k) \det(S^k)} = \frac{\gamma((x^k)^T s^k)^{2n}}{\prod_{i=1}^n (x_i^k s_i^k)}$$

where γ is the volume of the unit ball in \mathbb{R}^{2n} . In other words,

$$\ln V(E^k) = \phi(x^k, s^k) + \ln \gamma,$$

i.e., the logarithmic volume of the ellipsoid E^k only differs from the potential function by a constant. Therefore, via Theorem 1 and the above relations, we derive the following shrinking theorem for the containing ellipsoids $\{E^k\}$.

Theorem 2. *For all k , the ellipsoids defined above satisfy*

$$(x^*, s^*) \in E^k \quad \text{and} \quad V(E^{k+1})/V(E^k) \leq 2^{-\Omega(\sqrt{n})}. \quad \square$$

4. A criterion to identify the optimal basis

In order to use E^k to identify the optimal basis, one has to calculate the projections onto the null space of A scaled by X^k and by $(S^k)^{-1}$, which may be too costly to compute [16, 19]. However, we can do almost as well using a single projection, onto the null space of AD , where

$$D = (X^k)^{1/2}(S^k)^{-1/2}.$$

Indeed, if $Q = 0$ so that QP reduces to a linear programming problem, then this projection is used to solve (3). Let

$$\bar{A} = AD \quad \text{and} \quad \bar{P} = \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A},$$

so that \bar{P} is the projection onto the row space of \bar{A} and $I - \bar{P}$ the projection onto its null space. Then the solution to (3) when $Q = 0$ is

$$\Delta y = -(\bar{A}\bar{A}^T)^{-1}\bar{A}(\lambda\bar{e}^{-1} - \bar{e}) \quad \text{and} \quad \Delta x = D(I - \bar{P})(\lambda\bar{e}^{-1} - \bar{e}),$$

where

$$\bar{e}^{-1} = (X^k S^k)^{-1/2}e \quad \text{and} \quad \bar{e} = (X^k S^k)^{1/2}e.$$

We now construct an approximation \hat{E}^k to E^k by

$$\hat{E}^k = \{(x, s) \in \mathbb{R}^{2n} : \|Ds\|^2 + \|D^{-1}x\|^2 \leq n^2 z^k / (1 - \alpha)\}.$$

To verify that $E^k \subset \hat{E}^k$, note that

$$\begin{aligned} \|Ds\|^2 + \|D^{-1}x\|^2 &\leq \|(X^k S^k)^{-1/2}\|^2 (\|X^k s\|^2 + \|S^k x\|^2) \\ &\leq ((x^k)^T s^k)^2 / ((1 - \alpha)z^k) = n^2 z^k / (1 - \alpha) \end{aligned}$$

for any $(x, s) \in E^k$. Furthermore, the volume of \hat{E}^k shrinks at approximately the same rate as that of E^k , since it is easy to show that

$$\hat{E}^k \subset \sqrt{((1 + \alpha)/(1 - \alpha))} E^k.$$

Now we use \hat{E}^k to identify if possible the optimal basis. Consider the problem

$$\text{BP}_i: \quad \text{minimize} \quad s_i \quad (\text{or } x^i)$$

$$\text{subject to} \quad Ax = b,$$

$$s = Qx + c - A^T y,$$

$$(x, s) \in \hat{E}^k.$$

If the minimal value of BP_i is positive, then $s_i^* > 0$ ($x_i^* > 0$), showing that x_i is not in (or is in) the optimal basis [4]. Note that a closed-form (analytical) solution exists for BP_i . However, for simplicity, we assume that $Q = 0$, i.e., the QP problem

reduces to a LP problem. Then, similar to [2] we show how to solve BP_i . In fact, we can rewrite BP_i (up to a positive scale factor in the objective function) by using the scaling matrix D as

$$\begin{aligned} BP'_i: \quad & \text{minimize} \quad \bar{s}_i \quad (\text{or } \bar{x}^i) \\ & \text{subject to} \quad \bar{A}\bar{x} = b, \\ & \quad \quad \quad \bar{s} = \bar{c} - \bar{A}^\top y, \\ & \quad \quad \quad \|\bar{x}\|^2 + \|\bar{s}\|^2 \leq n^2 z^k / (1 - \alpha), \end{aligned}$$

where

$$\bar{c} = Dc, \quad \bar{x} = D^{-1}x \quad \text{and} \quad \bar{s} = Ds.$$

Since we are only interested in the sign of the minimal solution of BP_i , BP'_i gives the same information as BP_i .

To solve BP'_i , we first note that x^k and s^k are feasible for the primal and dual. Hence, $\bar{x} = \bar{e}$ and $\bar{s} = \bar{e}$ are feasible in BP'_i . Also,

$$\bar{A}\bar{P}\bar{e} = \bar{A}\bar{e} = b$$

and

$$(I - \bar{P})\bar{e} = \bar{c} - \bar{A}^\top y^k - \bar{A}^\top (\bar{A}\bar{A}^\top)^{-1} \bar{A}\bar{e}.$$

Thus any \bar{x} and \bar{s} satisfying the equality constraints in BP'_i can be expressed as

$$\bar{x} = \bar{P}\bar{e} + (I - \bar{P})\bar{u} \quad \text{and} \quad \bar{s} = (I - \bar{P})\bar{e} + \bar{P}\bar{v}$$

for some \bar{u} and \bar{v} in \mathbb{R}^n . Substituting \bar{x} and \bar{s} in BP'_i with the above two expressions and noting that the equality constraints are now redundant, we have the following problem that is equivalent to BP'_i :

$$\begin{aligned} & \text{minimize} \quad ((I - \bar{P})\bar{e})_i + (\bar{P}\bar{v})_i \quad (\text{or } (\bar{P}\bar{e})_i + ((I - \bar{P})\bar{u})_i) \\ & \text{subject to} \quad \|\bar{u}\|^2 + \|\bar{v}\|^2 \leq \frac{n^2 z^k}{1 - \alpha} - \|\bar{P}\bar{e}\|^2 - \|(I - \bar{P})\bar{e}\|^2. \end{aligned}$$

The right hand side of this inequality is denoted by

$$\varepsilon^k = \frac{n^2 z^k}{1 - \alpha} - \|\bar{P}\bar{e}\|^2 - \|(I - \bar{P})\bar{e}\|^2 = \frac{n^2 z^k}{1 - \alpha} - \|\bar{e}\|^2 = \left(\frac{n^2}{1 - \alpha} - n \right) z^k.$$

We can now show:

Theorem 3. *If*

$$((I - \bar{P})\bar{e})_i - \sqrt{\varepsilon^k \bar{p}_{ii}} > 0,$$

then x_i is not in the optimal basis, while if

$$(\bar{P}\bar{e})_i - \sqrt{\varepsilon^k(1 - \bar{p}_{ii})} > 0,$$

x_i is in the optimal basis, Here \bar{p}_{ii} is the i th diagonal entry of \bar{P} .

Proof. The minimal solution of BP'_i (minimizing \bar{s}_i) is to set $\bar{u} = 0$ and $\bar{v} = -\sqrt{\varepsilon^k} \bar{P}e^i / \|\bar{P}e^i\|$, where e^i is the i th unit vector. Then the minimal value is

$$((I - \bar{P})\bar{e})_i - \sqrt{\varepsilon^k} \|\bar{P}e^i\| = ((I - \bar{P})\bar{e})_i - \sqrt{\varepsilon^k} \bar{p}_{ii},$$

since

$$\|\bar{P}e^i\|^2 = (e^i)^T \bar{P}^T \bar{P} e^i = (e^i)^T \bar{P} e^i = \bar{p}_{ii} \geq 0.$$

This proves the first part of the theorem, and the second part follows similarly. \square

References

- [1] D. Bayer and J.C. Lagarias, "The non-linear geometry of linear programming, I. Affine and projective scaling trajectories, II. Legendre transform coordinates and central trajectories," *Transactions of the American Mathematical Society* 314 (1989) 499–581.
- [2] B.P. Burrell and M.J. Todd, "The ellipsoid method generates dual variables," *Mathematics of Operations Research* 10 (1985) 688–700.
- [3] R.W. Cottle and G.B. Dantzig, "Complementary pivot theory of mathematical programming", *Linear Algebra and its Applications* 1 (1968) 103–125.
- [4] G.B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, NJ, 1963).
- [5] C.C. Gonzaga, "An algorithm for solving linear programming problems in $O(n^3L)$ operations," in: N. Megiddo, ed., *Progress in Mathematical Programming* (Springer, New York, 1988) pp. 1–28.
- [6] C.C. Gonzaga, "Conical projection algorithms for linear programming," *Mathematical Programming* 43 (1989) 151–173.
- [7] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica* 4 (1984) 373–395.
- [8] L.G. Khachiyan, "A polynomial algorithm for linear programming," *Doklady Akademii Nauk SSSR* 244 (1979) 1093–96. [Translated in: *Soviet Mathematics Doklady* 20 (1979) 191–94.]
- [9] M. Kojima, S. Mizuno and A. Yoshise, "A polynomial-time algorithm for a class of linear complementarity problems," *Mathematical Programming* 44 (1989) 1–26.
- [10] M.K. Kozlov, S.P. Tarasov and L.G. Khachiyan, "Polynomial solvability of convex quadratic programming," *Doklady Akademii Nauk SSSR* 5 (1979) 1051–1053.
- [11] N. Megiddo, "Pathways to the optimal set in linear programming," *Proceedings of the 6th Mathematical Programming Symposium of Japan* (Nagoya, Japan, 1986) 1–35.
- [12] R.C. Monteiro and I. Adler, "Interior path following primal–dual algorithms. Part II: Convex quadratic programming," *Mathematical Programming* 44 (1989) 43–66.
- [13] C.H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Prentice-Hall, Englewood Cliffs, NJ, 1982).
- [14] J. Renegar, "A polynomial-time algorithm, based on Newton's method, for linear programming," *Mathematical Programming* 40 (1988) 59–93.
- [15] G. Sonnevend, "An 'analytic center' for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming," *Proceedings of the 12th IFIP Conference on System Modeling and Optimization* (Budapest, 1985).
- [16] M.J. Todd, "Improved bounds and containing ellipsoids in Karmarkar's linear programming algorithm," *Mathematics of Operations Research* 13 (1988) 650–659.
- [17] P.M. Vaidya, "An algorithm for linear programming which requires $O((m+n)n^2 + (m+n)^{1.5}n)L$ arithmetic operations," to appear in: *Mathematical Programming* 47 (1990) 175–201, next issue.

- [18] Y. Ye, "Interior algorithms for linear, quadratic, and linearly constrained convex programming," Ph.D. Thesis, Department of Engineering-Economic Systems, Stanford University (Stanford, CA, 1987).
- [19] Y. Ye, "Recovering optimal basis in Karmarkar's polynomial algorithm for linear programming," to appear in: *Mathematics of Operations Research* (1990).
- [20] Y. Ye, "Karmarkar's algorithm and the ellipsoid method," *Operations Research Letters* 4 (1987) 177-182.