Mathematical Programming 44 (1989) 157-179
North-Holland

AN EXTENSION OF KARMARKAR'S PROJECTIVE ALGORITHM FOR CONVEX QUADRATIC PROGRAMMING

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

Edison TSE
Department of Engineering-Economic Systems, Stanford University, Stanford, CA, USA

Received 1 July 1986
Revised manuscript received March 1988

We present an extension of Karmarkar's linear programming algorithm for solving a more general group of optimization problems: convex quadratic programs. This extension is based on the iterated application of the objective augmentation and the projective transformation, followed by optimization over an inscribing ellipsoid centered at the current solution. It creates a sequence of interior feasible points that converge to the optimal feasible solution in $O(Ln)$ iterations; each iteration can be computed in $O(Ln^3)$ arithmetic operations, where n is the number of variables and L is the number of bits in the input. In this paper, we emphasize its convergence property, practical efficiency, and relation to the ellipsoid method.

Key words: Convex quadratic programming, Karmarkar's LP algorithm, ellipsoid method, primal and dual.

Introduction

Optimization problems can generally be divided into two categories: combinatorial optimization problems and continuous optimization problems. Combinatorial optimization problems contain discrete variables while continuous optimization problems have continuous variables. These two categories generally are quite different in their characteristics, and the algorithms for solving them are very divergent. Therefore, quadratic programming (QP) plays a unique role in optimization theory: in one sense, it is a continuous optimization that includes linear programming (LP) and is a fundamental subroutine for general nonlinear programming, but it may also be considered one of the most challenging combinatorial optimization problems.

The combinatorial nature of QP is basically embedded in the existence of inequality constraints, which in general make inequality-constrained optimization (ICO) harder to solve than equality-constrained optimization (ECO). Naturally, most algorithms, such as simplex-type methods proposed by Beale [3], Cottle and Dantzig [5], Lemke [17] and Wolfe [24], gradient-projection (GP) method of Rosen [20] and Hildreth [11], and active-set method of Gill and Murray [9], solve a sequence of ECO's in order to approach the optimal solution for ICO. Geometrically, they move along the boundary of the feasible region to approach the optimal (boundary) feasible solution. Unfortunately, when the iterative solution arrives at vertices, a combinatorial decision has to be made to reselect the basic variables (or active-constraint set). In the worst case, the pivot and GP methods converge in an exponential number of iterations, and they are not polynomial-time algorithms. Thus, the question arises: does there exist a polynomial-time algorithm for QP?

Attempting to answer this question, Khachiyan in 1979 published a proof showing that a certain LP algorithm [14], called the ellipsoid method, is polynomial. Soon after, Kozlov, Tarasov and Khachiyan [16] extended this polynomial approach to solving convex quadratic programs in $O(L^2n^4)$ arithmetic operations. Unfortunately, the ellipsoid method behaves similar to its worst case complexity bound. The solution speed of the ellipsoid method does not compete with that of the simplex method in solving most real problems, and the method's significance remains theoretical.

In another attempt, Karmarkar in 1984 introduced a new polynomial-time algorithm for LP that sparked enormous interest in the mathematical programming community [13]. In contrast to the boundary-seeking nature of the simplex and GP methods, Karmarkar's projective method generates a sequence of points in the interior of the feasible region of a canonical LP form while converging to the optimal solution. Practically, his algorithm is competitive with the simplex method in terms of solution time for linear programming.

In this paper, we present an extension of Karmarkar's LP algorithm for convex quadratic programming. In Section 1, we review the optimality conditions and the combinatorial properties of convex quadratic programming. In Section 2, we introduce the interior ellipsoid (IE) method and discuss its solution strategy and convergence ratio. In Section 3, we modify the IE method using an objective augmentation technique and Karmarkar's projective transformation. We show that convexity of the objective function is invariant in the projective transformation and the objective augmentation. Consequently, the modified IE method is terminated in $O(Ln)$ iterations. In Section 4, we show that each iteration can be computed as systems of linear equations in $O(Ln^3)$ arithmetic operations using the trust region method (see, for example, Sorensen [21]). We also discuss this algorithm's relation to the ellipsoid method in Section 5: it turns out that they are closely related.

We have recently learned that Kapoor and Vaidya [12] also independently devised a similar extension of Karmarkar's algorithm for convex quadratic programming. Both our and Kapoor-Vaidya's methods generate a sequence of interior solution points. However, our approach is based on the convexity invariance lemma proved in this paper, which establishes a theoretical foundation for solving more general convex programming. In addition, our extension replaces a factor $O(\log(n))$ with a factor $O(L)$ in the complexity bound, and the factor $O(L)$ can usually be saved in practice. In this paper, we emphasize the algorithm's convergence property and practical efficiency, rather than its complexity bound and precision requirement. The latter issues were well-analyzed in Karmarkar's paper [13], Kozlov et al.'s paper [16] and Kapoor and Vaidya's paper [12].

1. Convex quadratic programming

In this paper, we solve the following convex quadratic program:

QP1 minimize $f(x) = x^{\mathrm{T}}Qx/2 + cx$

subject to $x \in X = \{x \in R^n\colon Ax = b,\ x \geq 0\}$

where $Q \in R^{n\times n}$, row vector $c \in R^n$, $A \in R^{m\times n}$, $b \in R^m$, Q is a positive semidefinite matrix, and superscript $^{\mathrm{T}}$ denotes the transpose operation. The dual problem of QP1 is

QD1 maximize $d(u, y) = yb - u^{\mathrm{T}}Qu/2$

subject to $(u, y) \in Y = \{(u, y)\colon yA \leq u^{\mathrm{T}}Q + c\}$

where $u \in R^n$ and row-vector $y \in R^m$. For all $(u, y) \in Y$ and $x \in X$

$$d(u, y) \leq z^* \leq f(x), \tag{1.1}$$

where z^* designates the optimal objective value of QP1.

Based on the Kuhn–Tucker conditions, x^* is an optimal feasible solution if and only if the following three optimality conditions hold:

(1) Primal feasibility: $x^* \in X$.

(2) Dual feasibility: $\exists y^*$, such that x^* and y^* are feasible for QD1: $(x^*, y^*) \in Y$.

(3) Complementary slackness: $(\nabla f(x^*) - y^*A)\ \mathrm{diag}(x^*) = 0$. (1.2)

As a result of the above conditions, if an optimal feasible solution exists for QP1, then there exists a basic optimal feasible solution such that

$$\begin{pmatrix} Q & -A^{\mathrm{T}} \\ A & 0 \end{pmatrix}\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} -c^{\mathrm{T}} \\ b \end{pmatrix} \tag{1.3}$$

where $x_i^* = 0$ if $i \in I_B$, an index subset of $\{1, 2, \ldots, n\}$ (Cottle and Dantzig [5] and Eaves [8]). Generally, the nonzero components of a basic feasible solution correspond to solutions of the linear system equations with d as the right-hand vector and B as the left-hand matrix, where d is a subvector of

$$\begin{pmatrix} -c^{\mathrm{T}} \\ b \end{pmatrix},$$

and B is a principal submatrix of

$$\begin{pmatrix} Q & -A^{\mathrm{T}} \\ A & 0 \end{pmatrix}.$$

Therefore, the combinatorial properties of QP are similar to that of LP. Let the coefficients in Q, A, c, and b all be integers. Then the basic feasible solutions of (1.3) are vectors of rational numbers, both the numerator and denominator of which are bounded by 2^L. In other words, for any basic feasible solution, x,

$$x_i \leq 2^L \quad \text{for } i = 1, 2, \ldots, n,$$

and either $x_i = 0$, or $x_i \geq 2^{-L}$, where L is the number of bits in the input. For quadratic programming,

$$L = n^2 + mn + [\log|P|],$$

and P is the product of the nonzero integer coefficients appearing in Q, c, A, and b. This fact adds a certain "discreteness" to the problem: many points in the feasible polytope that we wish to look at are not arbitrary points in R^n, but have entries that are rational numbers with bounded numerators and denominators.

We now assume that there exists an interior feasible solution x^0 for QP1 with

A1 $\quad x^0 \geq 2^{-L} e.$

We further make an implicit assumption that the optimal solution can be found in a bounded polytope, i.e.,

A2 $\quad x \leq 2^L e,$

where e is the vector of all one's. Therefore, for all $x \in X$:

$$-2^{2L} \leq z^* \leq f(x) \leq 2^{3L}. \tag{1.4}$$

In general, we say that QP1 is solved if and only if an $x \in X$ has been found such that

$$f(x) - z^* \leq M^{-1},$$

where $M = 2^{4L}$. Due to the above "discreteness" fact and two assumptions, if the input data are all integers, then the exact optimal feasible solution can be obtained by rounding the error from x, as is done in linear programming and convex quadratic programming of Kapoor and Vaidya [12], Karmarkar [13], Khachiyan [14] and Kozlov et al. [16].

2. Interior ellipsoid (IE) method

We now briefly review some existing algorithms for solving LP and QP. In the pivot-type methods, the solution moves from basic solution to basic solution, i.e., from vertex to vertex on the boundary of the feasible region, converging to the optimal solution point. In the GP or the active-set method, the solution may start from the interior of the feasible region. Soon after, the method would generate the boundary point, and then would move along the boundary of the feasible region. As the iterative solution reached the boundary, a "stalling" phenomenon would occur, and a combinatorial decision would have to be made to reform the base or the active constraint set at each step. In the worst case, the optimal solution would be reached in an exponential number of steps.

The issue is: how can we avoid hitting the "wrong" boundary? In other words, can we develop a mechanism to move the solution in the interior of the feasible region while reducing the objective function?

A geometric expression derived from the LP affine scaling algorithm (Dikin [7], Barnes [2], Kortanek and Shi [15], and Vanderbei, Meketon and Freedman [23]) can be drawn as an interior ellipsoid centered at the starting interior point in the feasible region. Then, the objective function can be minimized over this interior ellipsoid to generate the next interior solution point. A series of such ellipsoids can thus be constructed to generate a sequence of interior points converging to the optimal solution point that sits on the boundary. If the optimal solution point itself is an interior solution (which can happen if the objective is a nonlinear function), then the series terminates as soon as the optimal point is encircled by the newest ellipsoid.

The above geometric expression can be represented by the following optimization problem:

$$\text{QP2.1} \quad \text{minimize} \quad f(x)$$
$$\text{subject to} \quad Ax = b,$$
$$\|D^{-1}(x - x^k)\| \leq \beta < 1,$$

where D is an invertible diagonal matrix, and x^k is the interior feasible solution at the kth iteration. The last constraint, $\|D^{-1}(x-x^k)\| \leq \beta$, corresponds to an ellipsoid embedded in the positive orthant $\{x: x \geq 0\}$. Therefore, $\{x: Ax = b, \|D^{-1}(x-x^k)\| \leq \beta\}$ is an algebraic representation of the interior ellipsoid centered at x^k in X of QP1. The parameter β characterizes the size of the ellipsoid, and D affects the orientation and the shape of the ellipsoid. In this paper, like in the affine scaling method, we choose

$$D = \text{diag}(x^k). \tag{2.1}$$

With this D, the ellipsoid constraint can be rewritten as

$$\|D^{-1}x - e\| \leq \beta < 1. \tag{2.2}$$

Inequality (2.2) implies $x > 0$, i.e., any feasible solution for QP2.1 is an interior feasible solution for QP1. Overall, the algorithm can be described as

Algorithm 2.1. *The IE method*
At the kth iteration **do**
 begin
  $D = \text{diag}(x^k)$;
  let x^{k+1} be the minimal solution for QP2.1;
  $k = k + 1$;
 end.

To solve QP2.1, sometimes, it is more convenient to write QP2.1 into

$$\text{QP2.2} \quad \text{minimize} \quad \hat{f}(\hat{x}) = \hat{x}^{\mathrm{T}}\hat{Q}\hat{x}/2 + \hat{c}\hat{x}$$
$$\text{subject to} \quad \hat{A}\hat{x} = \hat{b},$$
$$\|\hat{x} - e\| \leq \beta,$$

where $\hat{x} = D^{-1}x$, $\hat{Q} = DQD$, $\hat{c} = cD$, $\hat{A} = AD$, and $\hat{b} = b$. Let $\hat{a}$ be the optimal solution for QP2.2. Then, $x^{k+1} = D\hat{a}$ and $\hat{a}$ meets the following optimality conditions:

$$\nabla\hat{f}(\hat{a}) - \hat{y}\hat{A} + \mu(\hat{a} - e)^{\mathrm{T}} = 0; \tag{2.3a}$$

$$\hat{A}\hat{a} = \hat{b}; \tag{2.3b}$$

$$\|\hat{a} - e\| \leqslant \beta \quad \text{and} \quad \mu \geqslant 0; \tag{2.3c}$$

and

$$\mu(\beta - \|\hat{a} - e\|) = 0. \tag{2.3d}$$

Conditions (2.3a) and (2.3b) can be written into a matrix form

$$\begin{pmatrix} \hat{Q} + \mu I & -\hat{A}^{\mathrm{T}} \\ \hat{A} & 0 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} -\hat{c}^{\mathrm{T}} + \mu e \\ \hat{b} \end{pmatrix} \tag{2.3e}$$

which can be solved by approximating the multiplier μ until $\hat{a}(\mu) > 0$ similar to the trust region method [21]. Analytically, we have

$$\hat{y}(\hat{A}\hat{A}^{\mathrm{T}}) = \nabla\hat{f}(\hat{a})\hat{A}^{\mathrm{T}}. \tag{2.4a}$$

Let

$$p^k = (\nabla\hat{f}(\hat{a}) - \hat{y}\hat{A})^{\mathrm{T}}, \tag{2.4b}$$

then

$$\mu = \frac{\|p^k\|}{\beta}, \tag{2.4c}$$

$$\nabla\hat{f}(\hat{a})p^k = \nabla f(x^{k+1})Dp^k = \|p^k\|^2, \tag{2.4d}$$

and, if $p^k \neq 0$, from (2.3a), (2.4b), (2.4c)

$$\hat{a} = e - \beta\frac{p^k}{\|p^k\|}. \tag{2.4e}$$

In terms of the original variables and coefficients,

$$x^{k+1} = D\hat{a} = x^k - \beta\frac{Dp^k}{\|p^k\|}, \tag{2.5a}$$

$$p^k = D(\nabla f(x^{k+1}) - \hat{y}A)^{\mathrm{T}}. \tag{2.5b}$$

In order to analyze convergence of the IE method, the following three lemmas are proved. Lemma 2.1 basically states that if the algorithm "stalls", then it arrives at a positive (interior) optimal feasible solution in finite iterations.

Lemma 2.1. *If* $p^k = 0$ (*i.e.* $\mu = 0$) *for* $k < \infty$, *then* x^{k+1} *and* $\hat{y}$ *are optimal for* QP1 *and* QD1.

Proof. Note from (2.5b) that

$$p^k = 0$$

implies

$$(\nabla f(x^{k+1}) - \hat{y}A)D = 0$$

which implies

$$(\nabla f(x^{k+1}) - \hat{y}A) = 0, \tag{2.6}$$

and

$$(\nabla f(x^{k+1}) - \hat{y}A)\ \mathrm{diag}(x^{k+1}) = 0. \tag{2.7}$$

Therefore, the conclusion in Lemma 2.1 follows from the optimality conditions—(1) x^{k+1} is feasible for QP1, (2) $(x^{k+1}, \hat{y})$ is feasible for QD1 from (2.6), and (3) complementary slackness is satisfied from (2.7). □

If $\|p^k\| > 0$ for all finite k, the second lemma claims that $\|p^k\| \to 0$, where $\to$ designates "converges to".

Lemma 2.2. *Let* $\|p^k\| > 0$ *for all* $k < \infty$ *and the optimal objective value of* QP1 *be bounded from below. Then* $\|p^k\| \to 0$ (*or* $\mu \to 0$).

Proof. Using (2.5a) and (2.4d), we have

$$\begin{aligned} f(x^k) &= f\left(x^{k+1} + \beta \frac{Dp^k}{\|p^k\|}\right) \\ &= f(x^{k+1}) + \frac{\beta}{\|p^k\|} \nabla f(x^{k+1}) Dp^k + \frac{\beta^2}{2\|p^k\|^2} (p^k)^{\mathrm{T}} DQDp^k \\ &= f(x^{k+1}) + \beta \|p^k\| + \frac{\beta^2}{2\|p^k\|^2} (p^k)^{\mathrm{T}} DQDp^k, \end{aligned} \tag{2.8}$$

where the Hessian Q is at least positive semi-definite. Therefore,

$$\beta \|p^k\| \leq f(x^k) - f(x^{k+1}). \tag{2.9}$$

Since $f(x^k)$ is monotonically decreasing and is bounded from below, $f(x^k)$ must converge and $f(x^k) - f(x^{k+1}) \to 0$, which implies $\|p^k\| \to 0$. □

Let y^{k+1} be $\hat{y}$ in (2.3a) at the kth iteration. Note that y^{k+1} always exists as a solution of (2.4a), even though it may not be unique. Therefore, even if x^k converges, y^k does not necessarily converge.

Lemma 2.3. *If* $\|p^k\|>0$ *for all* $k<\infty$, $x^k \to x^\infty$, $y^k \to y^\infty$, *and* $\|p\|^k \to 0$, *then* x^∞ *and* y^∞ *are feasible for* QD1.

Proof. We have

$$p_i^\infty = x_i^\infty(\nabla f(x^\infty) - y^\infty A)_i = 0 \quad \text{for } i=1,2,\ldots,n. \tag{2.10}$$

Suppose y^∞ is not feasible for QD1, i.e., $\exists \varepsilon>0$ and $1 \leqslant j \leqslant n$, such that

$$(\nabla f(x^\infty) - y^\infty A)_j \leqslant -\varepsilon < 0,$$

then $\exists K>0$ such that, for all $\infty > k > K$,

$$(\nabla f(x^{k+1}) - y^{k+1} A)_j < -\frac{\varepsilon}{2}.$$

At the kth ($k>K$) iteration of the algorithm,

$$x_j^{k+1} = x_j^k \left(1 - \beta \frac{x_j^k(\nabla f(x^{k+1}) - y^{k+1}A)_j}{\|p^k\|}\right) > x_j^k$$

hence,

$$x_j^{k+1} > x_j^k > x_j^K > 0 \quad \text{for all } k>K.$$

Thus, $\{x_j^k\}$ is a strictly increasing positive series for $k>K$. Since neither x_j^k nor $(\nabla f(x^k) - y^k A)_j$ converges to 0, $x_j^k(\nabla f(x^k) - y^k A)_j$ does not converge to 0. This contradicts (2.10). Therefore, it must be true that

$$\nabla f(x^\infty) - y^\infty A \geqslant 0,$$

i.e., x^∞ and y^∞ are feasible for QD1. □

Now, we can derive the following convergence theorem.

Theorem 2.1. *Let the optimal objective value of* QP1 *be bounded from below and let* x^k *and* y^k *converge. Then, Algorithm* 2.1 *generates solution sequences* x^k *and* y^k *that converge to the optimal solutions for both* QP1 *and* QD1.

Proof. If $p^k=0$ for $k<\infty$, the conclusion of Theorem 2.1 follows from Lemma 2.1; otherwise, from Lemma 2.3, x^∞ and y^∞ are feasible for QP1 and QD1, and from Lemma 2.2, $\|p^\infty\|=0$, which implies that complementary slackness is satisfied. Thus, x^∞ and y^∞ are optimal for QP1 and QD1. □

To evaluate the asymptotic convergence rate of the IE method, we have

Theorem 2.2. *Let* x^{k+1} *and* y^{k+1} *be feasible for* QD1. *Then*

$$f(x^{k+1}) - z^* \leqslant \left(1 - \frac{\beta}{\sqrt{n}}\right)(f(x^k) - z^*).$$

Proof. Since $f(x)$ is a convex function,

$$f(x^k)-f(x^{k+1})\geqslant \nabla f(x^{k+1})(x^k-x^{k+1}). \tag{2.11}$$

There is also given that y^{k+1} is feasible for QD1, so, from (1.1),

$$d(x^{k+1},y^{k+1})-z^*=f(x^{k+1})-\nabla f(x^{k+1})x^{k+1}+y^{k+1}b-z^*\leqslant 0. \tag{2.12}$$

From (2.11) and (2.12),

$$\nabla(f(x^{k+1})x^k-y^{k+1}b+d(x^{k+1},y^{k+1})-z^*\leqslant f(x^k)-z^*. \tag{2.13}$$

According to (2.4d) and (2.5a),

$$\nabla f(x^{k+1})x^{k+1}=\nabla f(x^{k+1})x^k-\beta\|p^k\|. \tag{2.14}$$

Using Hölder's inequality and (2.5b), and noting that $p^k\geqslant 0$,

$$\|p^k\|\geqslant\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}|p_i^k|\right)=\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}(p_i^k)\right)=\frac{1}{\sqrt{n}}(\nabla f(x^{k+1})x^k-y^{k+1}b). \tag{2.15}$$

Due to (2.12), (2.13), (2.14), and (2.15),

$$\begin{aligned}f(x^{k+1})-z^*&=\nabla f(x^{k+1})x^{k+1}-y^{k+1}b+d(x^{k+1},y^{k+1})-z^*\\&\leqslant\nabla f(x^{k+1})x^k-\beta\|p^k\|-y^{k+1}b+\left(1-\frac{\beta}{\sqrt{n}}\right)(d(x^{k+1},y^{k+1})-z^*)\\&\leqslant\left(1-\frac{\beta}{\sqrt{n}}\right)(\nabla f(x^{k+1})x^k-y^{k+1}b+d(x^{k+1},y^{k+1})-z^*)\\&\leqslant\left(1-\frac{\beta}{\sqrt{n}}\right)(f(x^k)-z^*).\quad\square\end{aligned}$$

Theoretically, the IE method is not a polynomial-time algorithm since there is no guarantee that x^k and y^k will be feasible for QD1 until they converge. Obviously, if $f(x^k)-z^*$ is reduced at the above ratio for all k, then the iterative solutions converge in a number of iterations that grow polynomially with the size of the problem. In the following sections, using Karmarkar's projective transformation, we propose a modified IE method for QP1 — a potential function is reduced at a fixed ratio, and each iteration can be computed in a polynomial of L and n.

3. Modified interior ellipsoid method

The original version of Karmarkar's algorithm solves a linear program of the special form

$$\begin{aligned}&\text{minimize}\quad \hat{f}(\hat{x})=\hat{c}\hat{x}\\&\text{subject to}\quad \hat{A}\hat{x}=0,\ e^{\mathrm{T}}\hat{x}=n+1,\ \hat{x}\geqslant 0,\end{aligned}$$

which we call Karmarkar's canonical LP form [13]. Here, $\hat{c} \in R^{n+1}$, $\hat{A} \in R^{m\times(n+1)}$. Let r be the radius of the sphere centered at e that inscribes the feasible region, and R be the radius of the sphere centered at e that circumscribes the feasible region. It is easy to show that $r \geq 1$ and $R < 1/(n+1)$, hence, the ratio of the two radii, i.e.,

$$\frac{r}{R} \geq \frac{1}{n+1}.$$

Thus, when minimizing the objective function, $\hat{c}\hat{x}$, with $\hat{x}$ being restricted in the inscribing sphere, the linear objective function will be reduced at the ratio of $(1-1/(n+1))$. However, this reduction should also be true even when $\hat{f}(\hat{x})$ is a general convex function. This observation led us to develop an objective augmentation technique that, coupled with Karmarkar's projective transformation, transforms a convex program to a new convex program in the canonical form. Therefore, we developed a similar algorithm for solving QP1.

To describe this algorithm, we will assume for the moment that the optimal objective value of QP1 is known to be zero, i.e.,

A3 $\quad z^* = 0$ (such an assumption will be removed later).

Then, we introduce a potential function associated with QP1 (like the one in Anstreicher [1], Karmarkar [13], Todd and Burrell [22] and Ye and Kojima [25])

$$P(x) = (n+1)\ln(f(x)) - \sum_{i=1}^{n} \ln(x_i)$$

where $0 < x \in X$ and $f(x) > 0$. By a simple calculation, we have the equality

$$\frac{f(x)}{f(x^0)} = \gamma(x)\exp\left(\frac{P(x)-P(x^0)}{n+1}\right)$$

where

$$\gamma(x) = \exp\left(\frac{\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\ln(x_i^0)}{n+1}\right).$$

By the assumptions A1 and A2, $\gamma(x) \leq 2^{2L}$ in the feasible region X. If $P(x^k)$ tends to $-\infty$ along some sequence $\{0 < x^k \in X\}$, then $f(x^k)$ converges to zero. The algorithm that is described in this section generates a sequence $\{0 < x^k \in X\}$, such that

$$P(x^k) \leq P(x^{k-1}) - \alpha \quad \text{for } k = 1, 2, \ldots,$$

where $\alpha \geq 0.2$. Hence,

$$P(x^k) \leq P(x^0) - k\alpha \quad \text{for } k = 1, 2, \ldots.$$

Thus,

$$f(x^k) \leq 2^{2L} f(x^0)\exp\left(\frac{-k\alpha}{n+1)}\right) \leq 2^{5L}\exp\left(\frac{-k\alpha}{n+1}\right).$$

Therefore,

$$f(x^k) \leq 2^{-4L} = M^{-1} \quad \text{for } k \geq 45L(n+1).$$

Now, let us look at the following problem related to QP1:

QP3.1 minimize $\hat{f}(\hat{x}) = \hat{x}_{n+1} f(T^{-1}(\hat{x}))$

subject to $\hat{x} \in \hat{X} = \{\hat{x}\colon AD\hat{x}[n] - \hat{x}_{n+1} b = 0,\ e^{\mathrm{T}}\hat{x} = n+1,$

$\hat{x}[n] \geq 0 \text{ and } \hat{x}_{n+1} > 0\},$

where $\hat{x}[n]$ is the vector of the first n components of $\hat{x} \in R^{n+1}$, D is defined by (2.1), and T^{-1} is the inverse projective transformation $T^{-1}\colon R^{n+1} \to R^n$ defined by

$$x = T^{-1}(\hat{x}) = \frac{D\hat{x}[n]}{\hat{x}_{n+1}}. \tag{3.1}$$

It can be verified that $\hat{x} \in \hat{X}$ implies $T^{-1}(\hat{x}) \in X$. Conversely, for any $x \in X$, an $\hat{x} \in \hat{X}$ can be obtained via $T\colon R^n \to R^{n+1}$ defined by

$$\hat{x}[n] = \frac{(n+1)D^{-1}x}{e^{\mathrm{T}}D^{-1}x + 1}, \tag{3.2a}$$

and

$$\hat{x}_{n+1} = \frac{n+1}{e^{\mathrm{T}}D^{-1}x + 1}. \tag{3.2b}$$

Particularly, let $\hat{x}^k = T(x^k)$, and then

$$\hat{x}^k = e.$$

In the rest contents of this paper, $\hat{x}$ always designates the variable in $\hat{X}$, and corresponds to $x \in X$. For example, $\hat{x}^k \leftrightarrow x^k$, $\hat{x}^* \leftrightarrow x^*$, $\hat{a} \leftrightarrow a$, and so on.

The augmented nonlinear objective function in QP3.1 plays a key role in our proposed algorithm. Note that $\hat{f}(\hat{x})$ is the product of $f(x)$ (≥ 0) multiplied by $\hat{x}_{n+1}$ (>0). Hence, $\hat{f}(\hat{x}) \geq 0$. Generally, we have

$$\frac{\hat{f}(\hat{x})}{\hat{x}_{n+1}} = f(T^{-1}(\hat{x})) = f(x).$$

Especially,

$$\hat{f}(\hat{x}^k) = \hat{f}(e) = f(x^k),$$

and

$$\hat{f}(\hat{x}^*) = \hat{x}^*_{n+1} f(x^*) = 0.$$

If $f(x)$ is a linear function, then $\hat{f}(\hat{x})$ is also a linear function; otherwise, $\hat{f}(\hat{x})$ looks complicated. But, fortunately, $\hat{f}(\hat{x})$ is merely a convex function according to the following convexity invariance lemma.

Lemma 3.1. *Let $f(x)$ be a convex function in X. Then $\hat{f}(\hat{x})$ is a convex function in $\hat{X}$.*

Proof. Let μ, $\nu \geqslant 0$, $\mu+\nu=1$, and $u=T^{-1}(\hat{u})$ and $v=T^{-1}(\hat{v})$, then for any $\hat{u}$, $\hat{v}\in\hat{X}$

$$\begin{aligned}\hat{f}(\mu\hat{u}+\nu\hat{v}) &= (\mu\hat{u}_{n+1}+\nu\hat{v}_{n+1})f\left(\frac{D(\mu\hat{u}[n]+\nu\hat{v}[n])}{\mu\hat{u}_{n+1}+\nu\hat{v}_{n+1}}\right)\\ &= (\mu\hat{u}_{n+1}+\nu\hat{v}_{n+1})f\left(\frac{(\mu\hat{u}_{n+1})u+(\nu\hat{v}_{n+1})v}{\mu\hat{u}_{n+1}+\nu\hat{v}_{n+1}}\right)\\ &\leqslant (\mu\hat{u}_{n+1})f(u)+(\nu\hat{v}_{n+1})f(v)\\ &= \mu\hat{f}(\hat{u})+\nu\hat{f}(\hat{v}). \quad \square\end{aligned}$$

Lemma 3.1 led to a very important conclusion: the convexity of the objective function remains invariant in the objective augmentation and the projective transformation. Using the IE method, we solve the following sub-optimization problem QP3.2 over an interior ellipsoid centered at $\hat{x}^k$, instead of solving QP3.1. Since the starting point $\hat{x}^k=e$, the interior ellipsoid happens to be an interior sphere in $\hat{X}$.

QP3.2 minimize $\hat{f}(\hat{x})=\hat{x}[n]^{\mathrm{T}}\hat{Q}\hat{x}[n]/(2\hat{x}_{n+1})-\hat{c}\hat{x}[n]$

subject to $\hat{A}\hat{x}=\hat{b}$,

$\|\hat{x}-e\|\leqslant\beta<1$,

where

$$\hat{Q}=DQD,$$

$$\hat{c}=cD,$$

$$\hat{A}=\begin{pmatrix}AD, & -b\\ e^{\mathrm{T}} & \end{pmatrix},$$

and

$$\hat{b}=\begin{pmatrix}0\\ n+1\end{pmatrix}.$$

As a result of Lemma 3.1, the following algorithm is introduced:

Algorithm 3.1.
Repeat do
  begin
    $D=\mathrm{diag}(x^k)$;
    let $\hat{a}$ be the minimal solution for QP3.2;
    $x^{k+1}=T^{-1}(\hat{a})$;
    $k=k+1$;
  end;
until $f(x^k)\leqslant 1/M$.

The next two lemmas are used to prove a theorem for Algorithm 3.1. The first lemma confirms that a fixed objective reduction can be made at each step.

Lemma 3.2.

$$\hat{f}(\hat{a}) \leq \left(1 - \frac{\beta}{n+1}\right) \hat{f}(e).$$

Proof. If $\hat{x}^* \in \{\hat{x}\colon \hat{A}\hat{x} = \hat{b}$ and $\|\hat{x} - e\| \leq \beta\}$, then

$$0 \leq \hat{f}(\hat{a}) \leq \hat{f}(\hat{x}^*) = 0,$$

and so Lemma 3.2 holds. Otherwise, since $\hat{X}$ is a convex polytope, the intersection point of the boundary of $\{\hat{x}\colon \hat{A}\hat{x} = \hat{b}$ and $\|\hat{x} - e\| \leq \beta\}$ and the line segment between e and $\hat{x}^*$, should be feasible for QP3.2. Let $\hat{a}'$ be the intersection point; then $\hat{a}'$ satisfies

$$\|\hat{a}' - e\| = \beta, \tag{3.3}$$

and

$$\hat{a}' = \theta\hat{x}^* + (1-\theta)e \quad \text{for some } 0 < \theta < 1. \tag{3.4}$$

Substituting $\hat{a}'$ in (3.4) for $\hat{a}'$ in (3.3),

$$\|\theta\hat{x}^* + (1-\theta)e - e\| = \beta,$$

then

$$\theta\|\hat{x}^* - e\| = \beta.$$

Note that

$$\|\hat{x}^* - e\|^2 = \|\hat{x}^*\|^2 - 2e^{\mathrm{T}}\hat{x}^* + \|e\|^2,$$

$$e^{\mathrm{T}}\hat{x}^* = n+1,$$

and

$$\|\hat{x}^*\|^2 \leq (e^{\mathrm{T}}\hat{x}^*)^2 = (n+1)^2.$$

Hence,

$$\theta \geq \frac{\beta}{\sqrt{n(n+1)}} > \frac{\beta}{n+1}. \tag{3.5}$$

In addition, due to the convexity of $\hat{f}(\cdot)$,

$$\hat{f}(\hat{a}') \leq \theta\hat{f}(\hat{x}^*) + (1-\theta)\hat{f}(e) = (1-\theta)\hat{f}(e).$$

Since $\hat{a}$ is the optimal feasible solution for QP3.2, $\hat{a}'$ is a feasible solution for QP3.2,

$$\hat{f}(\hat{a}) \leq \hat{f}(\hat{a}') \leq (1-\theta)\hat{f}(e). \tag{3.6}$$

Lemma 3.2 thus follows from (3.5) and (3.6). □

The second lemma is proved by Karmarkar [13], and it essentially measures the second term of the potential function introduced at the beginning of this section.

Lemma 3.3.

$$-\sum_{i=1}^{n+1} \ln(\hat{a}_i) \leq \frac{\beta^2}{2(1-\beta)^2}. \quad \square$$

Based on Lemma 3.2 and Lemma 3.3, we can derive:

Theorem 3.1. *In* $O(Ln)$ *iterations of Algorithm* 3.1, $f(x^k) < M^{-1}$.

Proof. It can be verified from (3.1), (3.2), Lemma 3.2, and Lemma 3.3 that

$$\begin{aligned} P(x^{k+1}) - P(x^k) &= (n+1)\ln\left(\frac{f(x^{k+1})}{f(x^k)}\right) - \sum_{i=1}^{n} \ln\left(\frac{x_i^{k+1}}{x_i^k}\right) \\ &= (n+1)\ln\left(\frac{\hat{f}(\hat{a})}{\hat{f}(e)}\right) - \sum_{i=1}^{n+1} \ln(\hat{a}_i) \\ &\leq (n+1)\ln\left(1-\frac{\beta}{n+1}\right) + \frac{\beta^2}{2(1-\beta)^2} \\ &\leq -\beta + \frac{\beta^2}{2(1-\beta)^2}. \end{aligned}$$

Let $\beta = 0.27$–0.36, and then

$$P(x^{k+1}) < P(x^k) - 0.2.$$

Therefore, Theorem 3.1 follows directly from the discussion at the beginning of this section. $\square$

4. Solving sub-optimization problem

Now we show that QP3.2, or the optimality conditions (2.3a)-(2.3d) (with new meanings of $\hat{f}$, $\hat{A}$, and $\hat{b}$), can be solved in $O(Ln^3)$ arithmetic operations. To solve QP3.2, we can check to see whether or not QP1 has a positive (interior) optimal feasible solution (POFS). If QP1 has no POFS, then neither does QP3.1. In this case, a unique optimal solution must exist for QP3.2. This is because (1) $\hat{f}(\hat{x}) > 0$ for all $\hat{x} \in \{\hat{x}\colon \hat{A}\hat{x} = \hat{b} \text{ and } \|\hat{x} - e\| \leq \beta\}$, i.e., $\hat{x}^* \notin \{\hat{x}\colon \hat{A}\hat{x} = \hat{b} \text{ and } \|\hat{x} - e\| \leq \beta\}$ (2) the sphere constraint $\{\hat{x}\colon \hat{A}\hat{x} = \hat{b} \text{ and } \|\hat{x} - e\| \leq \beta\}$ is a *strictly* convex set, and (3) $\hat{f}(\hat{x})$ is a convex function. By the well-known separating theorem, a unique optimal solution occurs at the boundary of the feasible region. In other words, there exists a unique fixed point $(\hat{a}, \hat{y}, \mu)$ that satisfied (2.3a)-(2.3d), where $\hat{y}$ is the Lagrange multiplier vector for the equality constraints, and μ is the multiplier for the sphere inequality constraint. Furthermore, we can see that the objective function $\hat{f}(\hat{x})$ of QP3.2 is almost quadratic, except that $\hat{x}_{n+1}$ appears in the denominator of the quadratic term. This makes it possible to solve (2.3a)-(2.3d) by using the multiplier

μ as a parameter like equation (2.3e), as described in Section 2. To be more specific, we first provide a bound for μ. Using (3.1) and (3.2),

$$\nabla\hat{f}(\hat{a})=\left(\frac{\hat{a}[n]^{\mathrm{T}}\hat{Q}}{\hat{a}_{n+1}}+\hat{c},\ \frac{-\hat{a}[n]^{\mathrm{T}}\hat{Q}\hat{a}[n]}{2(\hat{a}_{n+1})^2}\right).$$

Thus, we notice a nice property of $\hat{f}(\hat{a})$:

$$\nabla\hat{f}(\hat{a})\hat{a}=\hat{f}(\hat{a}). \tag{4.1}$$

Multiplying both sides of (2.4e) by $\nabla\hat{f}(\hat{a})$ from the left,

$$\nabla\hat{f}(\hat{a})\hat{a}=\nabla\hat{f}(\hat{a})e-\beta\|p^k\|$$

or

$$\beta\|p^k\|=\nabla\hat{f}(\hat{a})(e-\hat{a}). \tag{4.2}$$

This leads to

Lemma 4.1.

$$\frac{\hat{f}(\hat{a})}{\beta(n+1)}\leqslant\mu\leqslant\frac{\hat{f}(e)-\hat{f}(\hat{a})}{\beta^2}.$$

Proof. Due to the convexity of $\hat{f}(\cdot)$ and (4.1)

$$\hat{f}(\hat{x})\geqslant\hat{f}(\hat{a})+\nabla\hat{f}(\hat{a})(\hat{x}-\hat{a})=\nabla\hat{f}(\hat{a})\hat{x}\quad\forall\hat{x}\in\hat{X}.$$

In addition, the optimality conditions imply that $\hat{a}$ is also the optimal solution in minimizing the linear objective function, $\nabla\hat{f}(\hat{a})\hat{x}$, and subjecting $\hat{x}$ to the constraints of QP3.2. Hence, similar to the proof of Lemma 3.2

$$\nabla\hat{f}(\hat{a})\hat{a}\leqslant\theta\nabla\hat{f}(\hat{a})\hat{x}^*+(1-\theta)\nabla\hat{f}(\hat{a})e\leqslant\theta\hat{f}(\hat{x}^*)+(1-\theta)\nabla\hat{f}(a)e$$
$$=(1-\theta)\nabla\hat{f}(a)e\leqslant\left(1-\frac{\beta}{n+1}\right)\nabla\hat{f}(a)e.$$

Therefore, from (4.1), (4.2) and the above inequality

$$\frac{\beta\hat{f}(\hat{a})}{n+1}\leqslant\frac{\beta\nabla\hat{f}(\hat{a})e}{n+1}\leqslant\beta\|p^k\|=\nabla\hat{f}(\hat{a})(e-\hat{a})\leqslant\hat{f}(e)-\hat{f}(\hat{a}).$$

By combining this inequality with (2.4c) in Section 2, we derive the conclusion in Lemma 4.1. □

Now, we split (2.3a) into two groups: the first one is the first n equations of (2.3a)

$$\frac{\hat{Q}\hat{a}[n]}{\hat{a}_{n+1}}+\hat{c}^{\mathrm{T}}-\hat{A}[n]^{\mathrm{T}}\hat{y}^{\mathrm{T}}+\mu(\hat{a}[n]-e)=0 \tag{4.3a}$$

and the second one is the last equation of (2.3a)

$$\frac{-\hat{a}^{\mathrm{T}}[n]\hat{Q}\hat{a}[n]}{2(\hat{a}_{n+1})^2}-\hat{y}\hat{A}_{n+1}+\mu(\hat{a}_{n+1}-1)=0, \tag{4.3b}$$

where $\hat{A}[n]$ is the matrix of the first n columns of $\hat{A}$, and $\hat{A}_{n+1}$ is the last column of $\hat{A}$. Let

$$\hat{y}' = \hat{a}_{n+1}\hat{y}$$

and

$$\lambda = \hat{a}_{n+1}\mu.$$

Then, since $(1-\beta) \leqslant \hat{a}_{n+1} \leqslant (1+\beta)$, Lemma 4.1 imposes a bound for λ:

$$0 \leqslant \frac{(1-\beta)\hat{f}(\hat{a})}{\beta(n+1)} \leqslant \lambda \leqslant \frac{(1+\beta)(\hat{f}(e)-\hat{f}(\hat{a}))}{\beta^2}. \tag{4.4}$$

Particularly, (4.3a) and (4.3b) become

$$(\hat{Q}+\lambda I)\hat{a}[n] - \hat{A}[n]^{\mathrm{T}}\hat{y}'^{\mathrm{T}} = \lambda e - \hat{a}_{n+1}\hat{c}^{\mathrm{T}} \tag{4.5a}$$

and

$$\frac{-\hat{a}^{\mathrm{T}}[n]\hat{Q}\hat{a}[n]}{2} + \lambda(\hat{a}_{n+1})^2 - (\hat{y}'\hat{A}_{n+1}+\lambda)\hat{a}_{n+1} = 0, \tag{4.5b}$$

respectively. In addition, (2.3b) can be rewritten as

$$\hat{A}[n]\hat{a}[n] = \begin{pmatrix} \hat{a}_{n+1}b \\ n+1-\hat{a}_{n+1} \end{pmatrix}. \tag{4.5c}$$

Then, (4.5a) and (4.5c) form a system of linear equations similar to (2.3e):

$$P\begin{pmatrix} \hat{a}[n] \\ \hat{y}'^{\mathrm{T}} \end{pmatrix} = \hat{a}_{n+1}b^1 + b^2$$

where

$$P = \begin{pmatrix} \hat{Q}+\lambda I & -\hat{A}[n]^{\mathrm{T}} \\ \hat{A}[n] & 0 \end{pmatrix},$$

$$b^1 = \begin{pmatrix} -\hat{c}^{\mathrm{T}} \\ b \\ -1 \end{pmatrix} \quad \text{and} \quad b^2 = \begin{pmatrix} \lambda e \\ 0 \\ n+1 \end{pmatrix}.$$

Let

$$\lambda_{\min} = \frac{(1-\beta)^2}{M\beta(n+1)}$$

and

$$\lambda_{\max} = \frac{(1+\beta)f(x^k)}{\beta^2}.$$

Then, for any given $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, we can compute $P^{-1}b^1$ and $P^{-1}b^2$, and let

$$\begin{pmatrix} \hat{a}[n] \\ \hat{y}'^{\mathrm{T}} \end{pmatrix} = \hat{a}_{n+1}P^{-1}b^1 + P^{-1}b^2. \tag{4.6}$$

Then we substitute (4.6) for (4.5b) to compute $\hat{a}_{n+1}$ by solving a single quadratic equation. Thus, $\hat{a}$ will be precisely determined. Finally, we check to see if $\|\hat{a}(\lambda)-e\|$ is close to β. Let $h(\lambda)=\|\hat{a}(\lambda)-e\|-\beta$, then $h(\lambda)$ has a *unique* zero point bounded in (4.4). Obviously, $h(\infty)=\|e-e\|-\beta=-\beta<0$, which implies $h(\lambda_{\max})\leq 0$. On the other hand, if $h(\lambda_{\min})\leq 0$, then from Lemma 4.1 we obtain a positive (interior) optimal feasible solution $x^{k+1}=T^{-1}(\hat{a})$ for QP1 via (4.4):

$$f(x^{k+1})=\hat{f}(\hat{a})/\hat{a}_{n+1}\leq\frac{(1-\beta)}{M\hat{a}_{n+1}}\leq M^{-1}.$$

Generally, we expect $h(\lambda_{\min})>0$. Overall, the above process can be implemented by applying the well-known bisection technique to determine the unique zero λ^* of $h(\lambda)$, as shown in the following procedure:

Procedure 4.1. 1. Set $\beta=0.31$, $\lambda_1=\lambda_{min}$, $\lambda_3=\lambda_{max}$, and $\lambda_2=(\lambda_1+\lambda_3)/2$.

2. Let $\lambda=\lambda_2$, and then compute the two vectors $P^{-1}b^1$ and $P^{-1}b^2$ in $O(n^3)$ arithmetic operations.

3. Substitute (4.6) for (4.5b) to compute $\hat{a}_{n+1}$ by solving a single quadratic equation; hence, $\hat{a}$ will be determined.

4. If $\lambda_3-\lambda_1\leq 2^{-O(L)}$, then stop and return $\hat{a}$; else update the three points λ_1, λ_2 and λ_3, as in the well-known bisection technique, and go to 2.

Remark 4.1. An issue that needs clarifying in Procedure 4.1 is what to do if (4.5b) has non-real solutions. In fact, from the uniqueness of the optimal solution for QP3.2 and the boundness of λ in (4.4), $h(\lambda)$ has a unique zero in $[\lambda_{\min},\lambda_{\max}]$, and for any given $\lambda\in[\lambda^*,\infty)$, there must exist a solution to $\hat{a}_{n+1}$ in (4.5b), such that $\hat{a}_{n+1}$ is real and $|\hat{a}_{n+1}-1|\leq\beta$. Therefore, if (4.5b) has no real solutions, it must be true that $\lambda<\lambda^*$. Essentially, λ characterizes the size (β) of the interior ellipsoid (sphere) from (2.4c). Consequently, similar to the trust region method, searching for λ^* is equivalent to searching for the right size of the interior ellipsoid region.

In the following lemma, we justify that $O(L)$ precision of arithmetic operations is adequate to terminate Procedure 4.1.

Lemma 4.2. *$\hat{a}$ resulted from Procedure* 4.1 *satisfies*

$$|\|\hat{a}-e\|-\beta|\leq 2^{-O(L)}.$$

Proof. Since Procedure 4.1 is performed with λ in the bounded interval $[2^{-O(L)},2^{O(L)}]$, $\hat{Q}+\lambda I$ is positive definite. We also assumed that A has full rank. Therefore, it can be verified that the condition number of $P(\lambda)$ is bounded [12], or

$$\|P(\lambda)^{-1}\|\leq 2^{k_1L}$$

for some constant $k_1>0$. Additionally,

$$\|b^1\|\leq 2^{k_2L} \quad\text{and}\quad \|b^2\|\leq 2^{k_2L}$$

for some constant $k_2 > 0$. Thus, for $\lambda_3 - \lambda_1 \leq 2^{-(2k_1+k_2+2)L}$ we have

$$|\lambda_2 - \lambda^*| \leq 2^{-(2k_1+k_2+2)L},$$

$$\begin{aligned}\|P(\lambda_2)^{-1}b^1 - P(\lambda^*)^{-1}b^1\| &\leq \|P(\lambda_2)^{-1}\|\,\|P(\lambda_2) - P(\lambda^*)\|\,\|P(\lambda^*)^{-1}\|\,\|b^1\| \\ &\leq |\lambda_2 - \lambda^*|\,\|P(\lambda_2)^{-1}\|\,\|P(\lambda^*)^{-1}\|\,\|b^1\| \\ &\leq 2^{-2L}\end{aligned}$$

and similarly

$$\|P(\lambda_2)^{-1}b^2 - P(\lambda^*)^{-1}b^2\| \leq 2^{-2L}.$$

Furthermore, $\hat{a}(\lambda_2)_{n+1}$ was obtained via a square-root operation from $P(\lambda_2)^{-1}b^1$ and $P(\lambda_2)^{-1}b^2$, and then $\hat{a}(\lambda_2)$ is obtained via summation of (4.6), we must have

$$\|\hat{a}(\lambda_2) - \hat{a}(\lambda^*)\| \leq 2^{-O(L)}.$$

Therefore,

$$\begin{aligned}\|\hat{a}(\lambda_2) - e\| &= \|\hat{a}(\lambda_2) - \hat{a}(\lambda^*) + \hat{a}(\lambda^*) - e\| \\ &\geq \|\hat{a}(\lambda^*) - e\| - \|\hat{a}(\lambda_2) - \hat{a}(\lambda^*)\| \\ &\geq \beta - 2^{-O(L)},\end{aligned}$$

and

$$\begin{aligned}\|\hat{a}(\lambda_2) - e\| &= \|\hat{a}(\lambda) - \hat{a}(\lambda^*) + \hat{a}(\lambda^*) - e\| \\ &\leq \|\hat{a}(\lambda^*) - e\| + \|\hat{a}(\lambda_2) - \hat{a}(\lambda^*)\| \\ &\leq \beta + 2^{-O(L)}. \quad \square\end{aligned}$$

Lemma 4.2 established that $x^{k+1} = T^{-1}(\hat{a})$ resulted from Procedure 4.1 in $O(Ln^3)$ arithmetic operations is a valid solution to hold Theorem 3.1, since the range of β can be tolerated from 0.27 to 0.36 to make the potential function reduced by 0.2. Actually, we can use a looser line search tolerance to complete Procedure 4.1, i.e., set $|h(\lambda)| \leq 0.05$ as another criterion to terminate Procedure 4.1. In practice, λ (or μ) can be determined to minimize the potential function as long as the solution $\hat{a} > 0$. Our computational experience indicates that Procedure 4.1 is always terminated in few loops using a reasonable estimation on λ (or μ) from (4.4) (or Lemma 4.1). Moreover, since there is not need to solve QP3.2 exactly, we can use many existing iterative algorithms, such as the Newton-type method or the conjugate direction method to achieve further practical efficiency (see, for example, Luenberger [18]).

Directly from Theorem 3.1 and Lemma 4.2, we obtain

Corollary 4.1. *Let the optimal objective value of* QP1 *be zero. Then* QP1 *can be solved in* $O(L^2n^4)$ *by Algorithm* 3.1, *coupled with Procedure* 4.1. □

Furthermore, QD1 can be emerged into QP1 to form

$$\text{QPD} \quad \text{minimize} \quad F(x, y) = f(x) - d(x, y) = x^{\mathrm{T}}Qx + cx - yb$$

$$\text{subject to} \quad (x, y) \in \{(x, y) \colon Ax = b,\ yA \leqslant x^{\mathrm{T}}Q + c,\ x \geqslant 0\}.$$

In QPD, the optimal objective value is known to be zero (except when QPD is infeasible). In addition, the objective function of QPD remains convex quadratic and the constraints of QPD remain linear. In fact, let $s = x^{\mathrm{T}}Q + c - yA$. The objective function becomes sx and the potential function becomes

$$P(x, s) = (2n + 1)\ln(sx) - \sum_{i=1}^{n} \ln(s_i x_i).$$

Applying Algorithm 3.1 and Procedure 4.1 to QPD, we derive

Theorem 4.1. *Convex quadratic programming can be correctly solved in* $O(L^2n^4)$ *arithmetical operations using Algorithm* 3.1 *coupled with Procedure* 4.1. □

5. Relation to the ellipsoid method

Let us briefly review how the ellipsoid method solves linear and convex programs (Grötschel et al. [10], and Papadimitriou and Steiglitz [19]). The ellipsoid was designed for convex programming in QD form. Let the optimal solution set of QD be Y^*, i.e.,

$$Y^* = \{(y, y) \colon (u, y) \in Y \text{ and } d(u, y) \geqslant z^*\}.$$

Then, the ellipsoid method generates a sequence of ellipsoids $\{E^k\}$ containing at least one solution in Y^*. The volume of E^k is reduced globally at a fixed ratio. More precisely,

$$\frac{V(E^k)}{V(E^0)} \leqslant 2^{-O(k/n)}. \tag{5.1}$$

Therefore, after enough iterations either we must discover a solution, or else we must be certain that through successive shrinkings the ellipsoid has become too small to contain Y^* and conclude $Y^* = \emptyset$. Each update of the ellipsoid requires $O(Ln^2)$ arithmetic operations. One can analyze that the overall complexity is $O(L^2n^4)$.

We now derive a dual containing ellipsoid in Algorithm 3.1. For all optimal dual feasible solution $(u^*, y^*) \in Y^*$ with $z^* = 0$, we must have

$$y^* A \leq u^{*\mathrm{T}} Q + c \quad \text{and} \quad d(u^*, y^*) \geq 0. \tag{5.2}$$

Let

$$S^* = \{s^* \in R^{n+1}\colon s^* = (u^{*\mathrm{T}} Q + c - y^* A,\, d(u^*, y^*)) \text{ and } s^* \geq 0\}$$

containing all the optimal dual-slack solutions, and

$$\bar{D} = \begin{pmatrix} \operatorname{diag}(x^k) & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\|s^* \bar{D}\|^2 \leq (s^* \bar{D} e)^2 \leq (f(x^k))^2. \tag{5.3}$$

Let S^k be the ellipsoid

$$S^k = \{s \in R^{n+1}\colon \|s\bar{D}\| \leq f(x^k)\}.$$

Then, we must have

$$S^* \subset S^k. \tag{5.4}$$

Furthermore, the volume of S^k is

$$V(S^k) = \frac{\pi (f(x^k))^{n+1}}{\det(\bar{D})} = \frac{\pi (f(x^k))^{n+1}}{\prod_{i=1}^{n} x_i^k}.$$

where π is the volume of the unit ball in R^{n+1}. In other words,

$$\ln V(S^k) = P(x^k) + \ln(\pi), \tag{5.5}$$

i.e., the potential function correctly represents the logarithmic volume of the ellipsoid S^k (only differs by a constant). Recall the proof of Theorem 3.1, (5.4), and (5.5), we can derive

Theorem 5.1. *Let*

$$V(S^0) = \frac{\pi f(x^0)^{n+1}}{\prod_{i=1}^{n} x_i^0} \; (\leq 2^{\mathrm{O}(Ln)}).$$

Then, for all k,

$$S^* \subset S^k,$$

and

$$\frac{V(S^k)}{V(S^0)} \leq \exp(-k\alpha) \leq 2^{-O(k)}. \qquad \square$$

Theorem 5.1 indicates that S^k contains all the optimal dual-slack solutions (s^*) and the volume of S^k shrinks at the ratio $2^{-O(1)}$, comparing to the shrinking ratio $2^{-O(1/n)}$ of the ellipsoid method given in (5.1). However, each iteration of this extension of Karmarkar's algorithm costs $O(Ln^3)$ arithmetic operations, and each step of the ellipsoid method costs $O(Ln^2)$ arithmetic operations. Therefore, the average shrinking ratio of these two algorithms is identical. Kapoor and Vaidya [12] proposed the rank-one updating technique similar to that of Karmarkar's [13] to solve sub-optimization problem, which reduces the complexity by a factor $n^{0.33}$.

On the other hand, we can see the following two differences between the ellipsoid method and Karmarkar's algorithm.

(1) The initial volume of the containing ellipsoid in the ellipsoid method has to take the worst case bound $n^n 2^{Ln}$, but the initial volume of the containing ellipsoid in this extension of Karmarkar's algorithm is naturally bounded by the initial potential value in the primal space, which is much less than $n^n 2^{Ln}$.

(2) The step length in the ellipsoid method is limited to a certain size, therefore, the theoretical shrinking ratio is strictly true for the ellipsoid method. However, the step size β (or μ) in this extension of Karmarkar's algorithm can be taken larger (or smaller) to minimize the potential function, resulting in a much greater reduction in the volume of the containing ellipsoid.

6. Conclusion

Using Karmarkar's polynomial-time LP algorithm as a base [13], we developed a polynomial-time algorithm for convex quadratic programming. More precisely, the number of iterations of the algorithm is bounded by $O(Ln)$, and each iteration requires $O(Ln^3)$ arithmetic operations, where L is the number of input bits and n is the dimension of the problem.

We also showed that this extension is closely related to the ellipsoid method. The potential function, which is used to measure convergence of the primal solutions in this extension, correctly represents the logarithmic volume of the dual ellipsoid containing all the optimal dual-slack solutions. Like in the ellipsoid method, the volume of this containing ellipsoid uniformly shrinks to zero as the algorithm iterates.

Unlike the ellipsoid method, this algorithm has an efficient implementation in practice. Our computational experiments show that each iteration can be computed in $O(n^3)$ arithmetic operations. Regardless of the size of the program, the total number of iterations required to achieve 6-digit optimality accuracy is about 20 for a group of QP test problems (Ye [26]).

References

[1] K. M. Anstreicher, "A monotonic projective algorithm for fractional linear programming," *Algorithmica* 1 (1986) 483-498.

[2] E.R. Barnes, "A variation on Karmarkar's algorithm for solving linear programming," *Mathematical Programming* 36 (1986) 174-182.

[3] E.M.L. Beale, "On quadratic programming," *Naval Research Logistics Quarterly* 6 (1959) 227-243.

[4] A.R. Conn and J. W. Sinclair, "Quadratic programming via a non-differentiable penalty function," Report 75/15, Department of Combinatorics and Optimization, University of Waterloo (Waterloo, Canada, 1975).

[5] R.W. Cottle and G.B. Dantzig, "Complementary pivot theory of mathematical programming," *Linear Algebra and Its Applications* 1 (1968) 103-125.

[6] G.B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, 1963).

[7] I.I. Dikin, "Iterative solution of problems of linear and quadratic programming," *Soviet Mathematics Doklady* 8 (1967) 674-675.

[8] B.C. Eaves, "On the basic theorem of complementarity," *Mathematical Programming* 1 (1971) 68-75.

[9] P.E. Gill and W. Murray, "Newton type methods for unconstrained and linearly constrained optimization," *Mathematical Programming* 7 (1974) 311-350.

[10] M. Grötschel, L. Lovász and A. Schrijver, "The ellipsoid method and its consequences in combinatorial optimization," *Combinatorica* 1 (1981) 169-197.

[11] C. Hildreth, "A quadratic programming procedure," *Naval Research Logistics Quarterly* 4 (1957) 79-85.

[12] S. Kapoor and P. Vaidya, "Fast algorithms for Convex quadratic programming and multicommodity flows," *Proceedings of the 18th Annual ACM Symposium on the Theory of Computing* (1986) 147-159.

[13] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica* 4 (1984) 373-395.

[14] L.G. Khachiyan, "A polynomial algorithm for linear programming," *Soviet Mathematics Doklady* 20 (1979) 191-194.

[15] K.O. Kortanek and M. Shi, "Convergence results and numerical experiments on a linear programming hybrid algorithm," to appear in the European *Journal of Operational Research*, Dept. of Mathematics, Carnegie Mellon University (PA, 1985).

[16] M.K. Kozlov, S.P. Tarasov and L.G. Khachiyan, "Polynomial solvability of convex quadratic programming," *Soviet Mathematics Doklady* 20 (1979) 1108-1111.

[17] C.E. Lemke, "Bimatrix equilibrium points and mathematical programming," *Management Science* 11 (1965) 681-689.

[18] D.G. Luenberger, *Linear and Nonlinear Programming* (2nd Edition, Addison-Wesley, Menlo Park, 1984).

[19] C.H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Prentice-Hall, Englewood Cliffs, 1982).

[20] J. Rosen, "The gradient projection method for nonlinear programming, I. Linear constraints," *SIAM Journal on Applied Mathematics* 12 (1964) 74-92.

[21] D.C. Sorensen, "Trust region methods for unconstrained minimization," in: M.J.D. Powell, ed., *Nonlinear Optimization* (Academic Press, New York, 1981).

[22] M.J. Todd and B.P. Burrell, "An extension of Karmarkar's algorithm for linear programming using dual variables," *Algorithmica* 1 (1986) 409-424.

[23] R.J. Vanderbei, M.S. Meketon and B.A. Freedman, "On a modification of Karmarkar's linear programming algorithm," *Algorithmica* 1 (1986) 395-407.

[24] P. Wolfe, "The simplex algorithm for quadratic programming," *Econometrica* 27 (1959) 382–398.
[25] Y. Ye and M. Kojima, "Recovering optimal dual solutions in Karmarkar's algorithm for linear programming," *Mathematical Programming* 39 (1987) 305–317.
[26] Y. Ye, "Interior algorithms for linear, quadratic, and linearly constrained convex programming," Ph.D. Thesis, Dept. of Engineering-Economic Systems, Stanford University (CA, 1987).