ææ

An Algorithm for Convex Quadratic Programming that Requires $O(n^{3.5}L)$ Arithmetic Operations*

Sanjay Mehrotra and Jie Sun

Department of Industrial Engineering and Management Sciences
Northwestern University, Evanston, IL 60208, U.S.A.

Abstract

A new interior point method for minimizing a convex quadratic function over a polytope is developed. We show that our method requires $O(n^{3.5}L)$ arithmetic operations. In our algorithm we construct a sequence $P_{z^0}, P_{z^1}, \ldots, P_{z^k}, \ldots$ of nested convex sets that shrink towards the set of optimal solution(s). During iteration k we take a partial Newton step to move from an approximate analytic center of $P_{z^{k-1}}$ to an approximate analytic center of P_{z^k}. A system of linear equations is solved at each iteration to find the step direction. The solution that is available after $O(\sqrt{m}L)$ iterations can be converted to an optimal solution. Our analysis indicates that inexact solutions to the linear system of equations could be used in implementing this algorithm.

Key Words: analytic center, convex quadratic programming, interior point methods, Karmarkar's algorithm, method of centers.

* Mathematics of Operations Research 15 (1990) 342-363. This research has been supported in part by the ONR under grant N00014-87-K- 0214 and by the NSF under grant ECS-8721709.

1. Introduction

We consider the convex quadratic program

$$(QP)\begin{cases} \text{maximize} & q(x) \equiv \frac{1}{2}x^TQx + c^Tx \\ \text{subject to} & x \in P \equiv \{x \mid a_i^Tx \geq b_i,\ i = 1, 2, \ldots m\}, \end{cases} \tag{1.1}$$

where $x \in R^n,\ c \in Z^n,\ a_i \in Z^n$ and $Q \in Z^{n\times n}$ is a symmetric negative semidefinite matrix. We assume that $m = O(n)$ to simplify our presentation.

Several methods have been proposed to solve the above (QP). Among these methods the active set methods are computationally most effective [5]. Unfortunately, however, in the worst case the amount of work required to solve (QP) by active set methods may grow exponentially in n. Kozlov et. al. [15] proposed the first polynomial-time algorithm to solve (QP). They showed that the ellipsoid method of Iudin and Nemirovskii [10] and Shor [23] can be modified to solve (QP) in polynomial time. However, after Karmarkar's method for linear programming [12], the interior methods have caught the attention of many researchers.

One key notion that has emerged from the studies [1,16] on Karmarkar's algorithm is the notion of analytic centers. Let $\operatorname{int} P \equiv \{x \in R^n \mid a_i^Tx > b_i,\ i = 1, 2, \ldots m\}$ and consider the convex set

$$P_z \equiv \{x \in R^n \mid q(x) \geq z, x \in P\}. \tag{1.2}$$

Let us assume that P is bounded, and $\operatorname{int} P_z = \{x \in R^n \mid q(x) > z,\ x \in \operatorname{int} P\}$ is nonempty.

A point $\omega \in R^n$ is called the analytical center of P_z if it maximizes

$$F(x, z) \equiv m \ln(q(x) - z) + \sum_{i=1}^{m} \ln(a_i^Tx - b_i) \tag{1.3}$$

subject to $x \in \mathrm{int}\, P_z$. The function $F(x,z)$ defined in (1.3) is called the potential function. Note that $F(x,z)$ is a strictly concave function of x and the analytic center $\omega \in \mathrm{int}\, P_z$ is unique. The concept of centers introduced by Huard [9] and the notion of logarithmic barrier function [2,3] are related to analytic centers.

The first algorithm for (QP) that is based upon ideas of Karmarkar [12] for solving linear programs (LP) was due to Kapoor and Vaidya [11]. Their algorithm requires $O(n^{3.67} \ln n \; L \ln L)$ arithmetic operations, where

$$L = n^2 + mn + \ln |P|$$

is the input length of (QP). P is the product of the nonzero integer coefficients appearing in Q, c, and a_i, b_i, $i = 1, 2, \ldots m$. Several authors have recently proposed new algorithms for improving the worst case bound for solving (QP). All of these algorithms are motivated by a logarithmic barrier function approach [2,3]. Kojima Mizuno and Yoshise [14] developed an algorithm for solving a linear complementarity problem, in which they extended their previous work on linear programming [13]. Monteiro and Adler [21] proposed algorithms that work in a primal-dual framework. The algorithms in [14] and [21] take a Newton step to approximately solve for the KKT optimality conditions for the logarithmic barrier problem with a choice of barrier parameter. These algorithms appear to have been motivated from the work of Megiddo [16]. In a yet another development Ye [27] extended the results of Gonzaga [8] for solving (QP). Ye works directly with the logarithmic barrier problem. The worst case bound on (QP) that is proved by Ye is of the same order as ours. Goldfarb and Liu [7] also developed an algorithm that is similar to the one given by Ye [27]. However, Goldfarb and Liu [7] showed that Karmarkar's idea of using rank one updates can be incorporated in their analysis to get an $O(\sqrt{n})$ improvement in the worst case bound for solving (QP). Similar idea is also used by Kojima Mizuno and Yoshise [14] and Monteiro and Adler [21] to prove improved bounds for their algorithms.

One feature that is common to all the above mentioned algorithms is that they explicitly maintain primal and dual feasibility. The duality relationship is used in proving convergence and the improved bounds in these algorithms. In this paper we propose an algorithm for (QP) that does not depend on duality analysis and has a clear geometric interpretation. In our algorithm we construct a sequence $P_{z^0}, P_{z^1}, \ldots, P_{z^k}, \ldots$ of nested convex sets that shrink towards the set of optimal solution(s). During iteration k we take a partial Newton step to move from an approximate analytic center of $P_{z^{k-1}}$ to an approximate analytic center of P_{z^k}. A system of linear equations is solved at each iteration to find the step direction. The solution that is available after $O(\sqrt{m}L)$ iterations can be converted to an optimal solution. Our work is influenced by the work of Bayer and Lagarias [1] Renegar [22], Sonnevend [24], and Vaidya [25] in their study of methods of analytical centers for linear programming.

Although the worst case bound of our algorithm is not the best in comparison with the ones in [7], [14] and [21] ($O(\sqrt{n})$ slower), we feel that it is valuable in the following sense. It is the only primal algorithm that reaches the order of $O(\sqrt{m}n^3L)$ operations without explicitly requiring dual feasibility; and its analysis is based on estimating the Taylor expansion of the log-quadratic function (1.3). In our continuing research on interior point methods this approach has provided us an extremely useful framework for analyzing interior point methods in more general settings. The results obtained by us are reported in [17,19] for quadratically constrained convex quadratic programs and in [18] for convex programming problems. Furthermore, as a consequence of our analysis we also get a theoretical framework for computing inexact search directions. The ability to compute inexact search directions is desirable for efficient implementations of interior point methods.

This paper is organized as follows. In the next section we provide some basic properties of the potential function $F(x, z)$. In Section 3 we state our basic algorithm and give several results to establish its convergence. Several lemmas are proved in Section 4. Section 5 discusses an approach to satisfy the initial assumptions.

2. Analytic Centers and Potential Functions

Let $F(x, z)$ be as in (1.3) and ω be the corresponding analytic center. Let

$$f(x, z) \equiv F(\omega, z) - F(x, z).$$

The function $f(x, z)$ is called the normalized potential function. Note that the function $f(x, z)$ is convex in x and its minimizer over P_z is also ω. In this section we give several properties of the normalized potential function and the analytic center that are used frequently in our paper. Let $\nabla f(x, z)$ and $\nabla^2 f(x, z)$ represent the gradient and Hessian of $f(x, z)$ with respect to x. It is easy to see that

$$\nabla f(x, z) = -\frac{m \nabla q(x)}{q(x) - z} - \sum_{i=1}^{m} \frac{a_i}{a_i^T x - b_i} \tag{2.1}$$

and

$$\nabla^2 f(x, z) = \frac{m}{(q(x) - z)^2} \nabla q(x) (\nabla q(x))^T + \sum_{i=1}^{m} \frac{1}{(a_i^T x - b_i)^2} a_i a_i^T - \frac{m}{q(x) - z} Q. \tag{2.2}$$

Note that the gradient of $f(x,z)$ vanishes at ω, i.e., $\nabla f(\omega,z)=0$. The following lemma gives the complete Taylor's expansion of $f(y',z)$ at y. A proof of this lemma is provided in Section 4.

Lemma 2.1 Let $y' \in \operatorname{int} P_z$ and $h = y' - y$. The Taylor expansion of $f(y',z)$ at y is given by

$$f(y',z) - f(y,z) = \sum_{j=1}^{\infty}\Big[\frac{1}{j}\sum_{i=1}^{m}(C_i(y))^j + \frac{m}{j!}\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(y))^{j-2i}(B(y))^i\Big],$$

where $A(y)$, $B(y)$ and $C_i(y)$ are obtained by evaluating

$$A(x) = -\frac{\nabla q(x)^T h}{q(x)-z},\ B(x) = -\frac{h^TQh}{q(x)-z},\ C_i(x) = -\frac{a_i^T h}{a_i^T x - b_i},$$

at $x = y$ respectively. The coefficients α_{ji} are defined by the following recursive relationships:

$$\begin{cases} \alpha_{1,0} = 1 & \\ \alpha_{j+1,0} = j\alpha_{j,0} & j = 1,2,... \\ \alpha_{j+1,i} = (j-i)\alpha_{ji} + (j-2i+2)\alpha_{j,i-1} & j=1,2,...,i=1,2,...,\lfloor j/2\rfloor \\ \alpha_{ji} = 0 & j = 1,2,...,\quad i > \lfloor j/2\rfloor. \end{cases}$$

$\lfloor j/2\rfloor$ is the largest integer that is less than or equal to $j/2$.

Let $E(y,r) \equiv \{x \mid (x-y)^T\nabla^2 f(y,z)(x-y) \le r^2\}$ be the ellipsoid of radius r around y. This choice of ellipsoid provides us with several desirable properties. It can be shown that $E(y,1) \subset P_z \subset E(\omega, 2m)$. The next lemma shows that the normalized potential function is well behaved on points in ellipsoid $E(y,.5)$. A proof of this lemma is given in Section 4.

Lemma 2.2 If $x \in E(y,r)$, $0 \le r < 0.5$, then

$$\Big|f(x,z) - f(y,z) - \nabla f(y,z)^T(x-y) - \frac{1}{2}(x-y)^T\nabla^2 f(y,z)(x-y)\Big| \le \frac{5r^3}{3(1-2r)}.$$

The following are simple estimations of the objective function value and slacks in the constraints at $x \in E(y,r)$.

Lemma 2.3 If $x \in E(y,r)$, then

$$\Big|\frac{a_i^T(x-y)}{a_i^T y - b_i}\Big| \le r, \qquad 1 \le i \le m \tag{2.3}$$

and

$$\frac{\big|q(y)-q(x)\big|}{q(y)-z} \le \frac{r}{\sqrt{m}} + \frac{r^2}{2m}. \tag{2.4}$$

Proof: Since $x \in E(y,r)$, we have

$$m\Big(\frac{(\nabla q(y))^T(x-y)}{q(y)-z}\Big)^2 + \sum_{i=1}^m \Big(\frac{a_i^T(x-y)}{a_i^T y - b_i}\Big)^2 - \frac{m}{q(y)-z}(x-y)^T Q(x-y) \le r^2. \tag{2.5}$$

Since Q is symmetric negative semidefinite, (2.3) follows. In order to prove (2.4), note that

$$\frac{\big|q(y)-q(x)\big|}{q(y)-z} \le \frac{\big|\nabla q(y)^T(x-y)\big|}{q(y)-z} + \frac{1}{2}\frac{\big|(x-y)^T Q(x-y)\big|}{q(y)-z},$$

$$\frac{\big|\nabla q(y)^T(x-y)\big|}{q(y)-z} \le \frac{r}{\sqrt{m}}, \qquad \text{(by using (2.5))}$$

and

$$\frac{\big|(x-y)^T Q(x-y)\big|}{q(y)-z} \le \frac{r^2}{m}. \qquad \text{(by using (2.5))}$$

($\mathcal{Q.E.D.}$)

We now provide a lemma that shows that the analytical center ω is a "good representative" of the convex set P_z.

Lemma 2.4 The analytic center ω of the convex set P_z ensures that,

$$\frac{q(x)-z}{q(\omega)-z} \leq 2, \qquad \forall x \in P_z, \tag{2.6}$$

and

$$\frac{q(x)-z}{q(\omega)-z} \geq 1 - \frac{\delta}{\sqrt{m}} - \frac{\delta^2}{2m}, \quad \forall x \in E(\omega, \delta). \tag{2.7}$$

Proof: Note that

$$q(x) - q(\omega) = \nabla q(\omega)^T (x-\omega) + \frac{1}{2}(x-\omega)^T Q (x-\omega). \tag{2.8}$$

By using (2.8) and the fact that $\nabla f(\omega, z) = 0$, we have

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^{m} \frac{a_i^T x - b_i}{a_i^T \omega - b_i} + \frac{q(x)-z}{q(\omega)-z} \\
=&\frac{1}{m}\sum_{i=1}^{m}(1 + \frac{a_i^T (x-\omega)}{a_i^T \omega - b_i}) + \frac{1}{q(\omega)-z}\Big[q(\omega) - z + \nabla q(\omega)^T (x-\omega) + \frac{1}{2}(x-\omega)^T Q (x-\omega)\Big] \\
=&1 - \frac{\nabla f(\omega, z)^T (x-\omega)}{m} + 1 + \frac{1}{2}\frac{(x-\omega)^T Q (x-\omega)}{q(\omega)-z} \\
=&2 + \frac{1}{2}\frac{(x-\omega)^T Q (x-\omega)}{q(\omega)-z}.
\end{aligned}$$

Since Q is symmetric negative semidefinite, (2.6) follows. The inequality (2.7) is a simple consequence of Lemma 2.3.

3. Development of the Basic Algorithm

In this section we first study the following basic algorithm for (QP).

Algorithm 3.1

Input: z^0, $x^0 \in \operatorname{int} P_{z^0}$.

For $k = 0, 1, 2, \ldots,$ **until** a termination criterion is satisfied **do**:

Starting from x^k find ω^k the analytic center of P_{z^k} ;

$x^{k+1} \leftarrow \omega^k$;

$z^{k+1} \leftarrow z^k + \beta^{k+1}$, $0 < \beta^{k+1} < q(x^{k+1}) - q(x^k)$.

The Algorithm 3.1 is a restatement of the method of centers [9]. For a choice of β^{k+1} that is arbitrarily close to $q(x^{k+1}) - q(x^k)$, it can be shown from Lemma 2.4 that Algorithm 3.1 would require at most $O(L)$ iterations to produce a "good approximation" to the optimal solution of (QP). However, the problem of finding the analytic center (or a good approximation to it) at each iteration is "difficult". On the contrary the problem of finding the analytic center (or a good approximation to it) is "relatively easy" if β^{k+1} is close to zero but then Algorithm 3.1 would require a much larger number of steps to converge to the optimal solution.

The algorithm that we develop in this paper makes a compromise between these two extreme situations. At the beginning of iteration k we have a good approximation x^k of the analytic center ω^k of the convex set P_{z^k}. The normalized potential function provides the metric that is used to measure the distance. We take a partial Newton step at x^k and move to x^{k+1}. The point x^{k+1} is closer to ω^k than x^k. Finally z^k is appropriately increased to z^{k+1} to obtain $P_{z^{k+1}}$ such that x^{k+1} also serves as a good approximation to ω^{k+1}.

Let z^* denote the optimal objective function value of (QP). We make the following assumptions.

A1. P is a bounded polytope;

A2. $\operatorname{int} P$ is nonempty;

A3. A lower bound z_0 on the objective value of (QP) is available such that $z^* - z_0 \leq 2^{O(L)}$;

A4. A solution $x_0 \in \operatorname{int} P_{z_0}$ is known and $f(x_0, z_0) \leq .003$.

An approach to cast any given (QP) to the problem form that satisfies these assumptions is given in Section 5. We now outline our algorithm.

Algorithm 3.2

Initialization:

$$z^0 \leftarrow z_0 \quad \text{and} \quad x^0 \leftarrow x_0.$$

For $k = 0, 1, \ldots$ **until** termination criterion is satisfied **do**:

Determine a step direction p by solving

$$\nabla^2 f(x^k, z^k)\; p = -\nabla f(x^k, z^k).$$

Let

$$x^{k+1} \leftarrow x^k + \frac{\epsilon}{\sqrt{p^T \nabla^2 f(x^k, z^k) p}} p,$$

where $\epsilon = .03$, and

$$z^{k+1} \leftarrow z^k + \frac{\alpha}{\sqrt{m}} (q(x^{k+1}) - z^k),$$

with $\alpha = .001$,

The values $\epsilon = .03$ and $\alpha = .001$ are chosen for our analysis. In practice ϵ may be obtained by doing a one dimensional line search to maximize the potential function $F(x^k, z^k)$.

We now prove the following convergence theorem for Algorithm 3.2.

Theorem 3.3 Let z^* denote the optimal objective value of (QP), then at iteration k of Algorithm 3.2 we have

$$\frac{z^* - z^{k+1}}{z^* - z^k} \leq (1 - \frac{.0004}{\sqrt{m}}). \tag{3.1}$$

We need to state several lemmas that are used to prove Theorem 3.3. Proofs for all these lemmas are provided in the next section. Let us fix iteration k in Algorithm 3.2 and represent x^k, x^{k+1}, z^k and z^{k+1} by x, x^+, z and z^+ respectively.

The Lemma 3.4 shows that if at a given point x the value of the normalized potential function is "small", then x is "close to" ω. Lemma 3.5 shows that whenever x is "close to" ω, the value of the normalized potential function can be reduced by a "sufficient" amount, by taking a partial Newton step at x and moving to x^+. Lemma 3.6 shows that if x^+ is "sufficiently close" to ω then it remains "close to" the analytic center ω^+ that is defined for the convex set P_{z^+}. The improved lower bound z^+ is obtained by adding a fraction of $q(x^+) - z$ to z.

Lemma 3.4 Let $0 \leq \delta < 0.5$, and

$$f(x, z) \leq \frac{\delta^2}{2} - \frac{5\delta^3}{3(1 - 2\delta)},$$

then $x \in E(\omega, \delta)$.

Lemma 3.5 Let $x \in E(\omega, \delta)$, $0 \leq \delta < .243$ and ϵ be a parameter such that $0 \leq \epsilon < 0.5$. The point x^+ that minimizes $\nabla f(x,z)^T y$ over $E(x, \epsilon) \equiv \{y \mid (y-x)^T \nabla^2 f(x,z)(y-x) \leq \epsilon^2\}$ satisfies,

$$f(x^+, z) \leq f(x, z) - \epsilon\sqrt{f(x,z)\Big(1 - \frac{7\delta}{2} - \frac{5\delta^2}{2}\Big)} + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)}. \qquad (3.2)$$

Lemma 3.6 Let $z^+ = z + \frac{\alpha}{\sqrt{m}}(q(x^+) - z)$, $0 < \alpha < \sqrt{m}$ and $x^+ \in \operatorname{int} P_{z^+}$. If

$$f(x^+, z) \leq \frac{\delta^{+2}}{2} - \frac{5\delta^{+3}}{3(1-2\delta^+)},$$

where $0 \leq \delta^+ < 0.5$, then we have

$$f(x^+, z^+) \leq f(x^+, z) + \frac{m\alpha}{\sqrt{m} - \alpha}\Big(\frac{\delta^+}{\sqrt{m}} + \frac{\delta^{+2}}{2m}\Big) + \frac{\alpha^2(1 + \frac{\delta^+}{\sqrt{m}} + \frac{\delta^{+2}}{2m})^2}{1 - \frac{\alpha}{\sqrt{m}}(1 + \frac{\delta^+}{\sqrt{m}} + \frac{\delta^{+2}}{2m})}.$$

Figure 3.1 graphically represents ellipsoids $E(\omega, \delta)$, $E(\omega^+, \delta)$ and $E(x, \epsilon)$ used in Lemma 3.5, Lemma 3.6 and in the proof of Theorem 3.3.

Proof of Theorem 3.3 Without loss of generality let us assume that $f(x^k, z^k) \leq .003$. Since $f(x^k, z^k) \leq .003$ from Lemma 3.4 it is easy to see that $x^k \in E(\omega^k, .2)$. For the choice of $\epsilon = .03$, from Lemma 3.5 we can show that $f(x^{k+1}, z^k) \leq .00277$. This implies that $x^{k+1} \in E(\omega^k, \delta^+), \delta^+ < .2$. Let us take $\alpha = 0.001$. Since $m \geq 2$ for a bounded polytope, Lemma 3.6 implies that $f(x^{k+1}, z^{k+1}) \leq .003$. Now since by Lemma 2.4,

$$q(x^{k+1}) - z^k \geq (1 - \frac{\delta^+}{\sqrt{m}} - \frac{\delta^{+2}}{2m})(q(\omega^k) - z^k) \geq .84(q(\omega^k) - z^k) \geq .42(z^* - z^k),$$

we have

$$z^* - z^{k+1} = z^* - z^k - \frac{\alpha}{\sqrt{m}}(q(x^{k+1}) - z^k) \leq (1 - \frac{0.001 \times 0.42}{\sqrt{m}})(z^* - z^k).$$

The proof of Theorem 3.3 is now complete.

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Algorithm 3.2 finds the search direction at each iteration by solving a system of linear equations. This can be done in $O(n^3)$ arithmetic operations. From Theorem 3.3 and the fact that $z^* - z^0 \leq 2^{O(L)}$ it follows that $z^* - z^k \leq 2^{-\theta L}$ in $O(\sqrt{m}L)$ iterations for a given constant θ. From this point we may jump to an exact optimal solution of (QP) by using continued fractions (see e.g., [11,15]).

The main computational work in the implementation of the Algorithm 3.2 involves solving a system of linear equation:

$$Mp = -\nabla f(x, z), \tag{3.3}$$

where

$$M = \sum_{i=1}^{m} \frac{1}{(a_i^T x - b_i)^2} a_i a_i^T + \frac{m}{(q(x) - z)^2} \nabla q(x) \nabla q(x)^T - \frac{m}{(q(x) - z)} Q.$$

The matrix M defining the system of linear equation (3.3) is symmetric and positive definite. Direct (e.g., symmetric Gaussian elimination) and iterative methods (e.g., preconditioned conjugate gradient method) may be used to find the solution of (3.3). The detailed implementation of these methods would be much in the spirit of implementations of Karmarkar's algorithm (and related interior point methods) for linear programming. In particular, the fact that the system of equations that are solved at each iteration are symbolically and numerically related should be used for developing efficient implementations.

If iterative methods are used, it is computationally expensive to solve (3.3) exactly. The following lemma is important for developing implementations of Algorithm 3.2 that use iterative methods. In this lemma we observe that inexact solutions of (3.3) also provide search directions along which the normalized potential function is reduced by the desired amount.

Lemma 3.7 Let $x \in E(\omega, \delta),\ 0 \leq \delta < .243$ and ϵ be a parameter such that $0 \leq \epsilon < 0.5$. Let $\bar{x}$ be the point where the line joining x to the analytic center ω intersects with the boundary of ellipsoid $E(x, \epsilon)$. Let x^+ be a point that satisfies

$$\nabla f(x,z)^T x^+ \leq \nabla f(x,z)^T \bar{x}.$$

Then we have,

$$f(x^+, z) \leq f(x,z) - \epsilon \sqrt{f(x,z)\Big(1 - \frac{7\delta}{2} - \frac{5\delta^2}{2}\Big)} + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)}.$$

The above lemma is a restatement of Lemma 3.5 under a weaker condition. The proof of Lemma 3.7 is same as that of Lemma 3.5. The proof of Theorem 3.3 does not change if Lemma 3.5 is replaced by Lemma 3.7.

4. Analysis

We now provide several lemmas and their proofs that were stated in Section 2 and Section 3. At our convenience we list these lemmas in their logical order. However, if a lemma or a theorem was stated before, its old number is bracketed immediately after its new number.

Lemma 4.1 [Lemma 2.1] Let $y' \in \operatorname{int} P_z$ and $h = y' - y$. The Taylor expansion of $f(y', z)$ at y is given by

$$f(y', z) - f(y, z) = \sum_{j=1}^{\infty} \Big[\frac{1}{j} \sum_{i=1}^{m} (C_i(y))^j + \frac{m}{j!} \sum_{i=0}^{\lfloor j/2 \rfloor} \alpha_{ji} (A(y))^{j-2i} (B(y))^i \Big],$$

where $A(y)$, $B(y)$ and $C_i(y)$ are obtained by evaluating

$$A(x) = -\frac{\nabla q(x)^T h}{q(x) - z}, \; B(x) = -\frac{h^T Q h}{q(x) - z}, \; C_i(x) = -\frac{a_i^T h}{a_i^T x - b_i},$$

at $x = y$ respectively. The coefficients α_{ji} are defined by the following recursive relationships:

$$\begin{cases} \alpha_{1,0} = 1 & \\ \alpha_{j+1,0} = j\alpha_{j,0} & j = 1, 2, ... \\ \alpha_{j+1,i} = (j-i)\alpha_{ji} + (j - 2i + 2)\alpha_{j,i-1} & j = 1, 2, ..., i = 1, 2, ..., \lfloor j/2 \rfloor \\ \alpha_{ji} = 0 & j = 1, 2, ..., \quad i > \lfloor j/2 \rfloor. \end{cases}$$

Proof: Since

$$f(y', z) - f(y, z) = \sum_{j=1}^{\infty} \frac{1}{j!} (\frac{\partial}{\partial x_1} h_1 + \cdots + \frac{\partial}{\partial x_n} h_n)^j f(x, z)\Big|_{x=y}$$

and

$$(\frac{\partial}{\partial x_1} h_1 + \cdots + \frac{\partial}{\partial x_n} h_n)^j [-\ln(a_i^T x - b_i)] = (j-1)!(C_i(x))^j \quad j = 1, 2, ..., \qquad (4.1)$$

we only need to show that

$$(\frac{\partial}{\partial x_1}h_1+\cdots+\frac{\partial}{\partial x_n}h_n)^j\ln(q(x)-z)=-\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(x))^{j-2i}(B(x))^i. \tag{4.2}$$

We prove this by induction. It can be easily verified that (4.2) is valid for $j=1$ and $j=2$. Now suppose that (4.2) is valid for j. We have

$$\begin{aligned}&(\frac{\partial}{\partial x_1}h_1+\cdots+\frac{\partial}{\partial x_n}h_n)^{j+1}\ln(q(x)-z)\\=&(\frac{\partial}{\partial x_1}h_1+\cdots+\frac{\partial}{\partial x_n}h_n)(-\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(x))^{j-2i}(B(x))^i).\end{aligned}$$

The identities

$$(\frac{\partial}{\partial x_1}h_1+\cdots+\frac{\partial}{\partial x_n}h_n)A(x)=(A(x))^2+B(x)$$

and

$$(\frac{\partial}{\partial x_1}h_1+\cdots+\frac{\partial}{\partial x_n}h_n)B(x)=A(x)B(x)$$

are used in the sequel.

If $j = 2k$, we get

$$(\frac{\partial}{\partial x_1}h_1 + \cdots + \frac{\partial}{\partial x_n}h_n)[-\sum_{i=0}^{\lfloor j/2 \rfloor} \alpha_{ji}(A(x))^{j-2i}(B(x))^i]$$

$$\begin{aligned}
=& -\alpha_{2k,0}(2k)(A(x))^{2k-1}(A^2(x)+B(x)) \\
&\quad -\sum_{i=1}^{k-1} \alpha_{2k,i}\Big[(2k-2i)(A(x))^{2k-2i-1}(B(x))^i((A(x))^2+B(x)) \\
&\qquad + i(A(x))^{2k-2i}(B(x))^{i-1}A(x)B(x)\Big] - k(B(x))^{k-1}A(x)B(x)
\end{aligned}$$

$$\begin{aligned}
=& -2k\alpha_{2k,0}(A(x))^{2k+1} \\
&\quad -\sum_{i=1}^{k}\Big[(2k-i)\alpha_{2k,i} + (2k-2i+2)\alpha_{2k,i-1}\Big](A(x))^{2k-2i+1}(B(x))^i
\end{aligned}$$

$$= -\sum_{i=0}^{\lfloor (j+1)/2 \rfloor} \alpha_{j+1,i}(A(x))^{j+1-2i}(B(x))^i,$$

and if $j = 2k+1$, we get

$$(\frac{\partial}{\partial x_1}h_1 + \cdots + \frac{\partial}{\partial x_n}h_n)(-\sum_{i=0}^{\lfloor j/2 \rfloor} \alpha_{ji}(A(x))^{j-2i}(B(x))^i)$$

$$\begin{aligned}
= &- \alpha_{2k+1,0}(2k+1)(A(x))^{2k}(A^2(x)+B(x)) \\
&- \sum_{i=1}^{k} \alpha_{2k+1,i}\Big[(2k+1-2i)(A(x))^{2k-2i}(B(x))^i(A^2(x)+B(x)) \\
&\quad + i(A(x))^{2k+1-2i}(B(x))^{i-1}A(x)B(x)\Big]
\end{aligned}$$

$$\begin{aligned}
= &- (2k+1)\alpha_{2k+1,0}(A(x))^{2k+2} \\
&- \sum_{i=1}^{k}\Big[(2k+1-i)\alpha_{2k+1,i} + (2k+1-2i+2)\alpha_{2k+1,i-1}\Big](A(x))^{2k-2i+2}(B(x))^i \\
&- \alpha_{2k+1,k}(B(x))^{k+1}
\end{aligned}$$

$$= -\sum_{i=0}^{k} \alpha_{2k+2,i}(A(x))^{2k+2-2i}(B(x))^i - \alpha_{2k+1,k}(B(x))^{k+1}$$

$$= -\sum_{i=0}^{\lfloor (j+1)/2 \rfloor} \alpha_{j+1,i}(A(x))^{j+1-2i}(B(x))^i.$$

This completes the induction. The proof of Lemma 4.1 follows by combining (4.1) and (4.2).

($\mathcal{Q}.\mathcal{E}.\mathcal{D}.$)

Lemma 4.2 Let α_{ji}'s be defined by the recursive relationship in Lemma 4.1, then

$$\sum_{i=0}^{\lfloor j/2 \rfloor} \alpha_{ji} \leq 2^{j-1}(j-1)!$$

.

Proof: If $j = 2k$, we have

$$\sum_{i=0}^{k} \alpha_{2k,i} = \sum_{i=1}^{k-1} [(2k-1-i)\alpha_{2k-1,i} + (2k-2i+1)\alpha_{2k-1,i-1}] + (2k-1)\alpha_{2k-1,0} + \alpha_{2k-1,k-1}$$

$$\begin{aligned} &= \sum_{i=0}^{k-1} (4k-2-3i)\alpha_{2k-1,i} \\ &\leq (4k-2)\sum_{i=0}^{k-1} \alpha_{2k-1,i} \qquad \text{(because all } \alpha_{ji} \geq 0\text{).} \end{aligned}$$

Otherwise $j = 2k+1$, and we have

$$\begin{aligned} \sum_{i=0}^{k} \alpha_{2k+1,i} &= \sum_{i=1}^{k} [(2k-i)\alpha_{2k,i} + (2k+2-2i)\alpha_{2k,i-1} + \alpha_{2k,i-1}] + 2k\alpha_{2k,0} \\ &= \sum_{i=0}^{k} (4k-3i)\alpha_{2k,i} \\ &\leq 4k \sum_{i=0}^{k} \alpha_{2k,i}. \end{aligned}$$

Hence, by letting $\beta_j = \sum_{i=0}^{\lfloor j/2 \rfloor} \alpha_{ji}$, we have

$$\beta_j \leq (2j-2)\beta_{j-1}.$$

Therefore,

$$\beta_j \leq 2^{j-1}(j-1)!\beta_1 = 2^{j-1}(j-1)!$$

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Lemma 4.3 [Lemma 2.2] If $x \in E(y, r)$, $0 \leq r < 0.5$, then

$$\left| f(x,z) - f(y,z) - \nabla f(y,z)^T(x-y) - \frac{1}{2}(x-y)^T \nabla^2 f(y,z)(x-y) \right| \leq \frac{5r^3}{3(1-2r)}.$$

Proof: Since from Lemma 4.1 one has

$$f(x,z)-f(y,z)=\sum_{j=1}^{\infty}\Big[\frac{1}{j}\sum_{i=1}^{m}(C_i(y))^j+\frac{m}{j!}\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(y))^{j-2i}(B(y))^i\Big],$$

it is sufficient to show that

$$\Big|\sum_{j=3}^{\infty}\Big[\frac{1}{j}\sum_{i=1}^{m}(C_i(y))^j+\frac{m}{j!}\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(y))^{j-2i}(B(y))^i\Big]\Big|\le\frac{5r^3}{3(1-2r)},$$

where

$$A(y)=-\frac{\nabla q(y)^T(x-y)}{q(y)-z},\ B(y)=-\frac{(x-y)^TQ(x-y)}{q(y)-z},\ C_i(y)=-\frac{a_i^T(x-y)}{a_i^Ty-b_i}.$$

Now since $x\in E(y,r)$, we have

$$\sum_{i=1}^{m}(C_i(y))^2\le r^2,\ \ m(A(y))^2\le r^2\ \text{ and }\ mB(y)\le r^2.$$

Thus

$$\Big|\sum_{j=3}^{\infty}\Big[\frac{1}{j}\sum_{i=1}^{m}(C_i(y))^j+\frac{m}{j!}\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(y))^{j-2i}(B(y))^i\Big]\Big|$$

$$\le\sum_{j=3}^{\infty}\frac{r^{j-2}}{j}\sum_{i=1}^{m}(C_i(y))^2+\Big|m\sum_{j=3}^{\infty}\Big[\frac{1}{j!}\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(A(y))^{j-2i}(B(y))^i\Big]\Big|$$

$$\le\sum_{j=3}^{\infty}\frac{r^j}{j}+m\sum_{j=3}^{\infty}\Big[\frac{1}{j!}\sum_{i=0}^{\lfloor j/2\rfloor}\alpha_{ji}(\frac{r}{\sqrt{m}})^{j-2i}(\frac{r^2}{m})^i\Big]$$

$$\le\sum_{j=3}^{\infty}\frac{r^j}{j}+m^{-\frac{1}{2}}\sum_{j=3}^{\infty}\frac{r^j}{j}(2^{j-1})\quad\text{(by Lemma 4.2)}$$

$$\le\frac{r^3}{3(1-r)}+\frac{(2r)^3}{6(1-2r)}\le\frac{5r^3}{3(1-2r)}.$$

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Lemma 4.4 [Lemma 3.4] Let $0 \leq \delta < 0.5$, and

$$f(x,z) \leq \frac{\delta^2}{2} - \frac{5\delta^3}{3(1-2\delta)},$$

then $x \in E(\omega, \delta)$.

Proof: Since $f(x,z)$ is a non-negative, strictly convex function, and $f(\omega, z) = 0$, its minimum value over the region $\{x \in R^n | x \in P_z, x \notin \operatorname{int} E(w,\delta)\}$ occurs on the boundary of $E(\omega,\delta)$. It is therefore sufficient to show that

$$f(x,z) \geq \frac{\delta^2}{2} - \frac{5\delta^3}{3(1-2\delta)}$$

for all the points on the boundary of $E(\omega,\delta)$. The Taylor's expansion of $f(x,z)$ at ω and Lemma 4.3 gives,

$$\begin{aligned} f(x,z) &\geq f(\omega,z) + \nabla f(\omega,z)^T(x-\omega) + \frac{1}{2}(x-\omega)^T \nabla^2 f(\omega,z)(x-\omega) - \frac{5\delta^3}{3(1-2\delta)} \\ &\geq \frac{\delta^2}{2} - \frac{5\delta^3}{3(1-2\delta)}. \end{aligned}$$

The last inequality follows by using the facts that $f(\omega,x) = 0$, $\nabla f(\omega,x) = 0$ and x is on the boundary of $E(\omega,\delta)$.

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Lemma 4.5 Let $x \in E(\omega,\delta)$, and let $0 < \delta < 0.243$. Then

$$|f(x,z)| \leq 1 \tag{4.5.a}$$

$$\nabla f(x,z)^T(x-\omega) \geq f(x,z) \geq 0 \tag{4.5.b}$$

$$\nabla f(x,z)^T(x-\omega) \geq (1 - \frac{7}{2}\delta - \frac{5}{2}\delta^2)(x-\omega)^T \nabla^2 f(x,z)(x-\omega) \tag{4.5.c}$$

$$\nabla f(x,z)^T(x-w) \geq \left[(1 - \frac{7}{2}\delta - \frac{5}{2}\delta^2) f(x,z)(x-\omega)^T \nabla^2 f(x,z)(x-\omega)\right]^{\frac{1}{2}} \tag{4.5.d}$$

Proof of (4.5.a). Since $f(\omega, z) = 0$ and $\nabla f(\omega, z) = 0$, from Lemma 4.1 and Lemma 4.3 we have

$$\begin{aligned}
|f(x,z)| \leq & \frac{1}{2}\left|(x-\omega)^T \nabla^2 f(\omega,z)(x-\omega)\right| \\
& + \Big| \sum_{j=3}^{\infty} \frac{1}{j} \sum_{i=1}^{m} (C_i(\omega))^j + m \sum_{j=3}^{\infty} \frac{1}{j!} \sum_{i=0}^{\lfloor j/2 \rfloor} \alpha_{ji} A^{j-2i}(\omega) B^i(\omega) \Big| \\
\leq & \frac{1}{2}\delta^2 + \frac{5\delta^3}{3(1-2\delta)} \qquad \text{(by Lemma 4.3)} \\
= & \frac{3+4\delta}{6(1-2\delta)}\delta^2 < 1.
\end{aligned}$$

($\mathcal{Q}.\mathcal{E}.\mathcal{D}.$)

Proof of (4.5.b). Since $f(\omega, z) = 0$ and $f(x, z)$ is convex, we have

$$0 \leq f(x,z) = f(x,z) - f(\omega,z) \leq \nabla f(x,z)^T (x-\omega).$$

($\mathcal{Q}.\mathcal{E}.\mathcal{D}.$)

Proof of (4.5.c). The following three relationships together with (2.6) and (2.7) are frequently used in our proof of (4.5.c). Let $D_i(x) = \frac{a_i^T(x-\omega)}{a_i^T\omega - b_i}$, then $D_i(x) \geq -1$ for all i,

$$\frac{D_i(x)}{1+D_i(x)} = \frac{a_i^T(x-\omega)}{a_i^T x - b_i}, \tag{4.3}$$

and

$$q(x) - q(\omega) = \nabla q(\omega)^T (x-\omega) + \frac{1}{2}(x-\omega)^T Q (x-\omega), \qquad (4.4).$$

Now,

$$\nabla f(x,z)^T(x-\omega) = -\sum_{i=1}^m \frac{a_i^T(x-\omega)}{a_i^Tx-b_i} - \frac{m\nabla q(x)^T(x-\omega)}{q(x)-z}$$

$$= -\sum_{i=1}^m \frac{D_i(x)}{1+D_i(x)} - \frac{m\nabla q(x)^T(x-\omega)}{q(x)-z}$$

$$= \sum_{i=1}^m \frac{D_i^2(x)}{1+D_i(x)} - \frac{m\nabla q(x)^T(x-\omega)}{q(x)-z} + \frac{m\nabla q(\omega)^T(x-\omega)}{q(\omega)-z} \quad \text{(by using } \nabla f(\omega,z)(x-\omega)=0\text{)}$$

$$= \sum_{i=1}^m \frac{D_i^2(x)}{1+D_i(x)} - \frac{m}{q(x)-z}[\nabla q(\omega)^T(x-\omega)+(x-\omega)^TQ(x-\omega)] + \frac{m\nabla q(\omega)^T(x-\omega)}{q(\omega)-z}$$

$$= \sum_{i=1}^m \frac{D_i^2(x)}{1+D_i(x)} - \frac{m(x-\omega)^TQ(x-\omega)}{q(x)-z} + m\frac{[q(x)-q(\omega)][\nabla q(\omega)(x-\omega)]}{[q(x)-z][q(\omega)-z]}.$$

Thus,

$$
\begin{aligned}
&\frac{1}{m}\nabla f(x,z)^T(x-\omega)\\
=&\frac{1}{m}\sum_{i=1}^{m}\Big(\frac{D_i(x)}{1+D_i(x)}\Big)^2(1+D_i(x))-\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}\\
&\quad+\frac{q(x)-z}{q(\omega)-z}\Big(1+\frac{1}{2}\frac{(x-\omega)^TQ(x-\omega)}{\nabla q(\omega)^T(x-\omega)}\Big)\Big(\frac{\nabla q(\omega)^T(x-\omega)}{q(x)-z}\Big)^2 \qquad \text{(by (4.4))}\\
=&\frac{1}{m}\sum_{i=1}^{m}\Big(\frac{D_i(x)}{1+D_i(x)}\Big)^2(1+D_i(x))-\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}\\
&+\frac{q(x)-z}{q(\omega)-z}\Big(\frac{\nabla q(\omega)^T(x-\omega)}{q(x)-z}\Big)^2+\frac{1}{2}\frac{\nabla q(\omega)^T(x-\omega)}{q(\omega)-z}\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}\\
\geq&\frac{1-\delta}{m}\sum_{i=1}^{m}\Big(\frac{D_i(x)}{1+D_i(x)}\Big)^2-\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}\\
&+\frac{q(x)-z}{q(\omega)-z}\Big(\frac{\nabla q(x)^T(x-\omega)-(x-\omega)^TQ(x-\omega)}{q(x)-z}\Big)^2\\
&+\frac{\delta}{2\sqrt{m}}\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z} \qquad \text{(because } x\in E(\omega,\delta))\\
\geq&(1-\delta-\frac{\delta^2}{2})\Big[\frac{1}{m}\sum_{i=1}^{m}\Big(\frac{D_i(x)}{1+D_i(x)}\Big)^2-\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}+\Big(\frac{\nabla q(x)^T(x-\omega)}{q(x)-z}\Big)^2\Big]\\
&+\Big(\frac{-2\nabla q(x)^T(x-\omega)}{q(\omega)-z}+\frac{\delta}{2\sqrt{m}}\Big)\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z} \qquad \text{(by Lemma 2.4)}\\
=&\frac{1}{m}(1-\delta-\frac{\delta^2}{2})(x-\omega)^T\nabla^2 f(x,z)(x-\omega)\\
&\quad+\Big[2\Big(\frac{\nabla q(\omega)^T(\omega-x)}{q(\omega)-z}-\frac{(x-\omega)^TQ(x-\omega)}{q(\omega)-z}\Big)+\frac{\delta}{2\sqrt{m}}\Big]\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}\\
\geq&\Big[\frac{1}{m}(1-\delta-\frac{\delta^2}{2})-\frac{2}{m}\Big(\frac{\delta}{\sqrt{m}}+\frac{\delta^2}{m}+\frac{\delta}{4\sqrt{m}}\Big)\Big](x-\omega)^T\nabla^2 f(x,z)(x-\omega)\\
&\text{(because }\Big|\frac{(x-\omega)^TQ(x-\omega)}{q(x)-z}\Big|\leq\frac{1}{m}(x-\omega)^T\nabla^2 f(x,z)(x-\omega))\\
\geq&\frac{1}{m}(1-\frac{7\delta}{2}-\frac{5\delta^2}{2})(x-\omega)^T\nabla^2 f(x,z)(x-\omega).
\end{aligned}
$$

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Proof of (4.5.d). For $0\leq\delta<0.243$, we have $1-\frac{7\delta}{2}-\frac{5\delta^2}{2}>0$. Thus multiplying (4.5.b) with (4.5.c), we get (4.5.d).

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Lemma 4.6 Let $z^+ \geq z$ and let $f(x, z^+) = F(\omega^+, z^+) - F(x, z^+)$, where

$$F(x, z^+) = \sum_{i=1}^{m} \ln(a_i^T x - b_i) + m \ln(q(x) - z^+),$$

and ω^+ is the point that maximizes $F(x, z^+)$ over the convex set

$$P_{z^+} = \{x \in R^n | a_i^T x \geq b_i \ \forall i \text{ and } q(x) \geq z^+\}.$$

Then

$$0 \leq q(\omega^+) - q(\omega) \leq z^+ - z.$$

Proof. Let $z(t) = z + t(z^+ - z)$ and let $\omega(t)$ be the point that maximizes the function

$$F(x, z(t)) = \sum_{i=1}^{m} \ln(a_i^T x - b_i) + m \ln(q(x) - z(t)).$$

Since the gradient of $F(x, z(t))$ vanishes at $\omega(t)$, we have

$$\begin{aligned}
0 &= \frac{d}{dt}\Big[\sum_{i=1}^{m} \frac{a_i}{a_i^T \omega(t) - b_i} + m \frac{\nabla q(\omega(t))}{q(\omega(t)) - z(t)}\Big] \\
&= \sum_{i=1}^{m} \frac{-a_i a_i^T}{(a_i^T \omega(t) - b_i)^2} \frac{d\omega(t)}{dt} \\
&\quad + \frac{m}{(q(\omega(t)) - z(t))^2}\Big[\frac{d\nabla q(\omega(t))}{d\omega(t)} \frac{d\omega(t)}{dt}(q(\omega(t)) - z(t)) - \nabla q(\omega(t))\Big(\frac{dq(\omega(t))}{dt} - \frac{dz(t)}{dt}\Big)\Big] \\
&= \sum_{i=1}^{m} \frac{-a_i a_i^T}{(a_i^T \omega(t) - b_i)^2} \frac{d\omega(t)}{dt} \\
&\quad + \frac{m}{(q(\omega(t)) - z(t))^2}\Big[(q(\omega(t)) - z(t)) Q \frac{d\omega(t)}{dt} - \nabla q(\omega(t))\Big(\frac{dq(\omega(t))}{dt} - \frac{dz(t)}{dt}\Big)\Big].
\end{aligned}$$

Because $-a_i a_i^T$ and Q are negative semidefinite, and $q(\omega(t)) - z(t) > 0$, we have

$$0 \geq \Big(\frac{d\,\omega(t)}{d\,t}\Big)^T \nabla q(\omega(t)) \Big(\frac{d\,q(\omega(t))}{d\,t} - \frac{d\,z(t)}{d\,t}\Big) = \frac{d\,q(\omega(t))}{d\,t}\Big(\frac{d\,q(\omega(t))}{d\,t} - \frac{d\,z(t)}{d\,t}\Big),$$

so

$$\Big(\frac{d\,q(\omega(t))}{d\,t}\Big)^2 \leq \frac{d\,q(\omega(t))}{d\,t}\frac{d\,z(t)}{d\,t} = \frac{d\,q(\omega(t))}{d\,t}(z^+ - z).$$

This implies that

$$0 \leq \frac{d\,q(\omega(t))}{d\,t} \leq z^+ - z.$$

Hence,

$$0 \leq q(\omega^+) - q(\omega) = \int_0^1 \frac{d\,q(\omega(t))}{d\,t} d\,t \leq z^+ - z.$$

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Lemma 4.7 [Lemma 3.5] Let $x \in E(\omega, \delta), 0 \leq \delta < .243$ and ϵ be a parameter such that $0 \leq \epsilon < 0.5$. The point x^+ that minimizes $\nabla f(x,z)^T y$ over $E(x,\epsilon) \equiv \{y \mid (y-x)^T \nabla^2 f(x,z)(y-x) \leq \epsilon^2\}$ satisfies,

$$f(x^+, z) \leq f(x,z) - \epsilon\sqrt{f(x,z)\Big(1 - \frac{7\delta}{2} - \frac{5\delta^2}{2}\Big)} + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)}.$$

Proof: The Taylor's expansion of $f(x^+, z)$ at x and Lemma 4.3 gives,

$$\begin{aligned} f(x^+, z) &\leq f(x,z) + \nabla f(x,z)^T(x^+ - x) + \frac{1}{2}(x^+ - x)^T \nabla^2 f(x,z)(x^+ - x) + \frac{5\epsilon^3}{3(1-2\epsilon)} \\ &\leq f(x,z) + \nabla f(x,z)^T(x^+ - x) + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)}. \end{aligned} \tag{4.5}$$

Let $\bar{x}$ be the point where the straight line joining x to the analytic center ω intersects with the boundary of the ellipsoid $E(x,\epsilon)$. Since $\nabla f(x,z)^T x^+ \leq \nabla f(x,z)^T \bar{x}$, from

(4.5) we have

$$
\begin{aligned}
f(x^+, z) &\leq f(x,z) + \nabla f(x,z)^T(\bar{x} - x) + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)} \\
&= f(x,z) + \frac{\epsilon \nabla f(x,z)^T(\omega - x)}{\sqrt{(\omega - x)^T \nabla^2 f(x,z)(\omega - x)}} + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)} \\
&\leq f(x,z) - \epsilon\sqrt{f(x,z)\Big(1 - \frac{7\delta}{2} - \frac{5\delta^2}{2}\Big)} + \frac{\epsilon^2}{2} + \frac{5\epsilon^3}{3(1-2\epsilon)}.
\end{aligned}
$$

The last inequality follows by using (4.5.d).

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

Lemma 4.8 [Lemma 3.6] Let $z^+ = z + \frac{\alpha}{\sqrt{m}}(q(x^+) - z)$, $\sqrt{m} > \alpha > 0$ and $x^+ \in \operatorname{int} P_{z^+}$. If

$$
f(x^+, z) \leq \frac{{\delta^+}^2}{2} - \frac{5{\delta^+}^3}{3(1-2\delta^+)},
$$

where $0 \leq \delta^+ < 0.5$, then we have

$$
f(x^+, z^+) \leq f(x^+, z) + \frac{m\alpha}{\sqrt{m} - \alpha}\Big(\frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m}\Big) + \frac{\alpha^2(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})^2}{1 - \frac{\alpha}{\sqrt{m}}(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})}. \tag{4.6}
$$

Proof: We may write

$$
f(x^+, z^+) = f(x^+, z) + f(\omega, z^+) + m \ln \frac{(q(x^+) - z)(q(\omega) - z^+)}{(q(x^+) - z^+)(q(\omega) - z)}. \tag{4.7}
$$

Since,

$$
\frac{(q(x^+) - z)(q(\omega) - z^+)}{(q(x^+) - z^+)(q(\omega) - z)} = 1 + \frac{(z^+ - z)(q(\omega) - q(x^+))}{(q(x^+) - z^+)(q(\omega) - z)}
$$

$$
= 1 + \frac{\alpha(q(x^+) - z)(q(\omega) - q(x^+))}{\sqrt{m}(q(x^+) - z^+)(q(\omega) - z)} = 1 + \frac{\alpha}{\sqrt{m} - \alpha}\frac{q(\omega) - q(x^+)}{q(\omega) - z},
$$

by using (2.4) we have,

$$\frac{(q(x^+)-z)(q(\omega)-z^+)}{(q(x^+)-z^+)(q(\omega)-z)} \leq 1+\frac{\alpha}{\sqrt{m}-\alpha}\Big(\frac{\delta^+}{\sqrt{m}}+\frac{{\delta^+}^2}{2m}\Big). \tag{4.8}$$

We need the following to upper bound the value of $f(\omega, z^+)$.

$$\frac{q(\omega^+)-z^+}{q(\omega)-z^+} = \frac{q(\omega^+)-z}{q(\omega)-z}\Big(1+\frac{(z^+-z)(q(\omega^+)-q(\omega))}{(q(\omega^+)-z)(q(\omega)-z^+)}\Big).$$

Now by using Lemma 4.6 in the above equation we have

$$\begin{aligned}
&\frac{q(\omega^+)-z^+}{q(\omega)-z^+} \\
\leq&\frac{q(\omega^+)-z}{q(\omega)-z}+\frac{(z^+-z)^2}{(q(\omega)-z)(q(\omega)-z^+)} \\
=&\frac{q(\omega^+)-z}{q(\omega)-z}+\frac{\alpha^2(q(x^+)-z)^2}{m(q(\omega)-z)(q(\omega)-z^+)} \\
=&\frac{q(\omega^+)-z}{q(\omega)-z}+\frac{\alpha^2}{m}\Big(\frac{q(x^+)-q(\omega)}{q(\omega)-z}+1\Big)\Big(\frac{q(x^+)-z}{q(\omega)-z-\frac{\alpha}{\sqrt{m}}(q(x^+)-z)}\Big) \\
=&\frac{q(\omega^+)-z}{q(\omega)-z}+\frac{\alpha^2}{m}\Big(\frac{q(x)-q(\omega)}{q(\omega)-z}+1\Big)^2\Big(\frac{1}{1-\frac{\alpha(q(x^+)-z)}{\sqrt{m}(q(\omega)-z)}}\Big) \\
\leq&\frac{q(\omega^+)-z}{q(\omega)-z}+\frac{\frac{\alpha^2}{m}(1+\frac{\delta^+}{\sqrt{m}}+\frac{{\delta^+}^2}{2m})^2}{1-\frac{\alpha}{\sqrt{m}}(1+\frac{\delta^+}{\sqrt{m}}+\frac{{\delta^+}^2}{2m})}.
\end{aligned} \tag{4.9}$$

The inequality (4.10) below follows by using (4.9), and (4.11) follows from the fact that $\nabla f(\omega, z) = 0$. Now,

$$\begin{aligned}
f(\omega, z^+) &= \sum_{i=1}^m \ln \frac{a_i^T\omega^+ - b_i}{a_i^T\omega - b_i} + m \ln \frac{q(\omega^+) - z^+}{q(\omega) - z^+} \\
&\le \sum_{i=1}^m \Big(\frac{a_i^T\omega^+ - b_i}{a_i^T\omega - b_i} - 1\Big) + m\Big(\frac{q(\omega^+) - z^+}{q(\omega) - z^+} - 1\Big) \\
&\le \sum_{i=1}^m \frac{a_i^T\omega^+ - a_i^T\omega}{a_i^T\omega - b_i} + m\Big[\frac{q(\omega^+) - z}{q(\omega) - z} + \frac{\frac{\alpha^2}{m}(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})^2}{1 - \frac{\alpha}{\sqrt{m}}(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})} - 1\Big] \qquad (4.10) \\
&= \frac{m}{2}\frac{(\omega^+ - \omega)^T Q(\omega^+ - \omega)}{q(\omega) - z} + \frac{\alpha^2(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})^2}{1 - \frac{\alpha}{\sqrt{m}}(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})} \qquad (4.11) \\
&\le \frac{\alpha^2(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})^2}{1 - \frac{\alpha}{\sqrt{m}}(1 + \frac{\delta^+}{\sqrt{m}} + \frac{{\delta^+}^2}{2m})}. \qquad (4.12)
\end{aligned}$$

The proof of Lemma 4.8 is complete by combining (4.7), (4.8) and (4.12).

$(\mathcal{Q}.\mathcal{E}.\mathcal{D}.)$

5. Initial Problem Transformation

In this section we discuss an approach to cast any given convex quadratic program into the problem form that is desired in Algorithm 3.2. Let us consider the problem

$$(\overline{QP}) \begin{cases} \text{maximize} & x^T\bar{Q}x + \bar{c}^T x \\ \text{subject to} & \bar{a}_i^T x \ge \bar{b}_i, \ i = 1, 2, \ldots, \bar{m}, \\ & x \ge 0, \end{cases}$$

where $\bar{a}_i \in Z^{\bar{n}}$, $c \in Z^{\bar{n}}$ and $\bar{Q} \in Z^{\bar{n}\times\bar{n}}$ is a negative semidefinite matrix. Let $\bar{L}$ represent the input length of $(\overline{QP})$. The problem form $(\overline{QP})$ is general enough, since in any given convex quadratic program the free variables can be replaced by two non-negative variables.

The unboundedness of $(\overline{QP})$ may be verified by checking feasibility of the system of linear equations corresponding to the dual of $(\overline{QP})$(see e.g., [15]). A feasible point for $(\overline{QP})$ and its dual can be obtained by using the algorithms of Gonzaga [8] or Vaidya [25]. Therefore, without loss of generality we assume that a feasible solution $\hat{x}$ and an upper bound $\bar{z}$ of $(\overline{QP})$ are available. Given a feasible point $\hat{x}$ it is easy to find $(\hat{x}, \hat{x}_a)$ which is an interior feasible solution to the following problem:

$$(\overline{QP}_1)\begin{cases}\text{maximize} & x^T\bar{Q}x + \bar{c}^T x - M_a x_a \\ \text{subject to} & \bar{a}_i^T x + x_a \geq b_i,\ i = 1, 2, \ldots, \bar{m}. \\ & x \geq 0,\ x_a \geq 0,\end{cases}$$

where x_a is an artificial variable and M_a is a cost that is large enough so that $x_a = 0$ for all the optimal solutions of $(\overline{QP}_1)$. $M_a = 2^{O(\bar{L})}$ is sufficient. The fact that at least one optimal solution $y^* = (x, x_a)^*$ of $(\overline{QP}_1)$ must be a constrained minimizer and hence $\|(x, x_a)^*\| \leq 2^{O(\bar{L})}$ may be used to generate a large number M_1 such that at least one optimal solution for $(\overline{QP}_1)$ satisfies $x_1 + \cdots + x_{\bar{n}} + x_a \leq M_1$. The modified problem

$$(\overline{QP}_2)\begin{cases}\text{maximize} & x^T\bar{Q}x + \bar{c}^T x - M_a x_a \\ \text{subject to} & \bar{a}_i^T x + x_a \geq b_i,\ i = 1, 2, \ldots, \bar{m}. \\ & -x_1 - \cdots - x_{\bar{n}} - x_a \geq -M_1, \\ & x \geq 0,\ x_a \geq 0,\end{cases}$$

has the following form:

$$(QP)\begin{cases}\text{maximize} & x^TQx + c^T x \\ \text{subject to} & P \equiv \{x \mid a_i^T x \geq b_i,\ i = 1, 2, \ldots, m\},\end{cases}$$

where P is bounded and a point $\hat{x}_0 \in \operatorname{int} P$ is known. The input length L of (QP) is $O(\bar{L})$. We now discuss how to find a point $x_0 \in \operatorname{int}(P)$ and a lower bound z_0 on the objective value of (QP) such that $f(x_0, z_0) \leq .003$.

The analytic center ξ of a bounded polytope P with a nonempty interior is the point that solves the problem

$$\begin{aligned} &\text{maximize} \quad g(x) \equiv \sum_{i=1}^{m} \ln(a_i^T x - b_i) \\ &\text{subject to} \quad x \in \operatorname{int} P. \end{aligned}$$

In [26], Vaidya developed an algorithm that efficiently calculates the analytic center ξ starting from any interior feasible solution $\hat{x}_0$ of P. If $g(\xi) - g(\hat{x}_0) \le K$ then Vaidya's algorithm would find a point x_0 such that $g(\xi) - g(x_0) \le .001$ in $O(n^{2.5}K)$ arithmetic operations. We note that the initial point $\hat{x}_0$ constructed as above satisfies $g(\xi) - g(\hat{x}_0) \le O(mL)$. Since $\sum_{i=1}^{m} \ln(a_i^T x - b_i)$ is maximized at ξ we have,

$$\begin{aligned} f(x_0, z_0) &= m \ln(q(\omega) - z_0) + \sum_{i=1}^{m} \ln(a_i^T \omega - b_i) - m \ln(q(x_0) - z_0) - \sum_{i=1}^{m} \ln(a_i^T x_0 - b_i) \\ &\le .001 + m \ln(q(\omega) - z_0) - m \ln(q(x_0) - z_0) \\ &\le .001 + m \ln(\bar{z} - z_0) - m \ln(q(x_0) - z_0) \end{aligned}$$

We can now easily find a z_0 such that $f(x_0, z_0) \le .003$ and $\bar{z} - z_0 = 2^{O(L)}$.

References

[1] Bayer, D. A. and J. C. Lagarias (1986). The nonlinear geometry of linear programming. Preprints, AT&T Bell Laboratories, Murray Hills, N.J.

[2] Fiacco, A. V. and G. P. McCormick (1968). *Nonlinear Programming: Sequential Unconstrained minimization techniques*, John Wiley, New York.

[3] Frisch, K.R. (1955). The logarithmic potential method of convex programming. Memorandum of May 13, 1955, University Institute of Economics, Oslo, Norway.

[4] Gill, P.E., W. Murray, M.A. Saunders, J.A. Tomlin and M.H. Wright (1986). On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method. *Mathematical Programming*, 36, 183–209.

[5] Gill, P.E., Murray, W., and Wright, M.H. (1981). *Practical Optimization*, Academic Press, New York.

[6] Goldfarb, D. and A. Idnani, (1983). A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming 27*, 1-33.

[7] Goldfarb, D. and S. Liu (1988). An $O(n^3L)$ Primal Interior Point Algorithm for Convex Quadratic Programming. Technical Report, Dept. of IEOR, Columbia University, New York, NY 10027.

[8] Gonzaga, C.C. (1987). An Algorithm for Solving Linear Programming Problems in $O(n^3L)$ Operations. Tech. Report, Dept. of Electrical Engineering and Computer Sciences, Berkeley, California, 94720.

[9] Huard, P. (1967) Resolution of Mathematical Programming with Nonlinear Constraints by the Method of Centers. *Nonlinear Programming*, (Editor J. Abadie), 207-219, 1967, North-Holland, Amsterdam.

[10] Iudin, D.B., A.S. Nemirovskii. (1976). Informational Complexity and Effective Methods of Solution for Convex Extremal Problems. *Ekonomika i Matematicheskie Metody* 12, 128-142, (translated in Matekon: Translations of Russian and East European Math. Economics 13, 25-45, Sprint 1977)

[11] Kapoor, S. and Vaidya, P.M. (1985). Fast Algorithms for Convex Quadratic Programming and Multicommodity Flows. Proc. 18th Annual ACM Symposium on Theory of Computing, 147-159.

[12] Karmarkar, N. (1984). A New Polynomial-Time Algorithm for Linear Programming. *Combinatorica* 4, 373-395, 1984.

[13] Kojima, M., S., S. Mizuno and A. Yoshise (1987). A Primal-Dual Interior Point Algorithm for Linear Programming. Report No. B-188, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan.

[14] Kojima, M., S., S. Mizuno and A. Yoshise (1987). A Polynomial-Time Algorithm for a Class of Linear Complementarity Problems. Report No. B-193, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan.

[15] Kozlov, M.K., Tarasov, S.P. and Khachian, L.G. (1979). Polynomial Solvability of Convex Quadratic Programming. *Dokl. Akad. Nauk SSSR, 5*, 1051-1053.

[16] Meggido, N. (1986). Pathways to the Optimal Set in Linear Programming, Manuscript, IBM Almaden Research Center, San Jose, CA 95120.

[17] Mehrotra, S. and J. Sun (1988). A method of analytic centers for quadratically constrained convex quadratic programs. Technical Report 88-01, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208.

[18] Mehrotra, S. and J. Sun (1988). An algorithm for solving smooth convex programs based on Newton's Method. Technical Report 88-08, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208.

[19] Mehrotra, S. and J. Sun (1988). On finding the center of a convex quadratic program. Report 88-11 (in preparation), Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208.

[20] Monteiro, R. C. and I. Adler (1987). An $O(n^3)$ Primal-Dual Interior Point Algorithm for Linear Programming. Report ORC 87-4, Operations Research Center, Dept. of Operations Research, University of California, Berkeley, CA.

[21] Monteiro, R. C. and I. Adler (1987). An $O(n^3)$ Primal-Dual Interior Point

Algorithm for Convex Quadratic Programming. Report ORC 87-15, Operations Research Center, Dept. of Operations Research, University of California, Berkeley, CA.

[22] Renegar, J. (1985). A Polynomial-Time Algorithm Based on Newton's Method for Linear Programming. Working Paper, Department of Mathematics, Colorado State University.

[23] Shor, N.Z. (1977). Cut-off Method with Space Extension in Convex Programming Problems. *Kibernetika* 13(1), 94-95. (translated in Cybernetics 13(1) 94-96, 1977).

[24] Sonnevend, Gy. (1985). A New Method for Solving A Set of Linear (Convex) Inequalities and its Applications for Identification and Optimization. Working Paper, Department of Numerical Analysis, Institute of Mathematics, Eotvos University, 1088, Budapest, Muzeum Korut 6-8.

[25] Vaidya, P.M. (1987). An Algorithm for Solving Linear Programming which Requires $O(((m+n)n^2 + (m+n)^{1.5}n)L)$ Arithmetic Operations. Preprint, AT&T Bell Laboratories, Murray Hills, New Jersey 07974.

[26] Vaidya, P.M. (1987). A Locally Well-Behaved Potential Function and a Simple Newton-type Method for Finding the Center of a Polytope. Preprint, AT&T Bell Laboratories, Murray Hills, New Jersey 07974.

[27] Ye, Y. (1987). Further Development on the Interior Algorithm For Convex Quadratic Programming, Manuscript, Dept. of Engineering- Economic Systems, Stanford University, Stanford, CA.