J. of Inequal. & Appl., 1998, Vol. 2, pp. 181–194
Reprints available directly from the publisher
Photocopying permitted by license only

A Conjugate Direction Method for Approximating the Analytic Center of a Polytope*

MASAKAZU KOJIMA [a,†], NIMROD MEGIDDO [b]
and SHINJI MIZUNO [c,‡]

[a] *Departments of Mathematical and Computing Science, Tokyo Institute of Technology, Meguro-ku, Tokyo 152, Japan;* [b] *IBM Almaden Research Center, 650 Harry Road, San Jose, California 95120-6099, and School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel;* [c] *The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan*

(Received 21 November 1996)

The analytic center ω of an n-dimensional polytope $P = \{x \in R^n: a_i^T x - b_i \geq 0\ (i = 1, 2, \ldots, m)\}$ with a nonempty interior P_{int} is defined as the unique minimizer of the logarithmic potential function $F(x) = \sum_{i=1}^{m} \log(a_i^T x - b_i)$ over P_{int}. It is shown that one cycle of a conjugate direction method, applied to the potential function at any $\nu \in P_{\text{int}}$ such that $\epsilon = \sqrt{(\nu - \omega)^T \nabla^2 F(\omega)(\nu - \omega)} \leq 1/6$, generates a point $\hat{x} \in P_{\text{int}}$ such that $\sqrt{(\hat{x} - \omega)^T \nabla^2 F(\omega)(\hat{x} - \omega)} \leq 23\sqrt{n}\epsilon^2$.

Keywords: Linear inequality; Polytope; Linear programming; Interior-point method; Conjugate Direction Method

AMS Subject Classification Numbers: 15A39 (Linear inequalities), 52A25 (Polyhedra and polytopes) and 90C05 (Linear programming)

---

*Part of this research was done when M. Kojima and S. Mizuno visited at the IBM Almaden Research Center in the summer of 1991. Partial support from the Office of Naval Research under Contract N00014-91-C-0026 is acknowledged.

† Corresponding author. Supported by Grant-in-Aids for Co-operative Research (03832017) of The Japan Ministry of Education, Science and Culture.

‡ Supported by Grant-in-Aids for Encouragement of Young Scientist (03740125) and Co-operative Research (03832017) of The Japan Ministry of Education, Science and Culture.

1. INTRODUCTION

Let P denote a polytope of the form $\{x \in R^n\colon a_i^T x \geq b_i \ (i = 1, 2, \ldots, m)\}$. The symbol P_{int} stands for its interior $\{x \in R^n\colon a_i^T x > b_i \ (i = 1, 2, \ldots, m)\}$. We assume throughout that P_{int} is nonempty and bounded. Let F denote the logarithmic potential function on P_{int};

$$F(x) = -\sum_{i=1}^{m} \log(a_i^T x - b_i) \quad \text{for every } x \in P_{\text{int}}.$$

The analytic center ω of the polytope P [11] is defined as the unique minimizer of the potential function F over the interior P_{int} of P. We denote the gradient vector and the Hessian matrix at each $x \in P_{\text{int}}$ by $\nabla F(x)$ and $\nabla^2 F(x)$, respectively. By a simple calculation, we see that

$$\nabla F(x) = -\sum_{i=1}^{m} \frac{a_i}{a_i^T x - b_i} \quad \text{and} \quad \nabla^2 F(x) = \sum_{i=1}^{m} \frac{a_i a_i^T}{(a_i^T x - b_i)^2}.$$

Assuming there exist n linearly independent a_i's, the Hessian matrix $\nabla^2 F(x)$ is positive definite at every $x \in P_{\text{int}}$. Hence, the potential function F is strictly convex over P_{int}, and if $P_{\text{int}} \neq \emptyset$, the analytic center ω of the polytope P is the unique solution of the system of equations $\nabla F(x) = \mathbf{0}$. Furthermore, at any fixed $x \in P_{\text{int}}$, we can define a norm over R^n by

$$\|\xi\|_x = \sqrt{\xi^T \nabla^2 F(x) \xi}. \tag{1}$$

We use the norm $\|x - \omega\|_\omega$ to measure the distance from any $x \in P_{\text{int}}$ to the analytic center ω as in the papers by Renegar [10] and Vaidya [15].

Consider the linear program

$$\begin{array}{ll} \text{Minimize} & c^T x \\ \text{subject to} & a_i^T x \geq b_i \quad (i = 1, 2, \ldots, m) \end{array}$$

Here, $x, c \in R^n$, $a_i \in R^n$, and $b_i \in R$. We assume that the linear program has a bounded nonempty set of optimal solutions. Denote by λ^* the optimal value of the objective function. It follows that for every $\lambda > \lambda^*$ the interior $S_{\text{int}}(\lambda)$ of the parametric level set

$$S(\lambda) = \{x \in R^n\colon c^T x \leq \lambda,\ a_i^T x \geq b_i \ (i = 1, 2, \ldots, m)\}$$

is nonempty and bounded. For each $\lambda > \lambda^*$, let ω_λ denote the analytic center of $S(\lambda)$. It is well known that the set $\{\omega_\lambda : \lambda > \lambda^*\}$, consisting of all the analytic centers, forms a smooth curve which runs through the interior of the feasible region and converges to an optimal solution of the linear program as λ tends λ^*. The curve is called the central trajectory or the path of centers. Thus, if we numerically trace the central trajectory till λ gets sufficiently close to λ^*, then we obtain an approximate optimal solution of the linear program [11]. Renegar [10] embodied this idea in his polynomial-time algorithm for linear programs using Newton's method for approximating the analytic center of a polytope. Since then, the approximation of the central trajectory by Newton method has played a major role in many interior point algorithms developed for linear programs (see *e.g.* [2,3,5,7,8,10, 13,14,16]) and their extensions to convex quadratic programs (see *e.g.* [9]) and linear complementarity problems (see *e.g.* [6]).

The following theorem provides a theoretical basis for the polynomial-time convergence of Renegar's algorithm:

THEOREM 1.1 (Theorem 3.2 of [10]) *If* $\|\boldsymbol{v} - \boldsymbol{\omega}\|_\omega = \epsilon < 1$, *then the point* $\boldsymbol{x}^* = \boldsymbol{v} - \nabla^2 F(\boldsymbol{v})^{-1} \nabla F(\boldsymbol{v})$, *generated by one Newton iteration for minimizing the potential function F at the point* $\boldsymbol{v}$, *satisfies*

$$\|\boldsymbol{x}^* - \boldsymbol{\omega}\|_\omega \leq \frac{(1+\epsilon)^2 \epsilon^2}{1-\epsilon}. \tag{2}$$

Vaidya [15] also investigated the behavior of a Newton-type method for approximating the analytic center. Our research is motivated by the following questions: (i) Do conjugate direction and conjugate gradient methods (see, *e.g.* [1,4,12]) work as effectively as Newton's method for approximating the analytic center? (ii) Can these methods be utilized effectively in interior point algorithms, replacing Newton's method? Here we begin answering these questions. Let $\boldsymbol{v} \in \operatorname{int} P$, and let $\boldsymbol{d}^1, \boldsymbol{d}^2, \ldots, \boldsymbol{d}^n$ be n $\nabla^2 F(\boldsymbol{v})$-orthonormal vectors;

$$\begin{aligned} \|\boldsymbol{d}^i\|_v^2 = (\boldsymbol{d}^i)^{\mathrm{T}} \nabla^2 F(\boldsymbol{v}) \boldsymbol{d}^i = 1 \quad (i = 1, 2, \ldots, n), \\ (\boldsymbol{d}^i)^{\mathrm{T}} \nabla^2 F(\boldsymbol{v}) \boldsymbol{d}^j = 0 \quad \text{if } 1 \leq i < j \leq n. \end{aligned} \tag{3}$$

We are concerned with a conjugate direction method for approximating the analytic center.

Conjugate Direction Method (One Cycle)

Step 0: Let $\boldsymbol{y}^0=\boldsymbol{v}$ and $k=1$.
Step 1: Define $\boldsymbol{y}^k \in P_{\text{int}}$ by $F(\boldsymbol{y}^k)=\min\{F(\boldsymbol{y}^{k-1}+\nu \boldsymbol{d}^k)\colon \nu \in R\}$.
Step 2: If $k=n$, then stop. Otherwise, set $k=k+1$ and go to Step 1.

Our main result is:

THEOREM 1.2 *If* $\epsilon = \|\boldsymbol{v}-\boldsymbol{\omega}\|_\omega \leq 1/6$, *then* $\|\boldsymbol{y}^n-\boldsymbol{\omega}\|_\omega \leq 23\sqrt{n}\epsilon^2$.

In the remainder of this note we prove this theorem. As by-products, we derive several interesting properties of the potential function and its quadratic approximation. In particular, we will see in Corollary 2.5 that an inequality slightly stronger than (2) holds under the assumption of Theorem 1.1.

2. BASIC ANALYSIS

2.1. Symbols and Notation

We begin by introducing the notation. Let

$$f(\boldsymbol{x}) = F(\boldsymbol{x}) - F(\boldsymbol{\omega}) \quad \text{for every } \boldsymbol{x} \in P_{\text{int}}.$$

Obviously, the gradient and the Hessian of the function f at every $\boldsymbol{x} \in P_{\text{int}}$ coincide with those of the potential function F, respectively. By the definition of the analytic center $\boldsymbol{\omega}$ of the polytope P, we have

$$f(\boldsymbol{\omega}) = 0 \quad \text{and} \quad f(\boldsymbol{x}) \geq 0 \quad \text{for every } \boldsymbol{x} \in P_{\text{int}}.$$

We also call f a potential function.

Define a quadratic approximation of the potential function f at every $\boldsymbol{y} \in P_{\text{int}}$ by

$$q_y(\boldsymbol{x}) = f(\boldsymbol{y}) + \nabla F(\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y}) \\ + \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\nabla^2 F(\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y}) \quad \text{for every } \boldsymbol{x} \in R^n. \tag{4}$$

In particular, we have

$$q_v(\boldsymbol{x}) = q_v(\boldsymbol{x}^*) + \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}\nabla^2 F(\boldsymbol{v})(\boldsymbol{x}-\boldsymbol{x}^*) \quad \text{for every } \boldsymbol{x} \in R^n. \tag{5}$$

Here, $\boldsymbol{x}^* = \boldsymbol{v} - (\nabla^2 F(\boldsymbol{v}))^{-1}\nabla F(\boldsymbol{v})$ denotes the point generated by one Newton iteration for minimizing the potential function f at a point $\boldsymbol{v} \in P_{\text{int}}$ (see Theorem 1.1).

For each $\boldsymbol{x} \in P_{\text{int}}$, define an $n \times m$ matrix $\boldsymbol{U}_x$ by

$$\boldsymbol{U}_x = \left(\frac{\boldsymbol{a}_1}{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{x} - b_1}, \frac{\boldsymbol{a}_2}{\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{x} - b_2}, \ldots, \frac{\boldsymbol{a}_m}{\boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{x} - b_m}\right),$$

and for each pair $\boldsymbol{x}, \boldsymbol{y} \in P_{\text{int}}$, define the $m \times m$ diagonal matrix $\boldsymbol{R}_x^y$ by

$$\boldsymbol{R}_x^y = \operatorname{diag}\left(\frac{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{y} - b_1}{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{x} - b_1}, \frac{\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{y} - b_2}{\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{x} - b_2}, \ldots, \frac{\boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{y} - b_m}{\boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{x} - b_m}\right).$$

With this notation, we can rewrite the norm $\|\boldsymbol{\xi}\|_x$ of $\boldsymbol{\xi} \in R^n$ defined by (1) as

$$\|\boldsymbol{\xi}\|_x = \sqrt{\boldsymbol{\xi}^{\mathrm{T}}\nabla^2 F(\boldsymbol{x})\boldsymbol{\xi}} = \|\boldsymbol{U}_x^{\mathrm{T}}\boldsymbol{\xi}\| \quad (\boldsymbol{x} \in P).$$

This equality is used very often without any reference. Also, it is easy to verify the following equalities for every pair $\{\boldsymbol{x}, \boldsymbol{y}\} \subset P_{\text{int}}$:

$$\nabla F(\boldsymbol{x}) = -\boldsymbol{U}_x \boldsymbol{e}, \tag{6}$$

$$\nabla^2 F(\boldsymbol{x}) = \boldsymbol{U}_x \boldsymbol{U}_x^{\mathrm{T}}, \tag{7}$$

$$\boldsymbol{U}_y = \boldsymbol{U}_x \boldsymbol{R}_y^x, \tag{8}$$

$$\boldsymbol{e} - \boldsymbol{R}_y^x \boldsymbol{e} = \boldsymbol{U}_y^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}). \tag{9}$$

2.2. Some Properties of the Quadratic Approximation

In this subsection we present three lemmas. The first one estimates the difference between the norms $\|\boldsymbol{\xi}\|_y$ and $\|\boldsymbol{\xi}\|_x$ of $\boldsymbol{\xi} \in P_{\text{int}}$ when $\boldsymbol{y}$ and $\boldsymbol{x}$ are close. The second and the third ones evaluate the errors in the gradient vector ∇q_y and the Hessian matrix $\nabla^2 q_y$ of the quadratic approximation q_y of the potential function f, respectively.

LEMMA 2.1 *If* $\boldsymbol{\xi} \in R^n$, $\{\boldsymbol{x}, \boldsymbol{y}\} \subset P_{\text{int}}$, *and* $\|\boldsymbol{x} - \boldsymbol{y}\|_y \leq \epsilon$, *then*

$$(1 - \epsilon)\|\boldsymbol{\xi}\|_x \leq \|\boldsymbol{\xi}\|_y \leq (1 + \epsilon)\|\boldsymbol{\xi}\|_x. \tag{10}$$

Proof For every i $(i=1,2,\ldots,m)$, let r_i denote the ith diagonal element $(\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x}-b_i)/(\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{y}-b_i)$ of the diagonal matrix $\boldsymbol{R}_y^x$. By (9) and the assumption,

$$\sum_{i=1}^{m}(1-r_i)^2 = \|\boldsymbol{e}-\boldsymbol{R}_y^x\boldsymbol{e}\|^2 = \|\boldsymbol{U}_y^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{x})\|^2 = \|\boldsymbol{x}-\boldsymbol{y}\|_y^2 \le \epsilon^2.$$

This implies that

$$\sum_{i=1}^{m}(1-r_i)^2 \le \epsilon^2 \quad \text{and} \quad 1-\epsilon \le r_i \le 1+\epsilon \quad (i=1,2,\ldots,m). \tag{11}$$

On the other hand, we know by (8) that $\boldsymbol{U}_y^{\mathrm{T}}\boldsymbol{\xi} = \boldsymbol{R}_y^x\,\boldsymbol{U}_x^{\mathrm{T}}\boldsymbol{\xi}$. Hence, taking the Euclidean norm of both sides of this equality, we obtain

$$(1-\epsilon)\|\boldsymbol{U}_x^{\mathrm{T}}\boldsymbol{\xi}\| \le \|\boldsymbol{U}_y^{\mathrm{T}}\boldsymbol{\xi}\| = \|\boldsymbol{R}_y^x\,\boldsymbol{U}_x^{\mathrm{T}}\boldsymbol{\xi}\| \le (1+\epsilon)\|\boldsymbol{U}_x^{\mathrm{T}}\boldsymbol{\xi}\|.$$

Thus, the desired inequalities in (10) follow.

LEMMA 2.2 *If* $\boldsymbol{\xi}\in R^n$, $\{\boldsymbol{x},\boldsymbol{y}\}\subset P_{\mathrm{int}}$, *and either* $\|\boldsymbol{x}-\boldsymbol{y}\|_y\le\epsilon$ *or* $\|\boldsymbol{x}-\boldsymbol{y}\|_x\le\epsilon$, *then*

$$|(\nabla q_y(\boldsymbol{x})-\nabla F(\boldsymbol{x}))^{\mathrm{T}}\boldsymbol{\xi}| \le \frac{\epsilon^2}{1-\epsilon}\|\boldsymbol{\xi}\|_y.$$

Proof As in the proof of Lemma 2.1, we denote the ith diagonal element $(\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x}-b_i)/(\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{y}-b_i)$ of the diagonal matrix $\boldsymbol{R}_y^x$ by r_i $(i=1,2,\ldots,m)$. If $\|\boldsymbol{x}-\boldsymbol{y}\|_y\le\epsilon$, then the inequalities in (11) hold. Since $R_x^y=(R_y^x)^{-1}$, we see by symmetry that if $\|\boldsymbol{x}-\boldsymbol{y}\|_x\le\epsilon$, then the inequalities

$$\sum_{i=1}^{m}\left(1-\frac{1}{r_i}\right)^2 \le \epsilon^2 \quad \text{and} \quad 1-\epsilon\le\frac{1}{r_i}\le 1+\epsilon \quad (i=1,2,\ldots,m) \quad (12).$$

hold. We can easily verify that

$$\begin{aligned}
&|(\nabla q_y(\boldsymbol{x})-\nabla F(\boldsymbol{x}))^{\mathrm{T}}\boldsymbol{\xi}| \\
&\quad = |(\nabla F(\boldsymbol{y})+\nabla^2F(\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})-\nabla F(\boldsymbol{x}))^{\mathrm{T}}\boldsymbol{\xi}| \quad \text{(by (4))}
\end{aligned}$$

$$
\begin{aligned}
&= \left| \left(-U_y e + U_y U_y^{\mathrm{T}}(x - y) + U_x e \right)^{\mathrm{T}} \xi \right| \quad \text{(by (6) and (7))} \\
&= \left| \left(-U_y e + U_y(-e + R_y^x e) + U_y R_x^y e \right)^{\mathrm{T}} \xi \right| \quad \text{(by (8) and (9))} \\
&= \left| \left(-2e + R_y^x e + R_x^y e \right)^{\mathrm{T}} U_y^{\mathrm{T}} \xi \right| \\
&\leq \| -2e + R_y^x e + R_x^y e \| \cdot \| U_y^{\mathrm{T}} \xi \| \\
&\leq \left(\sum_{i=1}^{m} \left| 2 - r_i - \frac{1}{r_i} \right| \right) \|\xi\|_y \\
&= \left(\sum_{i=1}^{m} (1 - r_i)^2 \frac{1}{r_i} \right) \|\xi\|_y = \left(\sum_{i=1}^{m} \left(1 - \frac{1}{r_i} \right)^2 r_i \right) \|\xi\|_y .
\end{aligned}
$$

In either case ((11) or (12)) the desired inequality follows.

LEMMA 2.3 *If* $\xi \in R^n$, $\{x, y\} \subset P_{\text{int}}$ *and* $\|x - y\|_y \leq \epsilon$, *then*

$$|\xi^{\mathrm{T}}(\nabla^2 q_y(x) - \nabla^2 F(x))\xi| \leq (2 + \epsilon)\epsilon \|\xi\|_y^2 .$$

Proof By the definition (4) of the quadratic function q_y and (7),

$$\xi^{\mathrm{T}} \nabla^2 q_y(x) \xi = \|U_y^{\mathrm{T}} \xi\|^2 = \|\xi\|_y^2 \quad \text{and} \quad \xi^{\mathrm{T}} \nabla^2 F(x) \xi = \|U_x^{\mathrm{T}} \xi\|^2 = \|\xi\|_x^2$$

By Lemma 2.1,

$$
\begin{aligned}
&|\xi^{\mathrm{T}}(\nabla^2 q_y(x) - \nabla^2 F(x))\xi| \\
&\quad = |\|\xi\|_y^2 - \|\xi\|_x^2| \\
&\quad \leq \max\{|(1 - \epsilon)^2 \|\xi\|_y^2 - \|\xi\|_y^2|,\ |(1 + \epsilon)^2 \|\xi\|_y^2 - \|\xi\|_y^2|\} \\
&\quad \leq \max\{(2 - \epsilon)\epsilon,\ (2 + \epsilon)\epsilon\} \|\xi\|_y^2 \\
&\quad = (2 + \epsilon)\epsilon \|\xi\|_y^2 .
\end{aligned}
$$

2.3. Minimization of the Quadratic Approximation

The following lemma shows a relation between the minimizer x^f of the potential function f and the minimizer x^q of its quadratic approximation q_y over any affine subspace S that intersects P_{int}.

LEMMA 2.4 *Suppose S is an affine subspace of R^n which intersects P_{int}, and let $y \in P_{\text{int}}$ be any point. Let x^f be the minimizer of the potential function f over $S \cap P_{\text{int}}$, and let x^q be the minimizer over S of the quadratic approximation q_y of f at $y \in P_{\text{int}}$. Under these conditions, if either $\|x^f - y\|_y \leq \epsilon$ or $\|x^f - y\|_{x^f} \leq \epsilon$, then $\|x^q - x^f\|_y \leq \epsilon^2/(1-\epsilon)$.*

Proof From the definition of x^f and x^q,

$$(x - z)^{\mathrm{T}} \nabla F(x^f) = 0 \quad \text{and} \quad (x - z)^{\mathrm{T}} \nabla q_y(x^q) = 0$$

for any x and z in S. In particular, these inequalities hold for $x = x^q$ and $z = x^f$. Hence,

$$(x^q - x^f)^{\mathrm{T}} \nabla F(x^f) = 0 \quad \text{and} \quad (x^q - x^f)^{\mathrm{T}} \nabla q_y(x^q) = 0. \tag{13}$$

Therefore,

$$\begin{aligned}
\|x^q - x^f\|_y^2 &= \|U_y^{\mathrm{T}}(x^q - x^f)\|^2 \\
&= (x^q - x^f)^{\mathrm{T}} U_y U_y^{\mathrm{T}}(x^q - x^f) \\
&= (x^q - x^f)^{\mathrm{T}} \nabla^2 F(y)(x^q - x^f) \quad \text{(by (7))} \\
&= (x^q - x^f)^{\mathrm{T}} (\nabla q_y(x^q) - \nabla q_y(x^f)) \quad \text{(see (4))} \\
&= (x^q - x^f)^{\mathrm{T}} (\nabla F(x^f) - \nabla q_y(x^f)) \quad \text{(by (13))} \\
&\leq \frac{\epsilon^2}{1-\epsilon} \|x^q - x^f\|_y \quad \text{(by Lemma 2.2).}
\end{aligned}$$

Using Lemma 2.4, we can strengthen the inequality (2) given in Theorem 1.1.

COROLLARY 2.5 *Under the assumptions of Theorem* 1.1, $\|x^* - \omega\|_\omega \leq (1+\epsilon)\epsilon^2/(1-\epsilon)$.

Proof If we apply Lemma 2.4 to the case where $S = R^n$ and $y = v$, then $x^f = \omega$ and $x^q = x^*$, and hence $\|x^* - \omega\|_v \leq \epsilon^2/(1-\epsilon)$. On the other hand, by Lemma 2.1, $\|x^* - \omega\|_\omega \leq (1+\epsilon)\|x^* - \omega\|_v$. Thus the desired inequality follows.

2.4. Neighborhoods of the Analytic Center

We define an ellipsoidal neighborhood $E(\epsilon)$ of the analytic center ω by

$$E(\epsilon) = \{z: \|z - \omega\|_\omega \leq \epsilon\} = \{\omega + s\xi: \|\xi\|_\omega = 1 \text{ and } 0 \leq s \leq \epsilon\},$$

and the level set $L(c)$ of the potential function f by

$$L(c) = \{x: f(x) \leq c\}$$

for each $c \geq 0$. The following lemma shows a relation between the neighborhood $E(\epsilon)$ of the analytic center ω of the polytope P and the level set $L(c)$:

LEMMA 2.6 For every $\epsilon \in [0, \frac{1}{6}]$,

$$E(\epsilon) \subset L(2\epsilon^2/3) \subset E(\sqrt{2}\epsilon).$$

Proof By Lemma 5.1 of [15],

$$\left| f(\omega + \delta\xi) - \tfrac{1}{2}\delta^2 \xi^{\mathrm{T}} \nabla^2 F(\omega)\xi \right| \leq \frac{\delta^3}{3(1-\delta)}$$

for every $\delta \in [0, 1)$ and every $\xi \in R^n$ with $\|\xi\|_\omega = 1$. Since $\xi^{\mathrm{T}} \nabla^2 F(\omega)\xi = \xi^{\mathrm{T}} U_\omega U_\omega^{\mathrm{T}} \xi$, it follows that

$$\left| f(\omega + \delta\xi) - \frac{\delta^2}{2} \right| \leq \frac{\delta^3}{3(1-\delta)} \tag{14}$$

for every $\delta \in [0, 1)$ and every $\xi \in R^n$ with $\|\xi\|_\omega = 1$. This inequality will be used later.

Now, if

$$x = \omega + s\xi \in E(\epsilon), \; 0 \leq s \leq \epsilon < 1 \quad \text{and} \quad \|\xi\|_\omega = 1,$$

then

$$\begin{aligned} f(x) &= f(\omega + s\xi) \\ &\leq \frac{s^2}{2} + \frac{s^3}{3(1-s)} \quad \text{(by (14))} \\ &\leq \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3(1-\epsilon)} \quad (\text{since } s \leq \epsilon < 1) \\ &\leq \frac{2\epsilon^2}{3}. \end{aligned}$$

This implies that $x \in L(2\epsilon^2/3)$. Thus, we have shown the first inclusion relation $E(\epsilon) \subset L(2\epsilon^2/3)$ of the lemma.

We now prove the second inclusion relation $L(2\epsilon^2/3) \subset E(\sqrt{2}\epsilon)$. It suffices to show that if

$$x = \omega + s\xi \notin E(\sqrt{2}\epsilon) \quad \text{and} \quad \|\xi\|_\omega = 1,$$

then

$$x = \omega + s\xi \notin L(2\epsilon^2/3).$$

Since $s > \sqrt{2}\epsilon$ and f is a strictly convex function attaining its minimum value 0 at the analytic center ω of the polytope P, we have

$$f(\omega + s\xi) > f(\omega + \sqrt{2}\epsilon\xi).$$

On the other hand, by (14) and $0 \le \epsilon \le 1/6$,

$$f(\omega + \sqrt{2}\epsilon\xi) \ge \frac{2\epsilon^2}{2} - \frac{(\sqrt{2}\epsilon)^3}{3(1-\sqrt{2}\epsilon)} \ge 2\epsilon^2/3.$$

Hence, $x \notin L(2\epsilon^2/3)$.

3. PROOF OF THE MAIN THEOREM

By applying Lemma 2.4 with $S = R^n$, $x^f = \omega$ and $x^q = x^*$, we first observe that

$$\|x^* - \omega\|_v \le \frac{\epsilon^2}{1-\epsilon}. \tag{15}$$

We also see by Lemma 2.6 that $v \in E(\epsilon) \subset L(2\epsilon^2/3)$, which implies $f(y^0) = f(v) \le 2\epsilon^2/3$. It follows from the construction of the sequence (see the Conjugate Direction Method described in Section 1) that

$$f(y^k) \le f(y^0) \le \frac{2\epsilon^2}{3}, \quad \text{i.e., } y^k \in L(2\epsilon^2/3) \quad (k = 0, 1, \ldots, n).$$

Hence, for every $k = 0, 1, \ldots, n$,

$$\begin{aligned}\|\boldsymbol{y}^k - \boldsymbol{v}\|_v &\le \frac{1}{1-\epsilon}\|\boldsymbol{y}^k - \boldsymbol{v}\|_\omega \quad \text{(by Lemma 2.1)}\\ &\le \frac{1}{1-\epsilon}\left(\|\boldsymbol{y}^k - \boldsymbol{\omega}\|_\omega + \|\boldsymbol{v} - \boldsymbol{\omega}\|_\omega\right)\\ &\le \frac{1}{1-\epsilon}(\sqrt{2}\epsilon + \epsilon)\\ &\qquad \text{(by } \boldsymbol{y}^k \in L(2\epsilon^2/3)\text{, Lemma 2.6 and } \boldsymbol{v} \in E(\epsilon))\\ &\le 3\epsilon \quad (\text{since } 0 < \epsilon \le 1/6).\end{aligned}$$

Thus, we have shown that

$$\|\boldsymbol{y}^k - \boldsymbol{v}\|_v \le 3\epsilon \quad (k = 0, 1, \ldots, n). \tag{16}$$

Recall that $\boldsymbol{d}^1, \boldsymbol{d}^2, \ldots, \boldsymbol{d}^n$ are n $\nabla^2 F(\boldsymbol{v})$-orthonormal vectors satisfying (3).

Lemma 3.1 *If $\boldsymbol{y} = \boldsymbol{x}^* + \sum_{i=1}^n \lambda_i \boldsymbol{d}^i$, then for every fixed j ($j = 1, \ldots, n$), the minimum of the quadratic approximation q_v of the potential function f at $\boldsymbol{v} \in P_{\text{int}}$ on the line $\{\boldsymbol{y} + \alpha \boldsymbol{d}^j : \alpha \in R\}$ is attained at $\boldsymbol{x}^* + \sum_{i \ne j} \lambda_i \boldsymbol{d}^i$.*

Proof By (5) and (3),

$$\begin{aligned}q_v(\boldsymbol{y} + \alpha \boldsymbol{d}^j) &= q_v(\boldsymbol{x}^*) + \frac{1}{2}\left(\sum_{i=1}^n \lambda_i \boldsymbol{d}^i + \alpha \boldsymbol{d}^j\right)^{\mathrm{T}} \nabla^2 F(\boldsymbol{v}) \left(\sum_{i=1}^n \lambda_i \boldsymbol{d}^i + \alpha \boldsymbol{d}^j\right)\\ &= q_v(\boldsymbol{x}^*) + \frac{1}{2}\left(\sum_{i \ne j}^n \lambda_i^2 + (\lambda_j + \alpha)^2\right).\end{aligned}$$

Thus, the minimum of $q_v(\boldsymbol{y} + \alpha \boldsymbol{d}^j)$ with respect to $\alpha \in R$ is attained at $\alpha = -\lambda_j$.

Let μ_i ($i = 1, 2, \ldots, n$) be real numbers such that

$$\boldsymbol{y}^0 = \boldsymbol{v} = \boldsymbol{x}^* + \sum_{i=1}^n \mu_i \boldsymbol{d}^i.$$

For each k ($k = 1, 2, \ldots, n$), let ν_k denote the step length from $\boldsymbol{y}^{k-1}$ taken at Step 1 of the Conjugate Direction Method along the direction $\boldsymbol{d}^k$;

$$\boldsymbol{y}^k = \boldsymbol{y}^{k-1} + \nu_k \boldsymbol{d}^k. \tag{17}$$

Then each y^k $(k = 1, 2, \ldots, n)$ can be rewritten as

$$y^k = y^0 + \sum_{i=1}^{k} \nu_i d^i = x^* + \sum_{i=1}^{k} (\mu_i + \nu_i) d^i + \sum_{i=k+1}^{n} \mu_i d^i.$$

For each k $(k = 1, 2, \ldots, n)$, let $\bar{y}^k$ be the minimizer of q_v on the line $\{y^{k-1} + \alpha d^k : \alpha \in R\}$. By Lemma 3.1,

$$\bar{y}^k = x^* + \sum_{i=1}^{k-1} (\mu_i + \nu_i) d^i + \sum_{i=k+1}^{n} \mu_i d^i = y^k - (\mu_k + \nu_k) d^k.$$

Hence,

$$\begin{aligned}
\|y^n - x^*\|_v^2 &= \Big\| \sum_{i=1}^{n} (\mu_i + \nu_i) d^i \Big\|_v^2 \\
&= \sum_{i=1}^{n} \|(\mu_i + \nu_i) d^i\|_v^2 \quad \text{(by (3))} \\
&= \sum_{i=1}^{n} \|\bar{y}^i - y^i\|_v^2 \\
&\le \sum_{i=1}^{n} \left(\frac{(3\epsilon)^2}{1 - 3\epsilon} \right)^2 \quad \text{(by (16) and Lemma 2.4)} \\
&\le n \left(\frac{9\epsilon^2}{1 - 3\epsilon} \right)^2.
\end{aligned}$$

Since $0 \le \epsilon \le 1/6$,

$$\|y^n - x^*\|_v \le \sqrt{n} \left(\frac{9}{1 - 3\epsilon} \right) \epsilon^2 \le 18\sqrt{n}\epsilon^2. \tag{18}$$

Consequently,

$$\begin{aligned}
\|y^n - \omega\|_\omega &\le (1 + \epsilon) \|y^n - \omega\|_v \quad \text{(by } \|v - \omega\|_\omega \le \epsilon \text{ and Lemma 2.1)} \\
&\le (1 + \epsilon) \big(\|y^n - x^*\|_v + \|x^* - \omega\|_v \big) \\
&\le (1 + \epsilon) \left(18\sqrt{n}\epsilon^2 + \frac{\epsilon^2}{1 - \epsilon} \right) \quad \text{(by (15) and (18))} \\
&\le \frac{7}{6} \left(18\sqrt{n} + \frac{6}{5} \right) \epsilon^2 \quad \text{(since } 0 \le \epsilon \le 1/6) \\
&\le 23\sqrt{n}\epsilon^2.
\end{aligned}$$

4. CONCLUDING REMARKS

Our main theorem (Theorem 1.2) suggests a modification of Renegar's polynomial-time algorithm [10] for linear programs. Specifically, we can replace Newton's method, which is used to trace the central trajectory in Renegar's algorithm, by the Conjugate Direction Method. Such a modification does not, however, seem so attractive. The authors are not satisfied with the fact that in the Conjugate Direction Method the search directions $\boldsymbol{d}^1, \boldsymbol{d}^2, \ldots, \boldsymbol{d}^n$ are chosen as conjugate directions with respect to a fixed Hessian matrix $\nabla^2 F(\boldsymbol{v})$ of the potential function F at $\boldsymbol{v}$. Only the line search from $\boldsymbol{y}^{k-1}$ is performed along the direction $\boldsymbol{d}^k$ using the potential function f at each iteration of the method. In standard applications of conjugate gradient methods to nonlinear minimization, the search direction $\boldsymbol{d}^k$ is computed *adaptively* rather than fixed in advance. Our ultimate goal is to see whether conjugate gradient methods can be effectively incorporated in interior point algorithms. Thus it would require more effort to establish a similar result for conjugate gradient methods rather than the Conjugate Direction Method given in Theorem 1.2.

References

[1] A.I. Cohen, Rate of convergence of several conjugate gradient algorithms, *SIAM J. Numer. Anal.*, **9** (1972), 248–259.

[2] R.M. Freund and K.-C. Tan, A method for the parametric center problem, with a strictly monotone polynomial-time algorithm for linear programming, *Mathematics of Operations Research*, **16** (1991), 775–801.

[3] C.C. Gonzaga, An algorithm for solving linear programming programs in $O(n^3L)$ operations, in: N.Megiddo, Ed., *Progress in Mathematical Programming: Interior-Point and Related Methods*, Springer-Verlag, New York, 1989, pp. 1–28.

[4] M. Hestenes, *Conjugate Direction Methods in Optimization*, Springer-Verlag, New York, 1980.

[5] M. Kojima, S. Mizuno and A. Yoshise, A primal-dual interior point algorithm for linear programming, in: N. Megiddo, Ed., *Progress in Mathematical Programming: Interior-Point and Related Methods*, Springer-Verlag, New York, 1989, pp. 29–47.

[6] M. Kojima, S. Mizuno and A. Yoshise, A polynomial-time algorithm for a class of linear complementary problems, *Mathematical Programming*, **44** (1989), 1–26.

[7] N. Megiddo, Pathways to the optimal set in linear programming, in: N. Megiddo, Ed., *Progress in Mathematical Programming: Interior-Point and Related Methods*, Springer-Verlag, New York, 1989, pp. 131–158.

[8] R.D.C. Monteiro and I. Adler, Interior path following primal-dual algorithms. Part I: linear programming, *Mathematical Programming*, **44** (1989), 27–41.

[9] R.D.C. Monteiro and I. Adler, Interior path following primal-dual algorithms. Part II: convex quadratic programming, *Mathematical Programming*, **44** (1989), 43–66.

[10] J. Renegar, A polynomial-time algorithm based on Newton's method for linear programming, *Mathematical Programming*, **40** (1988), 59–94.
[11] G. Sonnevend, An 'analytical centre' for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, in: *Lecture Notes in Control and Information Sciences*, Vol. 84, Springer-Verlag, New York, 1985, pp. 866–876.
[12] J. Stoer, On the relation between quadratic termination and convergence properties of minimization algorithms, *Numer. Math.*, **28** (1977), 343–366.
[13] K. Tanabe, Centered Newton method for mathematical programming, in: M. Iri and K. Yajima, Eds., *System Modeling and Optimization*, Springer-Verlag, New York, 1988, pp. 197–206.
[14] K. Tanabe, Centered Newton method for linear programming: Interior and 'exterior' point method (in Japanese), in: K. Tone, Ed., *New Methods for Linear Programming*, **3**, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan, 1990, pp. 98–100.
[15] P.M. Vaidya, A locally well-behaved potential function and a simple Newton-type method for finding the center of a polytope, in: N. Megiddo, Ed., *Progress in Mathematical Programming: Interior-Point and Related Methods*, Springer-Verlag, New York, 1989, pp. 131–158.
[16] P.M. Vaidya, An algorithm for linear programming which requires $O((m+n)n^2+(m+n)^{1.5}nL)$ arithmetic operations, *Mathematical Programming*, **47** (1990), 175–201.