Convergence Property of the Iri-Imai Algorithm for Some Smooth Convex Programming Problems

S. Zhang*

Communicated by Z.Q. Luo

*Assistant Professor, Department of Econometrics, University of Groningen, Groningen, The Netherlands.

Abstract

In this paper, the Iri-Imai algorithm for solving linear and convex quadratic programming is extended to solve some other smooth convex programming problems. The globally linear convergence rate of this extended algorithm is proved, under the condition that the objective and constraint functions satisfy a certain type of convexity (called the harmonic convexity in this paper). A characterization of this convexity condition is given. In Ref. 14, the same convexity condition is used to prove the convergence of a path-following algorithm.

The Iri-Imai algorithm is a natural generalization of the original Newton algorithm to constrained convex programming. Other known convergent interior point algorithms for smooth convex programming are mainly based on the path-following approach.

Key Words: Convex programming, interior point method, Iri and Imai's algorithm, convergence.

1 Introduction

Since Karmarkar (Ref. 1) presented the first polynomial time interior point algorithm for linear programming, a large number of research papers have been devoted to the interior point method. The focus of the researches was first on developing theoretically and/or empirically more efficient interior point algorithms for linear programming. In this respect, mainly four classes of interior point algorithms have been developed. They are: the projective method (represented by the original Karmarkar's algorithm), the potential reduction method (e.g. Ye (Ref. 2), Freund (Ref. 3) and Gonzaga (Ref. 4)), the affine scaling method (cf. Dikin (Ref. 5), Barnes (Ref. 6) and Vanderbei *et al.* (Ref. 7)) and the path following method (cf. Renegar (Ref. 8), Megiddo (Ref. 9) and Den Hertog *et al.* (Ref. 10)). Some potential reduction algorithms and path following algorithms are shown to have better time complexity than the original Karmarkar algorithm. Numerical results show that the interior point method is indeed a promising approach for linear programming. Some of these algorithms have been proved to work for convex quadratic programming as well.

More recently, the interior point approach has been used to attack some combinatorial optimization problems (cf. Karmarkar (Ref. 11) and Mitchell (Ref. 12)) mainly based on the projective and the potential reduction methods. Other researchers have extended the interior point method to solve some convex programming problems. For the references of the second approach, see Jarre (Ref. 13), Mehrotra and Sun (Ref. 14), Den Hertog *et al.* (Ref. 15) and Den Hertog *et al.* (Ref. 16). To the best knowledge of the author, only the path-following method is so far successfully extended to solve convex programming problems.

Among many variants of the interior point method for linear programming, there is an interesting algorithm proposed by Iri and Imai (see Iri and Imai (Ref. 17)). That algorithm does not fall into the four conventional classifications of the interior point algorithms mentioned above. As a matter of fact, the idea of the Iri-Imai algorithm is based on a multiplicative barrier function approach for nonlinear programming. In simple words, it views a linear programming problem (in the form that all constraints are inequalities) as a constrained nonlinear programming problem (supposing that the interior of the feasible region is nonempty); and constructs a multiplicative barrier function for points inside the interior of the feasible region as it is usual for constrained nonlinear programming. After having such a multiplicative barrier function, Iri and Imai proposed to apply the Newton method using line search to optimize the barrier function. The multiplicative barrier function in the linear programming case, however, resembles very well the potential function. In Iri and Imai (Ref. 17), it was shown that this algorithm has a locally quadratic convergence rate. Numerical experiments presented in the same paper showed that this algorithm converges always globally, and it converges very fast indeed. A proof of the global convergence property was given in Zhang and Shi (Ref. 18) and Zhang (Ref. 19). Based on this convergence proof, the polynomiality of the Iri-Imai algorithm follows by replacing the exact line search with some fixed step searches. However, the number of iterations estimated in Zhang and Shi (Ref. 18) and Zhang (Ref. 19) is about $\mathcal{O}(m^8L)$ comparing to $\mathcal{O}(mL)$ of Karmarkar's algorithm, where m is the number of constraints and L is the inputlength of the problem. Later, Imai (Ref. 20) proved that the running time bound of the algorithm is at most $\mathcal{O}(m^4L)$ for linear programming. In Ref. 21, Imai further showed that the bound is at most $\mathcal{O}(m^2L)$ for linear programming. Recently, Iri (cf. (Ref. 22)) gave an elegant proof which shows that the Iri-Imai algorithm actually has the same order of polynomial running time bound as the original Karmarkar algorithm for linear programming. Moreover, he showed in (Ref. 22) that the Iri-Imai algorithm can be extended to solve convex quadratic programming with the same polynomial running time bound. In this paper, using similar approaches as in Ref. 22, the convergence result of the Iri-Imai algorithm applied to a larger class of smooth convex programming problems is presented. More precisely, under some smoothness and convexity assumptions we prove that the Iri-Imai algorithm has a globally linear convergence rate for convex programming. The main condition on the objective and constraint functions used to prove the convergence is called the harmonic convexity. The same condition was used in Mehrotra and Sun (Ref. 14) as well. A characterization of the harmonic convexity is given in this paper. This condition is easier to check and requires less continuity than the so called relative Lipschitz condition used in (Refs. 13, 15 and 16).

This paper is organized as follows. In Section 2, we introduce the Iri-Imai algorithm for convex programming. In Section 3, the convergence result is presented. We conclude the paper in Section 4.

2 The Iri-Imai Algorithm for Convex Programming

Consider the following convex programming problem

$$(P) \qquad \begin{array}{ll} \min & g_0(x) \\ \text{s.t.} & g_i(x) \geq 0,\ i = 1, 2, ..., m, \\ & x \in \mathcal{R}^n, \end{array}$$

where g_i is second order continuously differentiable, for $i = 0, 1, ..., m$, g_0 is convex and g_i is concave for $i = 1, 2, ..., m$.

We will assume from now on that $m \geq 1$. As we will see later, Iri and Imai's algorithm is an extension of Newton's method for constrained problems.

To simplify the analysis, we first make the following assumption on (P).

Assumption 2.1 The optimum value of (P) is known, for simplicity, to be zero.

We observe that this assumption is not essential (see Section 4) and can be dropped if the forthcoming algorithm is properly modified.

Now we define *harmonic convexity* as follows. Note that two square matrices M_1 and M_2 satisfy $M_1 \leq M_2$ iff $M_2 - M_1$ is a positive semi-definite matrix.

Definition 2.1 A second order continuously differentiable convex function f is called harmonically convex on its convex domain X iff there exists a positive constant λ such that

$$\frac{1}{\lambda} \nabla^2 f(y) \leq \nabla^2 f(x) \leq \lambda \nabla^2 f(y)$$

holds for any x and y in X, where $\nabla^2 f$ denotes the Hessian matrix of f. Such a constant λ is called a harmonic constant.

In convex analysis, a function is called uniformly convex if for any point in its domain the Hessian matrix exists and is positive definite, moreover, the largest and the smallest eigenvalues of the Hessian matrix are strictly bounded by some positive constants from both above and below respectively. The following lemma is readily seen.

Lemma 2.1 All linear functions, convex quadratic functions and uniformly convex functions are harmonically convex.

We will give a characterization of the harmonic convexity in the following lemma.

Lemma 2.2 A function f is harmonically convex on $\mathcal{R}^n$ iff there exists a nonsingular matrix A such that $f(Ax) = f_1(x') + f_2(x'')$, where x' and x'' form a partition of x, f_1 is a uniformly convex function and f_2 is a linear function.

Proof. Fix a point $y \in \mathcal{R}^n$. The Hessian $\nabla^2 f(y)$ is positive semi-definite and so there exists a nonsingular matrix A such that

$$\nabla^2 f(y) = A^{-T} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} A^{-1}.$$

Consider the function $\overline{f}(x) := f(Ax)$. Since $\nabla^2 \overline{f}(x) = A^T \nabla^2 f(Ax) A$, by the harmonic convexity of f we have $(1/\lambda)\nabla^2 f(y) \leq \nabla^2 f(Ax) \leq \lambda \nabla^2 f(y)$. Therefore

$$\frac{1}{\lambda} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \leq \nabla^2 \overline{f}(x) \leq \lambda \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

Based on the above inequalities, decompose the Hessian matrix $\nabla^2 \overline{f}(x)$ accordingly into four blocks. Let x' correspond to variables involved in the left-up block and x'' the variables involved in the right-lower block. Due to the above inequalities, it can be verified that only the left-up block, namely $\nabla^2_{x'} \overline{f}(x')$, is nonzero and satisfies $(1/\lambda)I \leq \nabla^2_{x'} \overline{f}(x') \leq \lambda I$. This proves that $f(Ax)$ can be indeed decomposed into two required separate parts.

It is easy to check that this condition is also sufficient for f to be harmonically convex. $\square$

Now we state two more assumptions on the problem (P).

Assumption 2.2 The functions g_0 and $-g_i$ $(1 \leq i \leq m)$ are all harmonically convex. For simplicity we let λ be their common harmonic constant.

Assumption 2.3 The convex programming problem (P) satisfies the Slater condition, i.e. there exists some $x \in \mathcal{R}^n$ such that $g_i(x) > 0$ for $i = 1, 2, ..., m$.

Since the functions g_i, $i = 1, 2, ..., m$, are all continuous, the Slater condition implies that the feasible region of (P) has a nonempty interior. In fact, the Slater condition is sufficient to guarantee that the set formed by the optimal Lagrange multipliers is nonempty and compact (cf. Bertsekas (Ref. 24) and Rockafellar (Ref. 25)).

Let the feasible set of (P) be $F := \{x : g_i(x) \geq 0, 1 \leq i \leq m\} \subseteq \mathcal{R}^n$. By Assumption 2.3 and due to the fact that g_i $(1 \leq i \leq m)$ is concave, we know that the set F is full dimensional and convex. We denote the nonempty interior of F by $\overset{o}{F}$, given by

$$\overset{o}{F} := \{x : g_i(x) > 0, 1 \leq i \leq m\} \subseteq \mathcal{R}^n.$$

Notice that $\overset{o}{F}$ is an open and convex set.

In order to simplify the analysis, we further make the following assumption.

Assumption 2.4 The feasible set F of (P) is bounded. Namely, there is a constant M such that $\|x\| \leq M$ for any $x \in F$.

Now we define a multiplicative barrier function G for the problem (P) as follows:

$$G(x) := \frac{(g_0(x))^{m+l}}{\prod_{i=1}^{m} g_i(x)}, \text{ for } x \in \overset{o}{F} \tag{1}$$

where $l > 1$ is some given positive integer.

We observe that the multiplicative barrier function G is well defined on the open and convex set $\overset{o}{F}$. Moreover, we will see in Lemma 2.3 that G is a strictly convex function on $\overset{o}{F}$ under the following assumption.

Assumption 2.5 The problem (P) is assumed to satisfy one of the following two conditions: 1) one of the functions g_0, $-g_1$, ..., $-g_m$ is strictly convex; 2) $rank\{\nabla g_i(x) : i = 0, 1, ..., m\} = n$ for all $x \in \overset{o}{F}$.

Lemma 2.3 If $l > 1$ and Assumption 2.5 holds, then the multiplicative barrier function G is strictly convex on the open and convex set $\overset{o}{F}$.

Proof. Cf. Theorem 5.16 of Avriel *et al.* (Ref. 23) and Iri (Ref. 22).

□

For $x \in \overset{o}{F}$ and $g_0(x) > 0$ (x not optimal), since $G(x)$ is positive in this case, we define

$$g(x) := \log G(x) = (m + l) \log g_0(x) - \sum_{i=1}^{m} \log g_i(x). \qquad (2)$$

The function g is called the *logarithmic barrier function*. Notice that g is a *quasi-convex* function since G is convex.

The following lemma shows that by using the multiplicative barrier function G or the logarithmic barrier function g, we have essentially converted the constrained problem (P) into an unconstrained problem.

Lemma 2.4 For any sequence $\{x^k : k \geq 1\}$ with $x^k \in \overset{o}{F}$, $k \geq 1$, suppose that $\lim_{k\to+\infty} G(x^k) = 0$, or equivalently $\lim_{k\to+\infty} g(x^k) = -\infty$, then any cluster point of $\{x^k : k \geq 1\}$ is an optimal solution of (P).

Proof. If $\lim_{k\to+\infty} G(x^k) = 0$, by (1) and Assumption 2.4 we conclude that

$$\lim_{k\to+\infty} g_0(x^k) = 0.$$

By Assumption 2.1, the claimed result follows. □

Based on Lemma 2.4, it is clear that to solve the problem (P) it suffices to minimize G or g in F. To minimize the twice differentiable convex function G, the well known Newton method is appropriate. This results in the following Iri-Imai algorithm for the convex programming problem (P) (cf. the original Iri-Imai algorithm for linear programming (Ref. 17)).

The Iri-Imai Algorithm for Convex Programming

For this algorithm, the input includes the initial interior point $x^0 \in \overset{o}{F}$ *and the precision parameter* $\epsilon > 0$. *The output consists in a sequence of solutions* $x^k \in \overset{o}{F}$, $k \geq 1$.

Step 0 Let $k := 0$.

Step 1 Solve the Newton equation

$$\nabla^2 G(x^k)\xi^k = -\nabla G(x^k).$$

Find $x^{k+1} := x^k + t_k \xi^k$ such that

$$G(x^k + t_k\xi^k) = \min_{t \geq 0} G(x^k + t\xi^k).$$

Go to Step 2.

Step 2 If $G(x^{k+1}) < \epsilon$, stop; otherwise, let $k := k + 1$ and go to Step 1.

Remark 2.1 The above described procedure requires an exact line search procedure (at Step 1). As we will see from the analysis presented in Section 3, the globally linear convergence holds even for some *inexact* search procedure.

Remark 2.2 For an non-optimal x on the boundary of F, i.e. $g_i(x) = 0$ for some $1 \leq i \leq m$ and $g_0(x) > 0$, it is easy to see that $\underline{\lim}_{y \in \overset{o}{F}, y \to x} G(y) = +\infty$. This implies by using the line search argument that if $x^k \in \overset{o}{F}$ and x^k is not optimal, then either x^{k+1} is optimal or $x^{k+1} \in \overset{o}{F}$. So if we let the precision parameter ϵ be 0, and if the whole sequence $\{x^k\}$ produced by the above algorithm is not finite, then the whole sequence will be contained in $\overset{o}{F}$.

3 Analysis

In this section we will first introduce some relations between the first order and second order derivatives of G and g.

For a given $x \in \overset{o}{F}$ (x not optimal) we have

$$\nabla g(x) = \frac{\nabla G(x)}{G(x)} = (m+l)\frac{\nabla g_0(x)}{g_0(x)} - \sum_{i=1}^{m} \frac{\nabla g_i(x)}{g_i(x)} \tag{3}$$

and

$$\begin{aligned} \nabla^2 g(x) &= \frac{\nabla^2 G(x)}{G(x)} - \frac{\nabla G(x)}{G(x)} \cdot \frac{\nabla G(x)^T}{G(x)} \\ &= (m+l)(\frac{\nabla^2 g_0(x)}{g_0(x)} - \frac{\nabla g_0(x)}{g_0(x)} \cdot \frac{\nabla g_0(x)^T}{g_0(x)}) \\ &\quad - \sum_{i=1}^{m}(\frac{\nabla^2 g_i(x)}{g_i(x)} - \frac{\nabla g_i(x)}{g_i(x)} \cdot \frac{\nabla g_i(x)^T}{g_i(x)}). \end{aligned} \tag{4}$$

To simplify the notations we denote the scaled gradient and Hessian by

$$\tilde{\nabla} f := \frac{\nabla f}{f} \text{ and } \tilde{\nabla}^2 f := \frac{\nabla^2 f}{f}.$$

Now let the Newton direction at the point $x \in \overset{o}{F}$ be ξ, i.e.,

$$\xi := -(\tilde{\nabla}^2 G(x))^{-1} \tilde{\nabla} G(x) \tag{5}$$

and let

$$h := -\nabla g(x)^T \xi. \tag{6}$$

It follows from (4), (5) and (6) that

$$h = \xi^T \nabla^2 g(x) \xi + h^2. \tag{7}$$

Concerning the Newton direction ξ we have the following lemma.

Lemma 3.1

$$\xi = arg \max_{\eta \in \mathcal{R}^n \setminus \{0\}} \frac{-\tilde{\nabla} G(x)^T \eta}{\sqrt{\eta^T \tilde{\nabla}^2 G(x) \eta}}.$$

Proof. See Iri (Ref. 22). □

Let the optimal solution of (P) be x^*. It follows from Lemma 3.1 that

$$\sqrt{h} \geq \frac{-\tilde{\nabla} G(x)^T(x^* - x)}{\sqrt{(x^* - x)^T \tilde{\nabla}^2 G(x)(x^* - x)}}. \tag{8}$$

Since $g_0(x^*) = 0$ and $g_i(x^*) \geq 0$ $(i = 1, 2, ..., m)$, it follows from the convexity of g_0 and the concavity of g_i $(i = 1, 2, ..., m)$ that

$$0 = g_0(x^*) \geq g_0(x) + \nabla g_0(x)^T(x^* - x) \tag{9}$$

and

$$0 \leq g_i(x^*) \leq g_i(x) + \nabla g_i(x)^T(x^* - x). \tag{10}$$

Moreover, by the mean value theorem and the harmonic convexity of g_0 and $-g_i$ $(i = 1, 2, ..., m)$, we obtain

$$\begin{aligned} 0 = g_0(x^*) &= g_0(x) + \nabla g_0(x)^T(x^* - x) + \frac{1}{2}(x^* - x)^T \nabla^2 g_0(\tilde{x}_0)(x^* - x) \\ &\geq g_0(x) + \nabla g_0(x)^T(x^* - x) \\ &\quad + \frac{1}{2\lambda}(x^* - x)^T \nabla^2 g_0(x)(x^* - x) \end{aligned} \tag{11}$$

and similarly

$$\begin{aligned} 0 \leq g_i(x^*) &= g_i(x) + \nabla g_i(x)^T(x^* - x) + \frac{1}{2}(x^* - x)^T \nabla^2 g_i(\tilde{x}_i)(x^* - x) \\ &\leq g_i(x) + \nabla g_i(x)^T(x^* - x) \\ &\quad + \frac{1}{2\lambda}(x^* - x)^T \nabla^2 g_i(x)(x^* - x), \end{aligned} \tag{12}$$

where $\tilde{x}_i$ $(i = 0, 1, ..., m)$ is a point in the segment formed by x and x^*.

Let

$$w_0 := -\tilde{\nabla} g_0(x)^T(x^* - x) - 1,$$

and

$$w_i := \tilde{\nabla} g_i(x)^T(x^* - x) + 1, \; i = 1, 2, ..., m.$$

We observe from (9) and (10) that $w_i \geq 0$ $(i = 0, 1, ..., m)$.

Moreover, from (11) and (12) we conclude that

$$(x^* - x)^T \tilde{\nabla}^2 g_0(x)(x^* - x) \leq 2\lambda w_0 \tag{13}$$

and

$$(x^* - x)^T \tilde{\nabla}^2 g_i(x)(x^* - x) \geq -2\lambda w_i \tag{14}$$

for $i = 1, 2, ..., m$.

We have now

$$\begin{aligned} -\tilde{\nabla} G(x)^T (x^* - x) &= -\nabla g(x)^T (x^* - x) = (m + l)(w_0 + 1) + \sum_{i=1}^{m} (w_i - 1) \\ &= (m + l) w_0 + \sum_{i=1}^{m} w_i + l \end{aligned} \tag{15}$$

and

$$\begin{gathered} (x^* - x)^T \tilde{\nabla}^2 G(x)(x^* - x) = \\ (x^* - x)^T \nabla^2 g(x)(x^* - x) + (\nabla g(x)^T (x^* - x))^2. \end{gathered} \tag{16}$$

Using (13) and (14), the first term on the right hand side of the equation (16) can be estimated as

$$\begin{aligned} &(x^* - x)^T \nabla^2 g(x)(x^* - x) = (m + l)(x^* - x)^T \tilde{\nabla}^2 g_0(x)(x^* - x) \\ &\quad - (m + l)(w_0 + 1)^2 - \sum_{i=1}^{m} (x^* - x)^T \tilde{\nabla}^2 g_i(x)(x^* - x) + \sum_{i=1}^{m} (w_i - 1)^2 \\ &\leq 2(m + l)\lambda w_0 - (m + l)(w_0 + 1)^2 + 2\lambda \sum_{i=1}^{m} w_i + \sum_{i=1}^{m} (w_i - 1)^2 \\ &\leq 2\lambda((m + l) w_0 + \sum_{i=1}^{m} w_i) + \sum_{i=1}^{m} w_i^2 - 2\sum_{i=1}^{m} w_i - l \\ &\leq 2\lambda((m + l) w_0 + \sum_{i=1}^{m} w_i) + (\sum_{i=1}^{m} w_i)^2 - 2\sum_{i=1}^{m} w_i - l \\ &\leq 2\lambda((m + l) w_0 + \sum_{i=1}^{m} w_i) + ((m + l) w_0 + \sum_{i=1}^{m} w_i)^2 \\ &\leq (2\lambda + 1)((m + l) w_0 + \sum_{i=1}^{m} w_i + l)^2. \end{aligned} \tag{17}$$

From (8), (15), (16) and (17) we obtain the following result.

Theorem 3.1 If the problem (P) satisfies Assumptions 2.1, 2.2, 2.3, 2.4 and 2.5, then for a non-optimal point x it holds that $h \geq 1/(2\lambda + 2)$, where h is defined by (6).

Now we proceed to estimate how much the logarithmic barrier function can be decreased by searching along the Newton direction.

We note that a feasible steplength $t\,(> 0)$ can be guaranteed if the following holds

$$\begin{aligned} g_i(x + t\xi) &= g_i(x) + t\nabla g_i(x)^T \xi + \frac{t^2}{2}\xi^T \nabla^2 g_i(\tilde{x})\xi \\ &\geq g_i(x) + t\nabla g_i(x)^T \xi + \frac{t^2}{2}\lambda \xi^T \nabla^2 g_i(x)\xi > 0, \end{aligned}$$

for $i = 1, 2, ..., m$. Or, equivalently

$$1 + t\tilde{\nabla} g_i(x)^T \xi + \frac{t^2}{2}\lambda \xi^T \tilde{\nabla}^2 g_i(x)\xi > 0, \tag{18}$$

for $i = 1, 2, ..., m$.

In order to determine how large t can be without violating (18) we introduce the following notations:

$$\begin{aligned} a_{01} &:= \tilde{\nabla} g_0(x)^T \xi \\ a_{02} &:= \xi^T \tilde{\nabla}^2 g_0(x)\xi \\ a_{i1} &:= \tilde{\nabla} g_i(x)^T \xi \\ a_{i2} &:= \xi^T \tilde{\nabla}^2 g_i(x)\xi \end{aligned}$$

for $i = 1, 2, ..., m$.

Now from (3), (4), (6) and (7) it follows that

$$h = -(m + l)a_{01} + \sum_{i=1}^{m} a_{i1} \tag{19}$$

$$h - h^2 = (m + l)a_{02} - \sum_{i=1}^{m} a_{i2} - (m + l)a_{01}^2 + \sum_{i=1}^{m} a_{i1}^2. \tag{20}$$

By introducing

$$\overline{A}_1 := \frac{1}{m}\sum_{i=1}^{m} a_{i1} \tag{21}$$

and

$$\sigma_1^2 := \frac{1}{m}\sum_{i=1}^{m}(a_{i1} - \overline{A}_1)^2 \tag{22}$$

$$c := (m+l)a_{02} - \sum_{i=1}^{m} a_{i2} \ (\geq 0) \tag{23}$$

the equations (19) and (20) can be rewritten as

$$h = -(m+l)a_{01} + m\overline{A}_1 \tag{24}$$

$$h - h^2 = -(m+l)a_{01}^2 + m\overline{A}_1^2 + m\sigma_1^2 + c. \tag{25}$$

Solving the equations (24) and (25) in terms of a_{01} and $\overline{A}_1$ we obtain

$$a_{01} = -\frac{h}{l} \pm \sqrt{\frac{m}{(m+l)l}(h - \frac{l-1}{l}h^2 - m\sigma_1^2 - c)} \tag{26}$$

$$\overline{A}_1 = -\frac{h}{l} \pm \sqrt{\frac{m+l}{ml}(h - \frac{l-1}{l}h^2 - m\sigma_1^2 - c)}. \tag{27}$$

Since a_{01} and $\overline{A}_1$ are real numbers, we conclude that

$$h - \frac{l-1}{l}h^2 - m\sigma_1^2 - c \geq 0. \tag{28}$$

The following two lemmas follow immediately from (28).

Lemma 3.2 If $l > 1$, then $h \leq l/(l-1)$.

Lemma 3.3 $m\sigma_1^2 + c \leq h(1 - h(l-1)/l)$.

By the definitions of $\overline{A}_1$ and σ_1 (cf. (21) and (22)) we now use the well known inequality

$$|\, a_{i1}\,| \le |\,\overline{A}_1\,| + \sqrt{m-1}\,\sigma_1$$

and so together with (27) we have the following inequalities

$$\begin{aligned} |\ a_{i1}\ | &\le |\ \overline{A}_1\ | + \sqrt{m}\,\sigma_1 \\ &\le \frac{h}{l} + \sqrt{\frac{m+l}{ml}(h - \frac{l-1}{l}h^2 - m\sigma_1^2 - c)} + \sqrt{m}\sigma_1, \end{aligned}$$

for $i = 1, 2, ..., m$.

Maximizing the right hand side of the above inequality in terms of σ_1, it follows that

$$\begin{aligned} |\ a_{i1}\ | &\le \frac{h}{l} + \sqrt{(\frac{m+l}{ml} + 1)(h - \frac{l-1}{l}h^2 - c)} \\ &\le \frac{h}{l} + \sqrt{(\frac{m+l}{ml} + 1)(h - \frac{l-1}{l}h^2)}\,, \end{aligned} \tag{29}$$

for $i = 1, 2, ..., m$.

Letting

$$u := \frac{h}{l} + \sqrt{(\frac{m+l}{ml} + 1)(h - \frac{l-1}{l}h^2)}\,, \tag{30}$$

the inequality (29) is now rewritten as

$$|\ a_{i1}\ | \le u, \tag{31}$$

for $i = 1, 2, ..., m$.

Moreover, by the definition of c (cf. (23)), using Lemma 3.3 and noticing the fact that $a_{02} \ge 0$ and $a_{i2} \le 0$ (since g_0 is convex and g_i is concave for $i = 1, 2, ..., m$), we obtain

$$a_{i2} \ge -c \ge -h(1 - \frac{l-1}{l}h) \tag{32}$$

for $i = 1, 2, ..., m$.

Now we let

$$v := \frac{2}{u + \sqrt{u^2 + 2\lambda(1 - \frac{l-1}{l}h)h}}. \tag{33}$$

Observe from (30) and (33), using Theorem 3.1 and Lemma 3.2, that u and v are strictly bounded from zero by some positive constants.

Furthermore, we have the following result

Theorem 3.2 For any $x \in \overset{o}{F}$, if the Newton direction ξ is defined according to (5), then for any $0 < t < v$ we have $x + t\xi \in \overset{o}{F}$.

Proof. It is easy to see that if $0 < t < v$ then

$$1 - ut - \frac{\lambda h(1 - \frac{l-1}{l}h)}{2} t^2 > 0$$

and so

$$1 + t a_{i1} + \frac{t^2}{2} \lambda a_{i2} \geq 1 - ut - \frac{\lambda h(1 - \frac{l-1}{l}h)}{2} t^2 > 0.$$

From (18) we know that the above inequality implies $x + t\xi \in \overset{o}{F}$.

□

Theorem 3.2 shows that in the Iri-Imai algorithm a certain search step along the Newton direction is allowed without violating the feasibility. This property is essential for our analysis. Now we will show that by properly choosing the steplength within the region given by the interval $(0, v)$, at least some fixed amount of reduction in the logarithmic barrier function can be obtained.

Let an interior point $x \in \overset{o}{F}$, and let $0 < t < v$. By the mean value theorem, we have

$$g(x + t\xi) - g(x) = t \nabla g(x)^T \xi + \frac{t^2}{2} \xi^T \nabla^2 g(\tilde{x}) \xi, \tag{34}$$

where $\tilde{x} = x + \mu t \xi$, for some $\mu \in (0, 1)$.

Notice by the convexity of g_0 that

$$g_0(x + \mu t\xi) \geq g_0(x) + \mu t \nabla g_0(x)^T \xi$$

and by the harmonic convexity of $-g_i$ for $i = 1, 2, ..., m$, that

$$g_i(x + \mu t\xi) \geq g_i(x) + \mu t \nabla g_i(x)^T \xi + \frac{(\mu t)^2}{2} \lambda \xi^T \nabla^2 g_i(x) \xi.$$

Thus we have

$$
\begin{aligned}
\xi^T \nabla^2 g(\tilde{x}) \xi &= (m+l)\xi^T \tilde{\nabla}^2 g_0(\tilde{x})\xi - \sum_{i=1}^{m} \xi^T \tilde{\nabla}^2 g_i(\tilde{x})\xi \\
&\quad + \sum_{i=1}^{m} (\xi^T \tilde{\nabla} g_i(\tilde{x}))^2 - (m+l)(\xi^T \tilde{\nabla} g_0(\tilde{x}))^2 \\
&\leq (m+l)\lambda \frac{a_{02}}{1+\mu t a_{01}} - \lambda \sum_{i=1}^{m} \frac{a_{i2}}{1+\mu t a_{i1} + \frac{(\mu t)^2}{2}\lambda a_{i2}} \\
&\quad + \sum_{i=1}^{m} \frac{(\mid a_{i1} \mid - \mu \lambda a_{i2})^2}{(1+\mu t a_{i1} + \frac{(\mu t)^2}{2}\lambda a_{i2})^2}\,. \qquad (35)
\end{aligned}
$$

To further estimate the right hand side of (35), we note the following lemma.

First, let

$$
\overline{v} := \frac{1}{u + \sqrt{u^2 + \lambda h(1 - \frac{l-1}{l} h)}}.
$$

Note that $\overline{v} \leq v$.

Lemma 3.4 If $0 < t \leq \overline{v}$ then $1+\mu t a_{i1} + (\mu t)^2 \lambda a_{i2}/2 \geq 1/2$ for $i = 1, 2, ..., m$, and $1 + \mu t a_{01} \geq 1/2$.

Proof. First it is easy to see that if $0 < t \leq \overline{v}$ then for $i = 1, 2, ..., m$,

$$
1 + \mu t a_{i1} + \frac{(\mu t)^2}{2}\lambda a_{i2} \geq 1 - ut - \frac{\lambda h(1 - \frac{l-1}{l}h)}{2} t^2 \geq \frac{1}{2}.
$$

Moreover, notice that $\mid a_{01} \mid \leq u$ and so

$$
\frac{1}{\overline{v}} > 2u \geq 2 \mid a_{01} \mid .
$$

This implies that if $0 < t \leq \overline{v}$ then $1 + \mu t a_{01} \geq 1/2$. The lemma is proved. □

Now let $0 < t \leq \overline{v}$. Using Lemma 3.3, Lemma 3.4, and noticing (23) and (35) we have

$$\begin{aligned}
\xi^T \nabla^2 g(\tilde{x})\xi &\leq 2(m+l)\lambda a_{02} - 2\lambda \sum_{i=1}^{m} a_{i2} + 4\sum_{i=1}^{m}(\mid a_{i1} \mid -\lambda a_{i2})^2 \\
&\leq 2\lambda c + 4(\sqrt{\sum_{i=1}^{m} a_{i1}^2} + \lambda\sqrt{\sum_{i=1}^{m} a_{i2}^2})^2 \\
&\leq 2\lambda h(1 - \frac{l-1}{l}h) + 4(\sqrt{\sum_{i=1}^{m} a_{i1}^2} + \lambda c)^2. \qquad (36)
\end{aligned}$$

Furthermore, by (21), (22) and (27) we have

$$\begin{aligned}
\sum_{i=1}^{m} a_{i1}^2 &= m\overline{A}_1^2 + m\sigma_1^2 \\
&\leq m(\frac{h}{l} + \sqrt{\frac{m+l}{ml}(h - \frac{l-1}{l}h^2 - c - m\sigma_1^2)})^2 + m\sigma_1^2 \\
&\leq m(\frac{h}{l} + \sqrt{\frac{m+l}{ml}(h - \frac{l-1}{l}h^2)})^2. \qquad (37)
\end{aligned}$$

Now replacing (37) into (36) we obtain

$$\begin{aligned}
\xi^T \nabla^2 g(\tilde{x})\xi &\leq 2\lambda h(1 - \frac{l-1}{l}h) + 4(\frac{\sqrt{m}\,h}{l} + \sqrt{\frac{m+l}{l}h(1 - \frac{l-1}{l}h)} + \lambda c)^2 \\
&\leq 2\lambda h(1 - \frac{l-1}{l}h) + 4(\frac{\sqrt{m}\,h}{l} + \sqrt{\frac{m+l}{l}h(1 - \frac{l-1}{l}h)} \\
&\quad + \lambda h(1 - \frac{l-1}{l}h))^2. \qquad (38)
\end{aligned}$$

To simplify the notation, let

$$r := 2\lambda h(1 - \frac{l-1}{l}h) + 4(\frac{\sqrt{m}\,h}{l} + \sqrt{\frac{m+l}{l}h(1 - \frac{l-1}{l}h)} + \lambda h(1 - \frac{l-1}{l}h))^2 \qquad (39)$$

and we rewrite (38) as

$$\xi^T \nabla^2 g(\tilde{x}) \xi \leq r. \tag{40}$$

Note that if the parameter $l > 1$ then u, v, $\overline{v}$, h and r are all positive.

By (6), (34) and (40) the following lemma is readily seen.

Lemma 3.5 For $t \in (0, \overline{v})$ it holds that

$$g(x + t\xi) - g(x) \leq -th + \frac{t^2}{2} r.$$

Now, we let the parameter $l = m+1$. Since $m \geq 1$, by noticing Lemma 3.2 we have $h(1 - h(l-1)/l) \leq l/(4(l-1)) \leq 1/2$. Therefore,

$$r \leq 2\lambda \cdot \frac{1}{2} + 4(1 + \sqrt{\frac{m+l}{l} \cdot \frac{1}{2}} + \lambda \cdot \frac{1}{2})^2 \leq \lambda^2 + 9\lambda + 16,$$

$$u = \frac{h}{l} + \sqrt{(\frac{m+l}{ml} + 1)h(1 - \frac{l-1}{l} h)} \leq 2$$

and

$$\overline{v} = \frac{1}{u + \sqrt{u^2 + \lambda h(1 - \frac{l-1}{l} h)}} \geq \frac{1}{2 + \sqrt{4 + \frac{\lambda}{2}}}.$$

Let

$$\bar{t} := \frac{1}{(\lambda^2 + 9\lambda + 16)(\lambda + 1)}.$$

It is clear that $0 < \bar{t} < \overline{v}$. From Lemma 3.5, it follows that

$$g(x + \bar{t}\xi) - g(x) \leq -\bar{t}h + \frac{\bar{t}^2}{2} r \leq -\frac{1}{2(\lambda^2 + 9\lambda + 16)(\lambda + 1)^2}.$$

Let

$$\delta := \frac{1}{2(\lambda^2 + 9\lambda + 16)(\lambda + 1)^2} (= \mathcal{O}(\frac{1}{(\lambda + 1)^4})). \tag{41}$$

Now we are ready to present the main result of this paper. First we note by Assumption 2.4 and the continuity of the constraint function g_i ($i = 1, 2, ..., m$) in the compact feasible region F that there exists some constant $N > 0$ such that $g_i(x) \leq N$ ($i = 1, 2, ..., m$) for all $x \in F$.

Theorem 3.3 For the convex programming problem (P), suppose that Assumptions 2.1, 2.2, 2.3, 2.4 and 2.5 hold, and $x^0 \in \overset{o}{F}$ is the initial interior point. We let the parameter $l = m + 1$. Then the Iri-Imai algorithm has at least a globally linear convergence rate in terms of the multiplicative barrier function value for solving (P), i.e. for the sequence of points $\{x^k : k \geq 0\}$ produced by the algorithm, it holds that

$$G(x^{k+1}) \leq \exp(-\delta)G(x^k)$$

for $k \geq 0$.

Moreover, for any given $p > 0$, we will have $g_0(x^K) < 2^{-p}$ in at most $K = \mathcal{O}((mp + m \log N + g(x^0))(\lambda + 1)^4)$ steps, where g is the logarithmic barrier function and g_0 is the objective function.

Proof. For any $x^k \in \overset{o}{F}$, we see from Lemma 3.5 that if a step length t is taken to be $\bar{t}$, then

$$\min_{t \geq 0}\{g(x^k + t\xi^k) - g(x^k)\} \leq g(x^k + \bar{t}\xi^k) - g(x^k) \leq -\delta.$$

This means that

$$g(x^{k+1}) - g(x^k) \leq -\delta \tag{42}$$

and so

$$G(x^{k+1}) \leq \exp(-\delta)G(x^k)$$

for $k \geq 0$.

This proves the first part of the theorem.

By (42) we have

$$g(x^k) \leq g(x^0) - k\delta$$

for $k \geq 0$.

The second part of Theorem 3.3 follows immediately from the above inequality, equation (41) and the following inequality:

$$(m + l) \log g_0(x^k) - m \log N \leq g(x^k).$$

□

Remark 3.1 In the case of linear programming or convex quadratic programming, where the harmonic constant λ can be chosen to be 1, Theorem 3.3 implies that the Iri-Imai algorithm needs at most $\mathcal{O}(mL)$ steps to get close enough to the optimal point (set $p := L$ in this case, where L is the inputlength of the problem), assuming that the initial point x^0 satisfies $g(x^0) = \mathcal{O}(mL)$. This gives exactly the same result as in Iri (Ref. 22).

4 Conclusions

Iri and Imai's algorithm seems to be a natural generalization of Newton's algorithm for constrained convex programming problems. Iri and Imai (Ref. 17) showed that under some non-degeneracy assumptions and if line-search is used, then the Iri-Imai algorithm actually has a locally quadratic convergence rate for linear programming. There is no reason to assume that such a locally fast convergence rate does not hold for some smooth convex programming problems. Certainly, to prove locally fast convergence, an exact line-search procedure and some continuity of the Hessian matrices should be required. It remains a topic for future research.

Assumption 2.1 in this paper is not essential. One needs only a lower bound on the optimal value. The lower bound can be updated at each step in such a way that $h \in [1/(2\lambda+2), l/(l-1)]$ (cf. Theorem 3.1 and Lemma 3.2). The other proofs remain the same. Notice that if a strict lower bound b of the optimal value is used, then the multiplicative barrier function $G_b(x) := (g_0(x) - b)^{m+l}/(\prod_{i=1}^{m} g_i(x))$ will have a unique minimum point in $\overset{o}{F}$, since in this case G_b remains strictly convex in $\overset{o}{F}$ and attains plus infinity on the boundary of F. The path formed by the minimum points when the lower bound goes up to the true optimal value resembles the path studied in the path-following approach.

Assumption 2.4 is not essential as well. We need only to assume that the set of optimal points is bounded. Because if the initial point is properly chosen, we may add some constraints using the information about the upper bound of the objective value. In this way, we may exclude some part of the feasible region where no optimal point will be contained and at the same time keep the new feasible region bounded.

In the existing literature, mainly only the path-following method in the interior point approach is generalized to solve convex programming (Refs. 13-16). In Refs. 13, 15 and 16, the so called Relative Lipschitz Condition on the objective and the constraint functions is required to prove the convergence. The Relative Lipschitz Condition is difficult to check and requires more continuity on the Hessian matrices.

References

1. KARMARKAR, N., *A New Polynomial Time Algorithm for Linear Programming*, Combinatorica, Vol. 4, pp. 373-395, 1984.

2. YE, Y., *An $\mathcal{O}(n^3L)$ Potential Reduction Algorithm for Linear Programming*, Mathematical Programming, Vol. 50, pp. 239-258, 1991.

3. FREUND, R. M., *Polynomial Algorithms for Linear Programming Based Only on Primal Scaling and Projected Gradients of a Potential Function*, Mathematical Programming, Vol. 51, pp. 293-222, 1991.

4. GONZAGA, C. C., *Polynomial Affine Algorithms for Linear Programming*, Report 139/88, Universidade Federal do Rio de Janeiro, 1988.

5. DIKIN, I. I., *Iterative Solution of Problems of Linear and Quadratic Programming*, Soviet Mathematics Doklady, Vol. 8, pp. 674-675, 1974.

6. BARNES, E. R., *A Variation of Karmarkar's Algorithm for Solving Linear Programming Problems*, Mathematical Programming, Vol. 36, pp. 174-182, 1986.

7. VANDERBEI, R. J., Meketon, M. S., and Freedman, B. A., *A Modification of Karmarkar's Linear Programming Algorithm*, Algorithmica, Vol. 1, pp. 395-407, 1986.

8. RENEGAR, J., *A Polynomial-Time Algorithm Based on Newton's Method for Linear Programming*, Mathematical Programming, Vol. 40, pp. 59-69, 1988.

9. MEGIDDO, N., *Pathways to the Optimal Set in Linear Programming*, Progress in Mathematical Programming, edited by N. Megiddo, Springer-Verlag, New York, New York, pp. 131-158, 1989.

10. DEN HERTOG, D., ROOS, C., and TERLAKY, T., *A Potential Reduction Variant of Renegar's Short-Step Path-Following Method for Linear Programming*, Linear Algebra and Its Applications, Vol. 84, pp. 43-68, 1991.

11. KARMARKAR, N., *Lecture Series on "Interior Point Methods for Discrete and Combinatorial Optimization"*, Indian Institute of Technology, Bombay, India, 1988.

12. MITCHELL, J., *Karmarkar's Algorithm and Combinatorial Optimization*, Ph.D Thesis, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 1988.

13. JARRE, F., *The Method of Analytic Centers for Smooth Convex Programs*, Ph.D Thesis, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, 1989.

14. MEHROTRA, S., and SUN, J., *An Interior Point Algorithm for Solving Smooth Convex Programs Based on Newton's Method*, Contempory Mathematics, Vol. 114, pp. 265-284, 1990.

15. DEN HERTOG, D., ROOS, C., and TERLAKY, T., *On the Classical Logarithmic Barrier Function Method for a Class of Smooth Convex Programming Problems*, Journal of Optimization Theory and Applications, Vol. 73, pp. 1-25, 1992.

16. DEN HERTOG, D., ROOS, C., and TERLAKY, T., *A Large-Step Analytic Center Method for a Class of Smooth Convex Programming Problems*, SIAM Journal on Optimization, Vol. 2, pp. 55-70, 1992.

17. IRI, M., and IMAI, H., *A Multiplicative Barrier Function Method for Linear Programming*, Algorithmica, Vol. 1, pp. 455-482, 1986.

18. ZHANG, S., and SHI, M., *On the Polynomiality of Iri and Imai's New Algorithm for Linear Programming*, Journal of Qinhua University (in Chinese), Vol. 28, pp. 121-126, 1988.

19. ZHANG, S., *On the Convergence Property of Iri and Imai's Method for Linear Programming*, Report 8917/A, Econometrics Institute, Erasmus University, Rotterdam, the Netherlands, 1989.

20. IMAI, H., *On the Polynomiality of the Time Complexity of Multiplicative Penalty Function Method for Linear Programming*, Working Paper, 1990.

21. IMAI, H., *On the Polynomiality of the Multiplicative Penalty Method for Linear Programming and Related Inscribed Ellipsoid*, IEICE Transactions, Vol. E74, pp. 669-671, 1991.

22. IRI, M., *A Proof of the Polynomiality of the Iri-Imai Method*, Paper Presented at the 14th International Symposium on Mathematical Programming, Amsterdam, the Netherlands, 1991.

23. AVRIEL, M., DIEWERT, W. E., SCHAIBLE, S., and ZANG, I., *Generalized Concavity*, Plenum Press, New York, New York, 1988.

24. BERTSEKAS, D. P., *Constrained Optimization and Lagrange Multiplier Methods,* Academic Press, New York, New York, 1982.

25. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.