A Value Estimation Approach to the Iri-Imai Method for Constrained Convex Optimization

Szewan Lam*, Duan Li†, Shuzhong Zhang‡

April 2003; revisions on October 2003, and March 2004

Abstract

In this paper, we propose an extension of the so-called Iri-Imai method to solve constrained convex programming problems. The original Iri-Imai method is designed for linear programs and assumes that the optimal objective value of the optimization problem is known in advance. Zhang [18] extends the method for constrained convex optimization, but the optimum value is still assumed to be known in advance. In our new extension this last requirement on the optimal value is relaxed; instead, only a lower bound of the optimal value is needed. Our approach uses a multiplicative barrier function for the problem with a univariate parameter that represents an estimated optimum value of the original optimization problem. An optimal solution to the original problem can be traced down by minimizing the multiplicative barrier function. Due to the convexity of this barrier function, the optimal objective value as well as the optimal solution of the original problem, are sought iteratively by applying Newton's method to the multiplicative barrier function. A new formulation of multiplicative barrier function is further developed to acquire computational tractability and efficiency. Numerical results are presented to show the efficiency of the new method.

Keywords: Constrained Convex Optimization, Iri-Imai's Algorithm, Value Estimation.

*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong.

†Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by Hong Kong RGC Earmarked Grant CUHK4214/01E.

‡Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by Hong Kong RGC Earmarked Grant CUHK4233/01E.

1 Introduction

Consider the following constrained convex global optimization problem:

$$
\begin{array}{lll}
(P) & \text{minimize} & f(x) \\
 & \text{subject to} & g_i(x) \leq 0,\ i = 1, 2, \ldots, m, \\
 & & x \in \Re^n,
\end{array}
$$

where $f : \Re^n \to \Re$, and $g_i : \Re^n \to \Re$, for $i = 1, 2, \ldots, m$, are twice continuously differentiable and convex functions.

In case $f(x)$ and $g_i(x)$, $i = 1, ..., m$, are all affine linear functions, (P) is a linear program, for which Iri and Imai [5] proposed an interior-point type algorithm to solve. The algorithm of Iri and Imai uses a multiplicative analogue of the potential function introduced by Karmarkar [6]. A number of research papers have been devoted to the Iri-Imai method. The focus of the research has been on improving the complexity of Iri-Imai's method for linear programming; see [4, 5, 9, 13, 14, 15, 17, 19]. Using a primal-dual potential reduction framework, the best complexity result for the Iri-Imai type algorithm for linear programs is $O(\sqrt{m}L)$ with L being the input-length of the linear program, which is also the best known complexity result for linear programs in general; see Sturm and Zhang [9]. Note that the primal-dual potential function was first introduced by Tanabe [11], and Todd and Ye [12]. Iri [4] is the first one to show that the Iri-Imai algorithm can be extended to solve convex quadratic programming with $O(m \log \frac{1}{\epsilon})$ iteration bound, where $\epsilon > 0$ is the required precision. Under the condition that the objective function f and the constraint functions g_i $(i = 1, ..., m)$ of (P) satisfy a certain type of convexity, termed the *harmonic convexity* in [18], Zhang [18] proves that the Iri-Imai algorithm for convex quadratic programming problems can be further extended to solve general convex programming problems (P) with complexity $O(m \log \frac{1}{\epsilon})$. Remark here that the harmonic convexity condition was also used in Mehrotra and Sun [8] in a different context. In the extended algorithm proposed in [18], the optimal objective value of the problem is still assumed to be known in advance, which causes practical cumbersomeness of the method.

The main contribution of the current paper is to extend the method of [18] by developing an efficient solution method for (P) via an optimal-value-estimation technique so that only a lower bound of the optimal value is needed. This type of value-estimation technique was first used in [10] and [16]. In our context, the proposed value-estimation function is an auxiliary unconstrained optimization problem with a univariate parameter that represents the estimated optimal value of (P). It turns out that the true optimal objective value of (P) is the unique root of the above mentioned value-estimation function. Therefore, the optimal objective value and the optimal solution can be searched in a two-level scheme that alternates between a root-finding phase at the upper level and an unconstrained optimization phase at the lower level. To achieve more computational tractability and efficiency, we further develop an approximate value-estimation function. We present numerical testing results for the proposed methods.

This paper is organized as follows. In Section 2, we introduce our value-estimation function. In Section 3, a smoothing approximation to the multiplicative barrier function is discussed, and the method is extended to the setting where there are both linear equality and convex inequality constraints. In Section 4, numerical results are presented and the paper is concluded.

2 The value-estimation approach

2.1 Formulation and properties

Consider the convex programming problem (P) as introduced in Section 1. To simplify the analysis, we make the following assumptions.

Assumption 1 Problem (P) has an optimal solution, and an initial lower bound of the optimum value is known, denoted by θ_0^l.

Assumption 2 The function f, and g_i for $i = 1, 2, \ldots, m$, are all convex and twice continuously differentiable. Moreover, f is not constant over the feasible region.

Assumption 3 The convex programming problem (P) satisfies the Slater condition; that is, there exists some $x \in \Re^n$ such that $g_i(x) < 0$ for $i = 1, 2, \ldots, m$.

Let the feasible set of (P) be

$$\mathcal{F} := \{x \mid g_i(x) \leq 0, 1 \leq i \leq m\} \subseteq \Re^n.$$

By Assumption 3, the interior of $\mathcal{F}$ is nonempty, denoted by

$$\overset{\circ}{\mathcal{F}} := \{x \mid g_i(x) < 0, 1 \leq i \leq m\} \subseteq \Re^n.$$

Assumption 4 The feasible set $\mathcal{F}$ of (P) is bounded; i.e., there is a constant M such that $\|x\| \leq M$ for any $x \in \mathcal{F}$.

We define a multiplicative barrier function Φ_θ for (P), and its logarithmic function, ϕ_θ, as follows:

$$\Phi_\theta(x) \quad := \quad \frac{(f(x)-\theta)^{m+h}}{\prod_{i=1}^m(-g_i(x))}, \text{ for } x \in \overset{\circ}{\mathcal{F}}, \tag{1}$$

$$\phi_\theta(x) \quad := \quad (m+h)\log(f(x)-\theta) - \sum_{i=1}^m \log(-g_i(x)), \text{ for } x \in \overset{\circ}{\mathcal{F}} \cap \{x \mid f(x) > \theta\}, \tag{2}$$

where $h > 1$ is some given positive integer, chosen in such a way that $m+h$ is an odd number, and θ is an estimation of the optimal value of (P).

Note that Φ_θ is well defined on the open and convex set $\overset{\circ}{\mathcal{F}}$. Moreover, we will see in Lemma 2.1 that $\Phi_\theta(x)$ has a nice convexity property on $\overset{\circ}{\mathcal{F}}$ under the following assumption.

Assumption 5 Problem (P) is assumed to satisfy one of the following two conditions: (i) at least one of the functions, f, g_i, $i = 1, 2, \ldots, m$, is strictly convex; (ii) $\text{rank}\{\nabla f(x), \nabla g_i(x) \mid i = 1, 2, \ldots, m\} = n$ for all $x \in \overset{\circ}{\mathcal{F}}$.

Denote by x^* the optimal solution and $\theta^* = f(x^*)$ the corresponding optimal objective value of (P).

Let us denote

$$\mathcal{R}_1^\theta := \overset{\circ}{\mathcal{F}} \cap \{x \mid f(x) \geq \theta\} \tag{3}$$

and

$$\mathcal{R}_2^\theta := \overset{\circ}{\mathcal{F}} \cap \{x \mid f(x) < \theta\}. \tag{4}$$

Lemma 2.1 *Suppose that $h > 1$ and Assumptions 3 and 5 hold. Then, the following statements hold true.*

(a) *If $\theta < \theta^*$, then the multiplicative barrier function $\Phi_\theta(x)$ of (P) is strictly convex on the open and convex set $\overset{\circ}{\mathcal{F}}$.*

(b) *If $\theta > \theta^*$, then $\nabla^2\Phi_\theta(x) \succ 0$ for $x \in \mathcal{R}_1^\theta$, and $\nabla^2\Phi_\theta(x) \prec 0$ for $x \in \mathcal{R}_2^\theta$.*

Proof. The convexity property in part **(a)** of the lemma actually follows immediately from Theorem 5.16 of [1].

The strict convexity proof was due to Iri and Imai [5]. For completeness we provide a proof below.

Note the following relationships for the functions (1) and (2):

$$\nabla^2\Phi_\theta(x) = \Phi_\theta(x)\left[\nabla^2\phi_\theta(x) + \nabla\phi_\theta(x)\nabla\phi_\theta(x)^{\mathrm{T}}\right] \tag{5}$$

$$\nabla\phi_\theta(x) = (m+h)\frac{\nabla f(x)}{f(x)-\theta} - \sum_{i=1}^{m}\frac{\nabla g_i(x)}{g_i(x)} \tag{6}$$

$$\begin{aligned}\nabla^2\phi_\theta(x) &= \frac{m+h}{f(x)-\theta}\nabla^2 f(x) - \frac{m+h}{[f(x)-\theta]^2}\nabla f(x)\nabla f(x)^{\mathrm{T}} \\ &\quad + \sum_{i=1}^{m}\frac{\nabla^2 g_i(x)}{-g_i(x)} + \sum_{i=1}^{m}\frac{\nabla g_i(x)\nabla g_i(x)^{\mathrm{T}}}{[g_i(x)]^2}.\end{aligned} \tag{7}$$

To prove assertion **(a)**, we need only to show the Hessian matrix $\nabla^2\Phi_\theta(x)$ is positive definite on $\mathring{\mathcal{F}}$, which is in turn equivalent to showing $\nabla^2\phi_\theta(x) + \nabla\phi_\theta(x)\nabla\phi_\theta(x)^{\mathrm{T}} \succ 0$ since $\Phi_\theta(x)$ is positive on $\mathring{\mathcal{F}}$.

Let us denote

$$M_2 := \frac{m+h}{f(x)-\theta}\nabla^2 f(x) + \sum_{i=1}^{m}\frac{\nabla^2 g_i(x)}{-g_i(x)} \succeq 0$$

and

$$M_1 := -\frac{m+h}{[f(x)-\theta]^2}\nabla f(x)\nabla f(x)^{\mathrm{T}} + \sum_{i=1}^{m}\frac{\nabla g_i(x)\nabla g_i(x)^{\mathrm{T}}}{[g_i(x)]^2} + \nabla\phi_\theta(x)\nabla\phi_\theta(x)^{\mathrm{T}}.$$

By (5), (6) and (7), we need only to show that $M_2 + M_1 = \nabla^2\phi_\theta(x) + \nabla\phi_\theta(x)\nabla\phi_\theta(x)^{\mathrm{T}} \succ 0$. To this end, we shall first prove $M_1 \succeq 0$.

Take any $\xi \in \Re^n$ and denote

$$\nu_0 = \left[\frac{\nabla f(x)}{f(x)-\theta}\right]^{\mathrm{T}}\xi, \text{ and } \nu_i = \left[\frac{\nabla g_i(x)}{g_i(x)}\right]^{\mathrm{T}}\xi \text{ for } i = 1, 2, \ldots, m.$$

Then

$$\xi^{\mathrm{T}} M_1 \xi = -(m+h)\nu_0^2 + \sum_{i=1}^{m}\nu_i^2 + \left[(m+h)\nu_0 - \sum_{i=1}^{m}\nu_i\right]^2. \tag{8}$$

By taking $\nu = (\nu_1, ..., \nu_m)^{\mathrm{T}}$ and $e = (1, ..., 1)^{\mathrm{T}}_{1\times m}$ equation (8) becomes

$$\xi^{\mathrm{T}} M_1 \xi = \left(\nu_0, \nu^{\mathrm{T}}\right) H \begin{pmatrix}\nu_0 \\ \nu\end{pmatrix} \tag{9}$$

where $H = \begin{pmatrix} (m+h)^2-(m+h), & -(m+h), \ldots, -(m+h) \\ -(m+h) & \\ \vdots & \mathrm{I} + ee^{\mathrm{T}} \\ -(m+h) & \end{pmatrix}_{(m+1)\times(m+1)}.$

Since $(m+h)^2-(m+h)>0$, we use Schur's lemma to conclude that $H \succ 0$ if and only if

$$\mathrm{I}+ee^{\mathrm{T}}-\frac{(m+h)^2}{(m+h)^2-(m+h)}\,ee^{\mathrm{T}}=\mathrm{I}-\frac{m+h}{(m+h)^2-(m+h)}\,ee^{\mathrm{T}}=\mathrm{I}-\frac{1}{m+h-1}\,ee^{\mathrm{T}}\succ 0.$$

However, the maximum (and the only nonzero) eigenvalue of the rank-one matrix $\frac{1}{m+h-1}\,ee^{\mathrm{T}}$ is $\frac{m}{m+h-1}$, which is less than 1 whenever $h>1$. This confirms that $H \succ 0$. Now, if condition (i) in Assumption 5 holds, then $M_2 \succ 0$. Otherwise, if condition (ii) holds, then for any $\xi \neq 0$ it follows that $(\nu_0, \nu^{\mathrm{T}})^{\mathrm{T}} \neq 0$. Hence in the latter case we must have $M_1 \succ 0$.

Part **(b)** of the lemma follows from the same matrix inequalities.

□

Note that $\mathcal{R}_1^\theta$ may not be a convex set in general; cf. (3) and (4). However, in the case of linear programming, i.e., when f and g_i's are affine linear function, then both $\mathcal{R}_1^\theta$ and $\mathcal{R}_2^\theta$ are half spaces. In that case, we conclude that $\Phi_\theta(x)$ is positive and convex in $\mathcal{R}_1^\theta$. Moreover, since $m+h$ is chosen to be an odd number, by replacing $f(x)$ with $-f(x)$ and θ with $-\theta$ for $x \in \mathcal{R}_2^\theta$, we can apply part **(b)** of Lemma 2.1 to the function $-\Phi_\theta(x)$, and this leads to the fact that $\Phi_\theta(x)$ is negative and concave in $\mathcal{R}_2^\theta$.

As a specialization, Lemma 2.1 leads to the following assertion.

Corollary 2.2 *Suppose that $f(\hat{x})>0$ and $\nabla^2 f(\hat{x}) \succeq 0$, and $g_i(\hat{x})<0$ and $\nabla^2 g_i(\hat{x}) \preceq 0$ for $i=1,...,m$. Moreover, suppose that $h>1$ and* $\operatorname{rank}\{\nabla f(\hat{x}), \nabla g_1(\hat{x}), \cdots, \nabla g_m(\hat{x})\}=n$*. Then,*

$$\nabla^2\left(\frac{(f(x))^{m+h}}{\prod_{i=1}^m(-g_i(x))}\right)\Bigg|_{x=\hat{x}} \succ 0.$$

This, however, further leads to the following result.

Proposition 2.3 *Suppose that $h>1$ and Assumption 3 holds. Moreover, suppose that*

$$\operatorname{rank}\{\nabla g_1(x), \cdots, \nabla g_m(x)\}=n \text{ for all } x \in \overset{\circ}{\mathcal{F}}.$$

Then

$$\chi(x,\theta):=\frac{(f(x)-\theta)^{m+h}}{\prod_{i=1}^m(-g_i(x))} \tag{10}$$

satisfies $\nabla^2_{(x,\theta),(x,\theta)}\chi(x,\theta) \succ 0$ for $(x,\theta) \in \overset{\circ}{\mathcal{F}} \times (-\infty,\theta^*)$.

Proof. We may view $f(x)-\theta$ and $g_i(x)$ as jointly convex functions for $(x,\theta) \in \overset{\circ}{\mathcal{F}} \times (-\infty, \theta^*)$. Hence we may apply Corollary 2.2 if its conditions are satisfied. In this context, we need only to particularly check that

$$\operatorname{rank}\left\{ \begin{pmatrix} \nabla f(x) \\ -1 \end{pmatrix}, \begin{pmatrix} \nabla g_1(x) \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} \nabla g_m(x) \\ 0 \end{pmatrix} \right\} = n+1 \tag{11}$$

for all $x \in \overset{\circ}{\mathcal{F}}$.

Now we prove by contradiction and assume that the above statement is not true, i.e., there exists $\bar{x} \in \overset{\circ}{\mathcal{F}}$ and $(\lambda^{\mathrm{T}}, \lambda_0) \neq [0^{\mathrm{T}}, 0]$ with $\lambda \in \Re^n$ such that

$$(\lambda^{\mathrm{T}}, \lambda_0) \left[\begin{pmatrix} \nabla f(\bar{x}) \\ -1 \end{pmatrix}, \begin{pmatrix} \nabla g_1(\bar{x}) \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} \nabla g_m(\bar{x}) \\ 0 \end{pmatrix} \right] = [0^{\mathrm{T}}, 0]$$

This implies that

$$\lambda^{\mathrm{T}}[\nabla g_1(\bar{x}), \cdots, \nabla g_m(\bar{x})] = 0 \text{ and } \lambda^{\mathrm{T}} \nabla f(\bar{x}) - \lambda_0 = 0.$$

Since $\operatorname{rank}\{\nabla g_1(\bar{x}), \cdots, \nabla g_m(\bar{x})\} = n$, we must have $\lambda = 0$ and consequently $\lambda_0 = 0$, which is in contradiction with the fact that $(\lambda^{\mathrm{T}}, \lambda_0) \neq [0^{\mathrm{T}}, 0]$. This means that the conditions of Corollary 2.2 are satisfied, and so its results apply. The proposition is proven. □

We now introduce a slightly stronger regularity condition to validate Proposition 2.3.

Assumption 6 Problem (P) is assumed to satisfy the following condition:

$$\operatorname{rank}\{\nabla g_1(x), \cdots, \nabla g_m(x)\} = n \text{ for all } x \in \overset{\circ}{\mathcal{F}}.$$

Obviously, Assumption 6 implies Assumption 5.

The key idea underlying the Iri-Imai method is to turn a constrained optimization problem into an unconstrained one, by virtue of the multiplicative barrier function, as the next lemma shows.

Lemma 2.4 *Suppose that Assumptions 3 and 4 holds. Then, for any sequence $\{x^k \mid k \geq 1\}$ with $x^k \in \overset{\circ}{\mathcal{F}}$, $k \geq 1$, satisfying $\lim_{k\to\infty} \Phi_{\theta^*}(x^k) = 0$ (or equivalently $\lim_{k\to\infty} \phi_{\theta^*}(x^k) = -\infty$), it follows that any cluster point of $\{x^k \mid k \geq 1\}$ is an optimal solution of (P).*

Proof. Since $\lim_{k\to\infty} \Phi_{\theta^*}(x^k) = 0$, by the definition of the multiplicative barrier function (1), and noting Assumption 4, it must hold that $\lim_{k\to\infty} f(x^k) = \theta^*$. Hence, the result follows. □

An important practical issue is to identify the value θ^*. For this purpose, let us consider the following auxiliary problem for (P):

$$\psi(\theta) := \inf_{x \in \overset{\circ}{\mathcal{F}}} \Phi_\theta(x) \tag{12}$$

where $\theta \in (-\infty, \theta^*]$. The function $\psi(\theta)$ is designed to provide a quality measure for the optimum value estimation θ. The following lemma shows nice properties of the function $\psi(\theta)$.

Lemma 2.5 *Suppose that Assumptions 1, 2, 3, 4 and 6 hold. Then, the value-estimation function $\psi(\theta)$ has the following properties:*

(a) *$\psi(\theta) > 0$ if and only if $\theta < \theta^*$;*

(b) *$\psi(\theta) = 0$ if and only if $\theta = \theta^*$;*

(c) *$\psi(\theta)$ is decreasing for $\theta \in (-\infty, \theta^*]$;*

(d) *$\psi(\theta)$ is a continuous function for $\theta \in (-\infty, \theta^*]$;*

(e) *$\psi(\theta)$ is strictly convex on $(-\infty, \theta^*]$.*

Proof.

(a) If $\theta < \theta^*$, then $\theta < \theta^* \le f(x)$ for all $x \in \overset{\circ}{\mathcal{F}}$. Hence, $f(x) - \theta \ge \theta^* - \theta > 0$ for all $x \in \overset{\circ}{\mathcal{F}}$. Moreover, for $x \in \overset{\circ}{\mathcal{F}}$, $-g_i(x) > 0$, for all $i = 1, 2, \ldots, m$. Noting Assumption 4 we have

$$\psi(\theta) = \inf_{x \in \overset{\circ}{\mathcal{F}}} \frac{(f(x) - \theta)^{m+h}}{\prod_{i=1}^m (-g_i(x))} > 0.$$

If $\psi(\theta) > 0$ then obviously we have $f(x) - \theta > 0$ for all $x \in \overset{\circ}{\mathcal{F}}$. Thus, $\theta \le \theta^*$. We now prove that $\theta \ne \theta^*$. Suppose for the sake of contradiction that $\theta = \theta^*$. Let x^* be an optimal solution of (P), which must exist due to Assumption 4. Moreover, because of Assumption 3, the optimal Lagrangian multipliers must exist at x^*. Let I be the set of active constraints at x^*, and $\lambda_i \ge 0$ be the Lagrangian multipliers such that

$$\nabla f(x^*) = -\sum_{i \in I} \lambda_i \nabla g_i(x^*).$$

Let $x^0 \in \overset{\circ}{\mathcal{F}}$. By Assumption 4, there is some constant $C \ge 0$ such that

$$f(x^* + t(x^0 - x^*)) - f(x^*) \le \nabla f(x^*)^T (x^0 - x^*) t + C \|x^0 - x^*\|^2 t^2$$

and

$$g_i(x^* + t(x^0 - x^*)) - g_i(x^*) \leq \nabla g_i(x^*)^T(x^0 - x^*)t + C\|x^0 - x^*\|^2 t^2, \tag{13}$$

for $t \in (0,1)$, and by convexity we also have

$$g_i(x^* + t(x^0 - x^*)) - g_i(x^*) \geq \nabla g_i(x^*)^T(x^0 - x^*)t$$

for $1 \leq i \leq m$. In particular, if $i \in I$ then $\nabla g_i(x^*)^T(x^0 - x^*) < 0$ and by (13),

$$-g_i(x^* + t(x^0 - x^*)) \geq -\nabla g_i(x^*)^T(x^0 - x^*)t - C\|x^0 - x^*\|^2 t^2.$$

This implies that if $t > 0$ is sufficiently small then we shall have

$$\begin{aligned}
&\frac{(f(x^* + t(x^0 - x^*)) - \theta^*)^{m+h}}{\prod_{i=1}^m(-g_i(x^* + t(x^0 - x^*)))} \\
= \;& O\left(\frac{(f(x^* + t(x^0 - x^*)) - f(x^*))^{m+h}}{\prod_{i\in I}(-g_i(x^* + t(x^0 - x^*)))}\right) \\
\leq \;& O\left(\frac{\left(\sum_{i\in I} -\lambda_i \nabla g_i(x^*)^T(x^0 - x^*)t + C\|x^0 - x^*\|^2 t^2\right)^{m+h}}{\prod_{i\in I}\left(-\nabla g_i(x^*)^T(x^0 - x^*)t - C\|x^0 - x^*\|^2 t^2\right)}\right) \\
= \;& O\left(\frac{\left(\sum_{i\in I} -\lambda_i \nabla g_i(x^*)^T(x^0 - x^*) + C\|x^0 - x^*\|^2 t\right)^{m+h}}{\prod_{i\in I}\left(-\nabla g_i(x^*)^T(x^0 - x^*) - C\|x^0 - x^*\|^2 t\right)} \times t^{m+h-|I|}\right).
\end{aligned}$$

The last term tends to zero as $t \downarrow 0$. Therefore,

$$\inf_{x \in \overset{\circ}{\mathcal{F}}} \frac{(f(x) - \theta)^{m+h}}{\prod_{i=1}^m(-g_i(x))} \leq 0.$$

This further implies that $\psi(\theta) \leq 0$, which is in contradiction with the condition $\psi(\theta) > 0$.

(b) Since $\psi(\theta) \geq 0$ for all $\theta \in (-\infty, \theta^*]$, the statement follows from **(a)**.

(c) By virtue of Lemma 2.1, we note that for any $\theta < \theta^*$, due to the barrier effect of $\Phi_\theta(x)$ and the fact that $\mathcal{F}$ is a compact set, there is a unique $x(\theta) \in \overset{\circ}{\mathcal{F}}$ such that $\Phi_\theta(x(\theta))$ is minimal over the domain $\overset{\circ}{\mathcal{F}}$. In fact, $x(\theta)$ coincides with the analytic central path (see also [18]); hence it is also a smooth curve.

Let $\theta_1 < \theta_2 < \theta^*$. Then, we have the following chain of inequalities

$$\begin{aligned}
f(x(\theta_1)) - \theta_2 &< f(x(\theta_1)) - \theta_1, \\
[f(x(\theta_1)) - \theta_2]^{m+h} &< [f(x(\theta_1)) - \theta_1]^{m+h},
\end{aligned}$$

and

$$\frac{[f(x(\theta_1))-\theta_2]^{m+h}}{\prod_{i=1}^m[-g_i(x(\theta_1))]}<\frac{[f(x(\theta_1))-\theta_1]^{m+h}}{\prod_{i=1}^m[-g_i(x(\theta_1))]}.$$

Moreover,

$$\frac{[(f(x(\theta_2))-\theta_2]^{m+h}}{\prod_{i=1}^m[-g_i(x(\theta_2))]}\quad\leq\quad\frac{[f(x(\theta_1))-\theta_2]^{m+h}}{\prod_{i=1}^m[-g_i(x(\theta_1))]}$$

by the definition of $x(\theta_2)$. Putting together these inequalities leads to $\psi(\theta_1)<\psi(\theta_2)$.

(d) This follows from the simple fact that if the cost function is continuous in some parameter (θ in this case), then the infimum value of the cost function over a domain is also continuous in the parameter; (see e.g. [2], Theorem 4.2.1).

(e) Let $\chi(x,\theta)=\Phi_\theta(x)$ (see (10)). By Proposition 2.3, the following property holds:

$$\nabla^2_{(x,\theta),(x,\theta)}\chi(x,\theta)=\begin{pmatrix}\nabla^2_{x,x}\chi(x,\theta) & \nabla^2_{x,\theta}\chi(x,\theta)\\ \nabla^2_{x,\theta}\chi(x,\theta)^{\mathrm{T}} & \nabla^2_{\theta,\theta}\chi(x,\theta)\end{pmatrix}\succ 0. \tag{14}$$

Moreover, we have

$$\psi(\theta)=\min_{x\in\overset{\circ}{\mathcal{F}}}\chi(x,\theta)=\chi(x(\theta),\theta). \tag{15}$$

Since $x(\theta)$ is the minimum point for $\Phi_\theta(x)$, we have

$$\nabla_x\chi(x(\theta),\theta)=0. \tag{16}$$

Differentiating (16) yields

$$\nabla^2_{x,x}\chi(x(\theta),\theta)^{\mathrm{T}}x'(\theta)+\nabla^2_{x,\theta}\chi(x(\theta),\theta)=0. \tag{17}$$

Now, consider (15) and differentiate the equation with respect to θ. This gives

$$\psi'(\theta)=\nabla_x\chi(x(\theta),\theta)^{\mathrm{T}}x'(\theta)+\nabla_\theta\chi(x(\theta),\theta)=\nabla_\theta\chi(x(\theta),\theta) \tag{18}$$

where in the second step we used (16). Further differentiating yields

$$\psi''(\theta)=\nabla^2_{x,\theta}\chi(x(\theta),\theta)^{\mathrm{T}}x'(\theta)+\nabla^2_{\theta,\theta}\chi(x(\theta),\theta). \tag{19}$$

Therefore,

$$\begin{aligned}
\psi''(\theta) &= ((x'(\theta))^{\mathrm{T}},1)\begin{pmatrix}0\\ \psi''(\theta)\end{pmatrix}\\
&= ((x'(\theta))^{\mathrm{T}},1)\begin{pmatrix}\nabla^2_{x,x}\chi(x(\theta),\theta)^{\mathrm{T}}x'(\theta)+\nabla^2_{x,\theta}\chi(x(\theta),\theta)\\ \nabla^2_{x,\theta}\chi(x(\theta),\theta)^{\mathrm{T}}x'(\theta)+\nabla^2_{\theta,\theta}\chi(x(\theta),\theta)\end{pmatrix}\\
&= ((x'(\theta))^{\mathrm{T}},1)\nabla^2_{(x,\theta),(x,\theta)}\chi(x(\theta),\theta)\begin{pmatrix}x'(\theta)\\ 1\end{pmatrix}\\
&> 0,
\end{aligned}$$

where in the second step we used (17) and (19), and in the last step we used (14). Hence, we proved that ψ is strictly convex on $\{\theta \mid \theta < \theta^*\}$. $\square$

2.2 The value estimation approach in the Iri-Imai method

The analysis so far suggests that $\psi(\theta) = 0$ if and only if $\theta = \theta^*$. That is, the optimal objective value of (P) is exactly the unique root of the function $\psi(\theta)$. Based on this fact, we propose the following *value-estimation* technique to modify the Iri-Imai algorithm. In simple terms, it searches for the optimal objective value and the optimal solution in a two-level framework that alternates between a root-finding phase and an unconstrained optimization phase.

Algorithm VE (Value Estimating Iri-Imai Method)

For this algorithm, the input includes the initial interior point $x^0 \in \overset{\circ}{\mathcal{F}}$, the initial lower bound of the optimal value θ_0^l and the precision parameter $\epsilon > 0$. The output consists of a sequence of solutions $x^k \in \overset{\circ}{\mathcal{F}}$, $k \geq 1$.

Step 0. Let $\theta := \theta_0^l$. Minimize the convex function $\Phi_\theta(x)$, (cf. (1)), using the Newton method with line-search starting from x^0. Return an optimal solution x^1. Set $\theta_1^l = \theta_0^l$ and $\theta_1^u = f(x^1)$. (Hence $\theta^* \in [\theta_1^l, \theta_1^u]$). Set $k := 1$.

Step 1. Let $\theta := (\theta_k^l + \theta_k^u)/2$. Minimize the function $\Phi_\theta(x)$ in the domain $\mathcal{R}_1^\theta$, (cf. (3)), using the Newton method with line-search starting from x^k. Return x^{k+1} either as the minimizer of $\Phi_\theta(\cdot)$ or when one detects that $x^{k+1} \notin \mathcal{R}_1^\theta$, namely when $f(x^{k+1}) < \theta$.

Step 2. *Termination Rule.*

If $\theta_k^u - \theta_k^l < \epsilon$ or $\Phi_\theta(x^{k+1}) = 0$, set $\theta^* = \theta$ and $x^* = x^{k+1}$, stop. Otherwise, go to **Step 3**.

Step 3. *Updating Rule.*

If $\Phi_\theta(x^{k+1}) > 0$, then set $\theta_{k+1}^l = \theta$, and furthermore if $f(x^{k+1}) < \theta_k^u$ then set $\theta_{k+1}^u = f(x^{k+1})$, else set $\theta_{k+1}^u = \theta_k^u$, $k := k+1$. Go to **Step 1**.

If $\Phi_\theta(x^{k+1}) < 0$, then set $\theta_{k+1}^l = \theta_k^l$ and $\theta_{k+1}^u = f(x^{k+1})$, $k := k+1$. Go to **Step 1**.

The following convergence result is immediate.

Theorem 2.6 *Suppose that Assumptions 1, 2, 3, 4 and 6 hold, and that we apply* **Algorithm VE** *to solve (P). Let* $\{x^k \mid k = 1, \cdots\}$ *be the generated iterative sequence with the bounds* $\{\theta_k^u \mid k = 1, \cdots\}$ *and* $\{\theta_k^l \mid k = 1, \cdots\}$. *Then, the Newton steps leading from* x^k *to* x^{k+1} *has a quadratic convergence rate, and the bounds* θ_k^u *and* θ_k^l *have the following property*

$$\theta_1^l \le \cdots \le \theta_k^l \le \theta^* \le \theta_k^u \le \cdots \le \theta_1^u \text{ with } \theta_k^u - \theta_k^l \le (\theta_1^u - \theta_1^l)/2^{k-1},$$

implying that

$$f(x^k) - \theta^* \le (\theta_1^u - \theta_1^l)/2^{k-1}$$

for any $k = 2, 3, \cdots$.

Proof. The quadratic convergence is due to the strong convexity of $\Phi_\theta(x)$ whenever θ is a lower bound; see Lemma 2.1. Certainly, if $\theta > \theta^*$, then the minimum value of $\Phi_\theta(x)$ in $\overset{\circ}{\mathcal{F}}$ is negative. Hence, it requires only finite number of Newton steps to reduce $\Phi_\theta(x)$ to a negative value. In other words, it requires only finite number of steps to move from x^k to x^{k+1}. The correctness of the lower and upper bounds $\{\theta_k^u \mid k = 1, \cdots\}$ and $\{\theta_k^l \mid k = 1, \cdots\}$ is due to the monotonicity of the value function established in Lemma 2.5, and last error estimation follows from the fact that $f(x^k) = \theta_k^u$ for all $k = 2, 3, \cdots$. □

As a remark, we mention that it is tempting to use the Newton method at the higher level as well (namely the zero-finding stage for $\psi(\theta)$). Note that the current approach can be viewed as a bisection method; hence only uses the monotonicity of $\psi(\theta)$. Our preliminary numerical experience shows, however, that one cannot achieve superlinear convergence at the higher level due to the degeneracy, i.e., $\psi'(\theta) \to 0$ as $\theta \uparrow \theta^*$.

3 Extensions

3.1 The plus function and its smoothing

One of the difficulties to be curbed is that the function Φ_θ is non-convex in the whole domain whenever θ is larger than the optimal objective value. This is reflected by the fact that during the execution of **Algorithm VE**, Step 1, one needs to constantly check whether x^{k+1} remains in $\mathcal{R}_1^\theta$ or not. In any case, the function ϕ_θ is not defined on $\mathcal{R}_2^\theta$. Since in our numerical procedure the function ϕ_θ is used exclusively, this is quite a nuisance. In addition to that, it is much more appealing, at least theoretically, to work with a *globally convex* function, even if we use the function Φ_θ instead of ϕ_θ. Note that the use of Φ_θ is not attractive numerically: it does cause overflows. In this section we shall propose a method to alleviate the problem. The idea is to truncate the original objective function for the part below the bound θ. This shall not affect the algorithm whenever θ is a lower bound of the optimal value. However, in case θ is not a lower bound, then the newly constructed multiplicative barrier function will not suffer from the non-convexity. A side effect is that the brutal truncation causes non-smoothness. Therefore we shall further use a smoothing method based on the Chen-Harker-Kanzow-Smale (cf. [3]) type approximation function.

Let $[y]_+$ be the plus function, i.e., $[y]_+ := \max\{y, 0\}$. For $x \in \overset{\circ}{\mathcal{F}}$, let us define

$$\Phi_\theta^+(x) := \frac{([f(x)-\theta]_+)^{m+h}}{\prod_{i=1}^m (-g_i(x))}, \tag{20}$$

where $h > 1$ is an integer, and θ is an estimation of the optimal value of (P). Specifically,

$$\begin{aligned} \Phi_\theta^+(x) &= \frac{\left(\frac{|f(x)-\theta|+f(x)-\theta}{2}\right)^{m+h}}{\prod_{i=1}^m (-g_i(x))} \\ &= \frac{(|f(x)-\theta|+f(x)-\theta)^{m+h}}{2^{m+h}\prod_{i=1}^m (-g_i(x))}. \end{aligned}$$

Note that the new multiplicative barrier function Φ_θ^+ is convex. However, Φ_θ^+ is not smooth at x with $f(x) = \theta$ when $\theta > \theta^*$ due to the non-smoothness of $[f(x) - \theta]_+$.

Let us further introduce one more parameter $\mu > 0$ and apply the Chen-Harker-Kanzow-Smale type approximation to the plus function; in particular, let us denote

$$[f(x)-\theta]_+^\mu := \frac{\sqrt{(f(x)-\theta)^2+\mu} + f(x) - \theta}{2}. \tag{21}$$

The following error estimation is well known:

$$\begin{aligned} 0 &< [f(x)-\theta]_+^\mu - [f(x)-\theta]_+ \\ &= \frac{\sqrt{(f(x)-\theta)^2+\mu} - |f(x)-\theta|}{2} \\ &= \frac{\mu}{2(\sqrt{(f(x)-\theta)^2+\mu} + |f(x)-\theta|)} \\ &\le \frac{\sqrt{\mu}}{2}. \end{aligned} \tag{22}$$

Consequently, for $\mu > 0$ and $x \in \overset{\circ}{\mathcal{F}}$ let us introduce the following smoothened multiplicative barrier function

$$\Phi_{\theta,\mu}^+(x) := \frac{([f(x)-\theta]_+^\mu)^{m+h}}{\prod_{i=1}^m (-g_i(x))}$$

$$= \frac{\left(\sqrt{(f(x)-\theta)^2+\mu}+f(x)-\theta\right)^{m+h}}{2^{m+h}\prod_{i=1}^m(-g_i(x))}.$$

The following property is immediate.

Lemma 3.1 *If $h > 1$ and Assumption 5 holds, then the multiplicative barrier functions Φ_θ^+ and $\Phi_{\theta,\mu}^+$ with $\mu > 0$ are convex and strictly convex respectively on the open and convex set $\overset{\circ}{\mathcal{F}}$ for any value θ.*

Proof. The convexity of Φ_θ^+ follows from Theorem 5.16 of [1], while the strict convexity of $\Phi_{\theta,\mu}^+$ can be shown in a similar way as in Lemma 2.1. □

Note that $\Phi_{\theta,\mu}^+(x) > 0$ for all $x \in \overset{\circ}{\mathcal{F}}$. Hence, we can define the logarithmic barrier function of $\Phi_{\theta,\mu}^+(x)$ as follows:

$$\begin{aligned}\phi_{\theta,\mu}^+(x) &:= \log \Phi_{\theta,\mu}^+(x) \\ &= (m+h)\log G_\mu^+(x) - \sum_{i=1}^m \log(-g_i(x)) - (m+h)\log 2 \end{aligned} \tag{23}$$

where $G_\mu^+(x) = [f(x)-\theta]_+^\mu = \sqrt{(f(x)-\theta)^2+\mu}+f(x)-\theta$.

For $\mu > 0$, we consider the following unconstrained global problem:

$$\psi_\mu^+(\theta) := \inf_{x\in\overset{\circ}{\mathcal{F}}} \Phi_{\theta,\mu}^+(x). \tag{24}$$

The function $\psi_\mu^+(\theta)$ is another value-estimation function. We intend to identify an approximate global optimal solution of (P) by solving (24) for sufficiently small $\mu > 0$. The unconstrained optimization problem (24) is computationally tractable due to the continuous differentiability of $\Phi_{\theta,\mu}^+(x)$. Moreover, Lemma 3.1 shows that problem (24) is an unconstrained convex optimization problem if $h > 1$ and Assumption 5 holds. Hence, we have the following modified Iri-Imai algorithm.

Algorithm VEA (Value Estimation and Approximation with Iri-Imai's Method)

The input of this algorithm includes the initial interior point $x^0 \in \overset{\circ}{\mathcal{F}}$, the initial lower bound of the optimal value θ_0^l, the smoothness parameter $\mu > 0$ and the precision parameter $\epsilon > 0$. The output consists of a sequence of solutions $x^k \in \overset{\circ}{\mathcal{F}}$, $k \geq 1$.

Step 0. Let $k := 1$. Solve the unconstrained global problem (24) by Newton's Method with line-search, where $\theta = \theta_0^l$, starting from $x = x^0$. Return an optimal solution x^1. Set $\theta_1^l = \theta_0^l$ and $\theta_1^u = f(x^1)$ such that $\theta^* \in [\theta_1^l, \theta_1^u]$.

Step 1. Solve (24) by Newton's Method with line-search, where $\theta = (\theta_k^l + \theta_k^u)/2$, starting from $x = x^k$. Return an optimal solution x^{k+1}.

Step 2. *Termination Rule.*

If $\theta_k^u - \theta_k^l < \epsilon$, then set $\theta^* = \theta$ and $x^* = x^{k+1}$, stop. Otherwise, go to **Step 3**.

Step 3. *Updating Rule.*

If $f(x^{k+1}) > \theta$, then set $\theta_{k+1}^l = \theta$, and furthermore if $f(x^{k+1}) < \theta_k^u$ then set $\theta_{k+1}^u = f(x^{k+1})$, else set $\theta_{k+1}^u = \theta_k^u$, $k := k+1$; go to **Step 1**.

If $f(x^{k+1}) \leq \theta$, then set $\theta_{k+1}^l = \theta_k^l$ and $\theta_{k+1}^u = f(x^{k+1})$, $k := k+1$; go to **Step 1**.

3.2 Linear equality constraints

Convex optimization problems are often naturally coupled with linear equality constraints. Although these constraints do not impose essential difficulties and can be eliminated by reducing the variables, it is likely to cause numerical complications in terms of spoiling the sparsity structure or inducing ill-conditioned computations. For this reason, we shall extend the idea of **Algorithm VE**, and **Algorithm VEA**, respectively, to directly solve convex constrained optimization problems with linear equality constraints:

$$\begin{array}{lll} (P^e) & \text{minimize} & f(x) \\ & \text{subject to} & g_i(x) \leq 0,\ i = 1, ..., m, \\ & & Ax = b \end{array}$$

where $A \in \Re^{m \times n}$ with rank m, $b \in \Re^m$, $f : \Re^n \to \Re$, and $g_i : \Re^n \to \Re$, for $i = 1, \ldots, m$, are twice continuously differentiable and convex functions.

Supposing that the optimal value of (P^e) is known in advance to be θ^*, then similar as before, one can show that (P^e) can be transformed into the following problem:

$$\begin{array}{ll} \text{minimize} & \Phi_{\theta^*}(x) \\ \text{subject to} & Ax = b \end{array}$$

where $\Phi_{\theta^*}(x) = \dfrac{(f(x) - \theta^*)^{m+h}}{\prod_{i=1}^m (-g_i(x))}$ is the multiplicative barrier function, and $h > 1$ is an integer such that $m + h$ is an odd number.

If only a strict lower bound θ of the optimal value θ^* is used in the multiplicative barrier function instead. The multiplicative barrier function

$$\Phi_\theta(x) = \frac{(f(x) - \theta)^{m+h}}{\prod_{i=1}^m (-g_i(x))} \tag{25}$$

is still strictly convex in $\overset{\circ}{\mathcal{F}}$. Hence, it will have the unique minimum point $x(\theta)$ in $\overset{\circ}{\mathcal{F}}$. As θ tends to the optimal objective value θ^*, the solution $x(\theta)$ of the minimization problem with objective function $\Phi_\theta(x)$ and constraints $Ax = b$ tends to the solution x^* of the original minimization problem (P^e). Therefore, the crux is to solve the problem indexed by the parameter θ in the the following form

$$\begin{array}{lll} (P^e_\theta) & \text{minimize} & \Phi_\theta(x) \\ & \text{subject to} & Ax = b \end{array}$$

for which the lower bound θ of the optimal objective value is updated each time until it reaches the true optimum value θ^*.

For a fixed parameter θ, given any point x with $x \in \overset{\circ}{\mathcal{F}}$ and $Ax = b$, we propose to find a projected Newton direction for (P^e_θ), i.e., the direction Δx solving the following quadratic problem

$$\begin{array}{ll} \text{minimize} & \Phi_\theta(x) + \nabla\Phi_\theta(x)^{\mathrm{T}}\Delta x + \dfrac{1}{2}\Delta x^{\mathrm{T}}\nabla^2\Phi_\theta(x)\Delta x \\ \text{subject to} & A\Delta x = 0. \end{array}$$

The above problem has an explicit and unique solution

$$\begin{aligned} \Delta x &= -\left[I - \nabla^2\Phi_\theta(x)^{-1}A^{\mathrm{T}}(A\nabla^2\Phi_\theta(x)^{-1}A^{\mathrm{T}})^{-1}A\right]\nabla^2\Phi_\theta(x)^{-1}\nabla\Phi_\theta(x) \\ &= \left[I - \nabla^2\Phi_\theta(x)^{-1}A^{\mathrm{T}}(A\nabla^2\Phi_\theta(x)^{-1}A^{\mathrm{T}})^{-1}A\right]d_N(x) \end{aligned} \tag{26}$$

where

$$d_N(x) = -\nabla^2\Phi_\theta(x)^{-1}\nabla\Phi_\theta(x).$$

In our implementation, we use the line-search technique such as the golden-section method to determine the optimal step-length along the projected Newton direction.

We can certainly use the smoothing approximation as introduced for the unconstrained case, i.e., we replace Φ_θ by $\Phi^+_{\theta,\mu}$ with a sufficiently small $\mu > 0$, resulting in the following problem indexed by the parameter θ for a fixed value μ:

$$\begin{array}{lll} (P^e_{\theta,\mu}) & \text{minimize} & \Phi_{\theta,\mu}(x) \\ & \text{subject to} & Ax = b. \end{array}$$

By the same argument, the corresponding projected Newton direction is

$$\Delta x = \left[I - \nabla^2\Phi^+_{\theta,\mu}(x)^{-1}A^{\mathrm{T}}(A\nabla^2\Phi^+_{\theta,\mu}(x)^{-1}A^{\mathrm{T}})^{-1}A\right]d_{N,\mu}(x) \tag{27}$$

where

$$d_{N,\mu}(x) = -\nabla^2\Phi^+_{\theta,\mu}(x)^{-1}\nabla\Phi^+_{\theta,\mu}(x).$$

3.3 Value estimation under linear constraints

As in the unconstrained case, if we define a value function

$$\tilde{\psi}(\theta) = \inf_{Ax=b;\, x\in\overset{\circ}{\mathcal{F}}} \Phi_\theta(x) \tag{28}$$

for $\theta \in (-\infty, \theta^*]$, then similarly we have the following property.

Lemma 3.2 *The value-estimation function $\tilde{\psi}(\theta)$ has the following properties:*

(a) $\tilde{\psi}(\theta) > 0$ *if and only if* $\theta < \theta^*$*;*

(b) $\tilde{\psi}(\theta) = 0$ *if and only if* $\theta = \theta^*$*;*

(c) $\tilde{\psi}(\theta)$ *is decreasing for* $\theta \in (-\infty, \theta^*]$*;*

(d) $\tilde{\psi}(\theta)$ *is a continuous function for* $\theta \in (-\infty, \theta^*]$*;*

(e) $\tilde{\psi}(\theta)$ *is strictly convex on* $(-\infty, \theta^*]$*.*

In the same vein, the value function for the approximated function can be introduced as follows, which has similar properties as stipulated in Lemma 3.2:

$$\tilde{\psi}_\mu^+(\theta) := \inf_{Ax=b;\, x\in\overset{\circ}{\mathcal{F}}} \Phi_{\theta,\,\mu}^+(x). \tag{29}$$

The extensions to the linear equality constraints of **Algorithm VE** and **Algorithm VEA** as discussed in Subsection 3.2 are called **Algorithm VEL** and **Algorithm VEAL** respectively, for references in the numerical tests. Certainly another straightforward way to eliminate the linear constraints would be to substitute the variables. However, these two approaches lead to different iterates.

4 Numerical results and conclusions

In order to evaluate the performance of our proposed methods, we implement our algorithms to solve a set of randomly generated convex quadratic feasible test-problems based on Lenard [7]. The numerical values of the multiplicative barrier functions Φ_θ and $\Phi_{\theta,\,\mu}^+$ can be very large and so numerical overflows are often encountered. Therefore, in our experimental results, we actually minimize the logarithmic barrier functions along the descent Newton's direction of the corresponding multiplicative barrier function, or the projection of the Newton direction. We compare the accuracy of the numerical results using different values of μ and conclude that $\mu = 10^{-10}$ works best for both **Algorithm VEA** and **Algorithm VEAL**. Those will be the results we present for these two algorithms.

The test problems are classified into four categories.

Type 1: *Convex Quadratic Programming Problems with Linear Inequality Constraints.*

The precise format is as follows:

$$
\begin{array}{lll}
(P_1) & \text{minimize} & x^{\mathrm{T}}Qx + p^{\mathrm{T}}x + r \\
 & \text{subject to} & Ax \leq b, \\
 & & x \in \Re^n,
\end{array}
$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$, and $Q \in \Re^{n \times n}$.

After applying **Algorithm VE** and **Algorithm VEA** respectively to 20 randomly generated test problems for each size by the method described in [7] with $\epsilon = 10^{-6}$, we get the overall numerical results which are summarized in Table 1 and Table 2, respectively.

Type 2: *Convex Quadratic Programming Problems with Convex Quadratic Inequality Constraints.*

The precise format is:

$$
\begin{array}{lll}
(P_2) & \text{minimize} & x^{\mathrm{T}}Q^0x + {p^0}^{\mathrm{T}}x + r^0 \\
 & \text{subject to} & x^{\mathrm{T}}Q^ix + {p^i}^{\mathrm{T}}x + r^i \leq 0, \text{ for } i = 1, 2, \ldots, m, \\
 & & x \in \Re^n,
\end{array}
$$

where $Q^i \in \Re^{n \times n}$, $p^i \in \Re^{n \times 1}$, and $r^i \in \Re$, for $i = 0, 1, \ldots, m$.

After applying **Algorithm VE** and **Algorithm VEA** respectively to 20 randomly generated test problems for each size by the method described in [7] with $\epsilon = 10^{-6}$, we get the overall numerical results which are summarized in Table 3 and Table 4, respectively.

Type 3: *Convex Quadratic Programming Problems with Linear Inequality and Equality Constraints.*

The format is

$$
\begin{array}{lll}
(P_3) & \text{minimize} & x^{\mathrm{T}}Qx + p^{\mathrm{T}}x + r \\
 & \text{subject to} & Ax \leq b \\
 & & Cx = d \\
 & & x \in \Re^n,
\end{array}
$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$, $C \in \Re^{l \times n}$, $d \in \Re^l$, $Q \in \Re^{n \times n}$, $p \in \Re^n$, and $r \in \Re$.

After applying **Algorithm VEL** and **Algorithm VEAL** respectively to 20 randomly generated test problems for each size by the method described in [7] with $\epsilon = 10^{-6}$, we get the overall numerical results which are summarized in Table 5 and Table 6, respectively.

Type 4: *Convex Quadratic Programming Problems with Convex Quadratic Inequality Constraints and Linear Equality Constraints.*

The format is as follows:

$$\begin{array}{lll} (P_4) & \text{minimize} & x^{\mathrm{T}}Q^0x + {p^0}^{\mathrm{T}}x + r^0 \\ & \text{subject to} & x^{\mathrm{T}}Q^ix + {p^i}^{\mathrm{T}}x + r^i \le 0, \text{ for } i = 1, 2, \ldots, m, \\ & & Ax = b, \\ & & x \in \Re^n, \end{array}$$

where $A \in \Re^{l \times n}$, $b \in \Re^l$, $Q^i \in \Re^{n \times n}$, $p^i \in \Re^n$, and $r^i \in \Re$, for $i = 0, 1, \ldots, m$.

After applying **Algorithm VEL** and **Algorithm VEAL** respectively to 20 randomly generated test problems for each size by the method described in [7] with $\epsilon = 10^{-6}$, we get the overall numerical results which are summarized in Table 7 and Table 8, respectively.

In the tables, the columns are partitioned into three parts, separated by double lines. The first part indicates the size of the random problem instances. The second part consists of two columns. The first column in the second part indicates the average number of times that the parameter θ is reduced to reach the desired precision. The column next to that is the standard deviation of the number of times to reduce θ. Note that the reduction of θ is not strictly bisection each time as the value of the objective function of the iterates are also used to update the upper and lower bounds, and so some fluctuation exists. The last part is to show the total number of Newton steps used. Again, the first column in that part is the average, and the second column is the standard deviation.

It appears that the number of times required to reduce θ (namely the upper level operation) is fairly stable, say around 10 to 20. It is remarkable that neither the size, nor the structure of the problem (different types of random problems generated), influences too much the number of iterations. In fact, the number of Newton steps required to re-center after the reduction of θ is typically 2 to 4, as it can be seen from the ratio between the second last column and the fourth last column. It is remarkable that the standard deviations are typically very small. We believe that it provides ground to conclude that the method is stable and reliable for convex quadratic functions. The method remains to be further tested for non-quadratic functions.

Size of A	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
201 × 100	10.850	0.366	29.950	2.540
401 × 200	10.800	0.410	29.800	1.361
601 × 300	11.400	0.680	32.900	1.224
801 × 400	11.000	0.649	34.000	1.026
1001 × 500	11.700	0.657	34.500	0.827
1201 × 600	10.850	0.366	31.950	3.120
1401 × 700	10.400	1.354	32.233	2.812
1601 × 800	10.550	0.511	33.400	1.354
1801 × 900	10.600	0.503	33.400	1.759
1801 × 950	10.400	0.503	34.050	2.089
1901 × 950	10.400	0.503	32.950	2.139
2001 × 1000	10.450	0.511	33.850	1.843

Table 1: Numerical results for QP with linear constraints (P_1) based on **Algorithm VE**

Size of A	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
201 × 100	16.100	0.912	35.350	1.461
401 × 200	16.300	0.924	34.550	1.317
601 × 300	16.300	0.801	33.700	1.780
801 × 400	15.450	1.146	33.800	2.505
1001 × 500	16.150	0.875	35.300	1.922
1201 × 600	15.450	1.146	35.200	1.853
1401 × 700	15.600	1.231	33.300	2.597
1601 × 800	15.400	1.392	32.700	2.250
1801 × 900	15.700	1.129	32.200	1.642
1801 × 950	16.400	1.046	32.900	2.075
1901 × 950	15.900	0.852	32.100	1.804
2001 × 1000	15.400	1.465	31.800	2.331

Table 2: Numerical results for QP with linear constraints (P_1) based on **Algorithm VEA**

m	n	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
201	100	9.800	0.523	30.300	1.657
401	200	9.000	0.795	30.450	2.373
601	300	8.700	0.571	30.650	2.059
801	400	8.300	0.470	30.150	1.565
1001	500	7.900	1.210	30.250	3.007
1201	600	8.250	0.716	31.400	1.429
1401	700	7.750	1.164	29.900	2.553
1601	800	7.950	0.686	30.350	2.796
1801	900	8.000	0.795	31.550	2.064
1801	950	7.900	0.308	30.750	1.916
1901	950	7.750	0.639	27.800	1.673
2001	1000	7.900	0.447	29.250	2.023

Table 3: Numerical results for QP with convex quadratic constraints (P_2) based on **Algorithm VE**

m	n	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
201	100	11.600	0.821	30.350	1.387
401	200	11.700	1.720	32.700	4.231
601	300	10.750	1.372	31.700	1.235
801	400	10.650	1.756	32.050	1.605
1001	500	11.150	2.183	32.350	1.756
1201	600	12.000	3.277	37.500	4.310
1401	700	11.800	2.931	39.800	5.085
1601	800	12.300	3.246	38.800	5.297
1801	900	10.600	3.378	35.200	2.118
1801	950	11.000	2.675	34.300	1.735
1901	950	12.200	3.888	36.200	2.840
2001	1000	11.200	2.587	34.300	3.063

Table 4: Numerical results for QP with convex quadratic constraints (P_2) based on **Algorithm VEA** ($\mu = 10^{-10}$)

Size of A	Size of C	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
151 × 100	50 × 100	16.400	0.503	41.500	0.827
301 × 200	100 × 200	16.300	0.657	40.100	0.968
451 × 300	150 × 300	16.600	0.503	40.500	0.688
601 × 400	200 × 400	16.200	0.410	40.000	0.918
751 × 500	250 × 500	16.050	0.686	36.000	1.622
901 × 600	300 × 600	16.100	0.308	36.800	0.616
1051 × 700	350 × 700	16.100	0.718	36.500	1.638
1201 × 800	400 × 800	16.150	0.587	36.450	1.317
1351 × 900	450 × 900	15.800	0.616	35.500	1.606
1351 × 950	450 × 950	16.300	0.470	36.550	1.099
1426 × 950	475 × 950	16.400	0.503	36.750	1.019
1501 × 1000	500 × 1000	16.350	0.489	36.150	1.348

Table 5: Numerical results for problem (P_3) based on **Algorithm VEL**

Size of A	Size of C	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
151 × 100	50 × 100	18.400	0.503	43.150	1.040
301 × 200	100 × 200	18.500	0.513	41.450	1.235
451 × 300	150 × 300	18.150	0.489	40.600	1.095
601 × 400	200 × 400	18.250	0.550	41.000	1.257
751 × 500	250 × 500	18.350	0.671	40.750	1.070
901 × 600	300 × 600	18.300	0.657	40.400	0.821
1051 × 700	350 × 700	18.200	0.616	40.100	1.166
1201 × 800	400 × 800	18.300	0.657	39.100	1.252
1351 × 900	450 × 900	18.200	0.616	38.900	0.968
1351 × 950	450 × 950	18.400	0.503	39.400	1.667
1426 × 950	475 × 950	18.400	0.503	39.200	1.005
1501 × 1000	500 × 1000	18.400	0.503	38.200	1.281

Table 6: Numerical results for problem (P_3) based on **Algorithm VEAL** ($\mu = 10^{-10}$)

m	Size of A	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
151	50 × 100	10.250	0.444	30.450	1.100
301	100 × 200	9.550	0.511	30.200	1.472
451	150 × 300	9.050	0.224	30.150	1.461
601	200 × 400	8.800	0.616	29.500	1.933
751	250 × 500	8.400	0.680	29.000	0.795
901	300 × 600	8.000	0.649	29.800	1.197
1051	350 × 700	8.300	0.801	29.900	1.334
1201	400 × 800	8.100	0.718	29.000	1.947
1351	450 × 900	7.500	1.147	28.900	2.89
1351	450 × 950	8.550	0.511	26.500	1.987
1426	475 × 950	8.600	0.598	27.900	1.971
1501	500 × 1000	8.500	0.513	26.600	1.314

Table 7: Numerical results for problem (P_4) based on **Algorithm VEL**

m	Size of A	updating θ (mean nr.)	updating θ (S.D.)	Newton steps (mean nr.)	Newton steps (S.D.)
151	50 × 100	12.550	0.511	31.600	1.142
301	100 × 200	11.700	0.470	31.000	1.522
451	150 × 300	11.350	0.587	31.100	1.619
601	200 × 400	11.200	0.696	31.050	1.701
751	250 × 500	10.750	0.444	30.400	0.994
901	300 × 600	10.800	0.410	30.000	1.451
1051	350 × 700	10.600	0.680	29.900	2.024
1201	400 × 800	10.800	0.410	30.400	1.231
1351	450 × 900	11.200	2.093	33.200	3.270
1351	450 × 950	10.800	0.410	29.600	1.142
1426	475 × 950	10.700	0.470	28.300	1.593
1501	500 × 1000	10.800	0.410	28.200	1.881

Table 8: Numerical results for problem (P_4) based on **Algorithm VEAL** ($\mu = 10^{-10}$)

Let us now summarize and conclude the paper.

In this paper, we relaxed the requirement of the original Iri-Imai method for convex programming problems [18] so that only a lower bound of the optimal objective value is needed. Note that in the conclusion part of Zhang [18], it was remarked that only a lower bound is needed, since the lower bound can be updated based on the local Hessian induced norm of the Newton direction. That approach resembles a short step path-following method, whereas the method in the current paper at least halves θ at each iteration, and so it is a long-step type path following method. It is known that the short-step method may have nice theoretical appeals, but it would not be efficient in practice, whereas the long-step variants usually are.

For our new value estimation method, we proved that the optimal objective value of the original optimization problem is the unique root of our proposed value-estimation function. Hence, the optimum value as well as the optimal solution can be searched in a two-level scheme that alternates between a root-finding phase by bisection method at the upper level and an unconstrained optimization phase by Newton's method at the lower level. Due to the convexity of $\psi(\theta)$ on $\{\theta \mid \theta \leq \theta^*\}$, the optimal objective value θ^* is not only the unique root of the function ψ but also the unique minimum point of that function. Hence, we may propose two other schemes to update the lower bound of the optimal value instead of the bisection method that we used. The first scheme is to update the value of θ by using Newton's method to solve $\psi(\theta) = 0$ in such a way that $\hat{\theta} = \theta - \frac{\psi(\theta)}{\psi'(\theta)}$ where $\hat{\theta}$ is the new lower bound of the optimal value. However, due to the degeneracy of the slope of the function ψ, this scheme is not efficient. The second scheme is to update the value of θ by using Newton's method to solve $\psi'(\theta) = 0$ so that $\hat{\theta} = \theta - \frac{\psi'(\theta)}{\psi''(\theta)}$ and $\hat{\theta} \leq \theta^*$. However, our experiments show that this approach still suffers from degeneracy. It remains to be studied how to efficiently exploit the convexity of $\psi(\theta)$ at the interval $\{\theta \mid \theta \leq \theta^*\}$. Based on our numerical experience so far, we conclude that the modified Iri-Imai algorithms, using bisection on θ, are already very stable, accurate and efficient.

References

[1] M. Avriel, W.E. Diewert, S. Schaible, and I. Zang, *Generalized Concavity*, Plenum Press, New York, 1988.

[2] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer, *Non-Linear Parametric Optimization*, Birkhauser Verlag, Basel, 1983.

[3] C. Chen and O.L. Mangasarian, *Smoothing methods for convex inequalities and linear complementarity problems*, Mathematical Programming 71: 51–69, 1995.

[4] M. Iri, *A proof of the polynomiality of the Iri-Imai method*, Journal of Complexity 9: 269–290, 1993.

[5] M. Iri and H. Imai, *A multiplicative barrier function method for linear programming*, Algorithmica 1: 455–482, 1986.

[6] N. Karmarkar, *A new polynomial-time algorithm for linear progarmming*, Combinatorica 4: 373–395, 1984.

[7] M.L. Lenard and M. Minkoff, *Randomly generated test problems for positive definite quadratic programming*, ACM Transactions on Mathematical Software 10: 86–96, 1984.

[8] S. Mehrotra and J. Sun, *An interior-point algorithm for solving smooth convex programs based on Newton's method*, Contemporary Mathematics 114: 265–284, 1990.

[9] J.F. Sturm and S. Zhang, *New complexity results for the Iri-Imai method*, Annals of Operations Research 62: 539–564, 1996.

[10] X.L. Sun and D. Li, *Value-estimation function method for constrained global optimization*, Journal of Optimization Theory and Applications 102: 385–409, 1999.

[11] K. Tanabe, *Centered Newton method for mathematical programming*, Proceedings of the 13th IFIP Conference, pages 197–206, 1987.

[12] M.J. Todd and Y. Ye, *A centered projective algorithm for linear programming*, Mathematics of Operations Research 15: 508–529, 1990.

[13] T. Tsuchiya, *Quadratic convergence of the Iri-Imai algorithm for degenerate linear porgramming problems*, Journal of Optimization Theory and Applications 87: 703–726, 1995.

[14] T. Tsuchiya and K. Tanabe, *Local convergence properties of new methods in linear programming*, Journal of the Operations Research Society of Japan 33: 22–45, 1990.

[15] R.H. Tutuncu, *A primal-dual variant of the Iri-Imai algorithm for linear progarmming*, Mathematics of Operations Research 25: 195–213, 2000.

[16] X.Q. Yang and D. Li, *Successive optimization mehod via parametric monotone composition formulation*, Journal of Global Optimization 16: 355–369, 2000.

[17] S. Zhang, *On the convergence property of Iri and Imai's method for linear progarmming*, Report 8917/A, Econometrics Institute, Erasmus Univeriesty, Rotterdam, The Netherlands, 1989.

[18] S. Zhang, *Convergence property of the Iri-Imai algorithm for some smooth convex programming problems*, Journal of Optimization Theory and Applications 82: 121–138, 1994.

[19] S. Zhang and M. Shi, *On the polynomiality of Iri and Imai's new algorithm for linear programming*, Journal of Qinhua University 28: 121–126, 1988.