



Continuous Optimization

A canonical dual approach for solving linearly constrained quadratic programs

Wenxun Xing^{a,*}, Shu-Cherng Fang^b, Ruey-Lin Sheu^c, Ziteng Wang^b^a Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China^b Department of Industrial and Systems Engineering, North Carolina State University, USA^c Department of Mathematics, National Cheng Kung University, Tainan, Taiwan

ARTICLE INFO

Article history:

Received 4 August 2010

Accepted 8 September 2011

Available online 17 September 2011

Keywords:

Quadratic programming

Global optimization

Canonical duality theory

ABSTRACT

This paper provides a canonical dual approach for minimizing a general quadratic function over a set of linear constraints. We first perturb the feasible domain by a quadratic constraint, and then solve a “restricted” canonical dual program of the perturbed problem at each iteration to generate a sequence of feasible solutions of the original problem. The generated sequence is proven to be convergent to a Karush–Kuhn–Tucker point with a strictly decreasing objective value. Some numerical results are provided to illustrate the proposed approach.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In this paper, we study the linearly constrained quadratic program (LCQP) of the following form:

$$\begin{aligned} Z_0 \equiv \quad & \text{Min} \quad p(x) = \frac{1}{2}x^T Qx + f^T x, \\ \text{s.t.} \quad & Ax \leq b, \end{aligned} \quad (1)$$

where x is an $n \times 1$ vector, Q an $n \times n$ real symmetric matrix, A an $m \times n$ real matrix, f an $n \times 1$ vector, and b an $m \times 1$ vector. We assume that problem (1) has a nonempty feasible domain and its objective function $p(x)$ is bounded from below over the feasible set.

Linearly constrained quadratic programming problems are commonly seen. A linear complementarity problem can be easily reformulated as such a problem [2]. When the Q matrix is indefinite, solving (LCQP) is in general NP-hard [13].

To find a global optimizer of problem (1), branch-and-bound techniques are often used. For instance, Serali and Tuncbilek [14] used the “reformulation convexification technique” to derive lower and upper bounds of the problem and partition the bounded polyhedral domain into subsets through branching. Barrientos and Correa [1] used the same branching idea but adopt the Lagrangian duality to provide lower bounds. Burer and Vandenbussche [3] used a semi-definite programming relaxation technique to provide bounds and enforce the first order Karush–Kuhn–Tucker (KKT) condition through branching. Notice that when one is contented with only a KKT point of problem (1), this goal can be achieved with less demanding computational efforts [10].

Recently, Gao [8,9] pioneered a canonical duality theory for nonconvex programming problems. The primal–dual relationship provides new insights to generating near-optimal solutions when the corresponding canonical dual problem is solved. Fang et al. [5,6] adopted this approach to tackle the 0–1 quadratic programming and sum-of-quadratic-ratios problems, Wang et al. [16] further extended the approach for handling multi-integer quadratic programming problems. Their computational results showed that the approach could be computationally efficient, but much more theoretical analysis is needed. The details of the canonical duality theory can be referred to [8,9,17].

In this paper, we develop an iterative scheme for solving linearly constrained quadratic programming problems based on the canonical duality theory. Starting from an arbitrarily selected initial point (may be infeasible to problem (1)), we first perturb problem (1) into a parameterized form, and then transform the perturbed problem to a canonical dual problem. This dual problem is further relaxed to a simple “restricted” convex quadratic programming problem that can be solved by using any known algorithms. Its optimal solution provides a corresponding feasible solution to problem (1). Using this feasible solution, a new restricted convex program then either provides a KKT point of problem (1) or generates a better feasible solution to problem (1) with a strictly less objective value such that the iterations of feasible points eventually converge to a KKT point of problem (1). Since the objective value keeps decreasing strictly at each iteration, the ending KKT point can be a good solution (local minimizer) to problem (1). Note that the very initial starting point can be arbitrarily re-selected, there is a chance for us to get a global minimizer of problem (1) by using this iterative scheme.

The rest of the paper is organized as follows. The properties of canonical duality programming are presented in Section 2. An

* Corresponding author.

E-mail addresses: wxing@math.tsinghua.edu.cn (W. Xing), fang@eos.ncsu.edu (S.-C. Fang), rsheu@mail.ncku.edu.tw (R.-L. Sheu), zwang23@ncsu.edu (Z. Wang).

iterative scheme and its convergence result are provided in Section 3. Numerical examples are shown in Section 4. Discussion and conclusion are given in Section 5.

2. Canonical duality

Let x_k be an arbitrary point in R^n and $\mu_k > 0$ be sufficiently large such that $\{x \in R^n | Ax \leq b\} \cap \{x \in R^n | \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\} \neq \emptyset$. Parameterized by (x_k, μ_k) , we consider the following perturbed problem of (1):

$$\begin{aligned} Z^p(x_k, \mu_k) \equiv \min_x \quad & p(x) = \frac{1}{2}x^T Qx + f^T x, \\ \text{s.t.} \quad & Ax \leq b, \\ & \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k. \end{aligned} \quad (2)$$

It is easy to see the next result.

Lemma 1. Let x_k be an arbitrary point in R^n and $\mu'_k, \mu''_k > 0$ be sufficiently large. If $\mu''_k \geq \mu'_k$, then $Z \leq Z^p(x_k, \mu'_k) \leq Z^p(x_k, \mu''_k)$.

For any $x \in \{x \in R^n | Ax \leq b\} \cap \{x \in R^n | \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\}$, we can change the variable x by $y = x - x_k$. Then the parameterized problem (2) becomes:

$$\begin{aligned} Z^p(x_k, \mu_k) \equiv \min_y \quad & p(x_k + y) = \frac{1}{2}y^T Qy + (f + Qx_k)^T y + p(x_k), \\ \text{s.t.} \quad & Ay \leq b - Ax_k, \\ & \frac{1}{2}y^T y \leq \mu_k. \end{aligned} \quad (3)$$

We shall refer to this problem as $(LCQP)_k$ in the remaining part of this paper.

Notice that if y_k^* is an optimal solution of problem (3), then $x_k + y_k^*$ must be a feasible solution to problem (1). Moreover, if x_k is feasible to problem (1), then $y = 0$ is feasible to problem (3). As a result, we have $p(x_k + y_k^*) \leq p(x_k)$, and hence y_k^* is a descent direction of the original objective function $p(x)$. We aim to use the canonical duality theory to find an approximate solution of y_k^* . We shall call it $\bar{y}_k$ and expect that $\bar{y}_k$ retains the above mentioned properties.

For this purpose, we begin with the Lagrangian function associated with problem (3), namely:

$$\begin{aligned} L_k(y, \sigma, \lambda) = \frac{1}{2}y^T(\lambda I + Q)y + (f + Qx_k + A^T \sigma)^T y - \sigma^T(b - Ax_k) \\ - \lambda \mu_k + p(x_k), \end{aligned} \quad (4)$$

where $(\sigma, \lambda) \in R_+^m \times R_+ = \{\sigma \geq 0 | \sigma \in R^m\} \times \{\lambda \geq 0 | \lambda \in R\}$. The key idea of the canonical duality theory [8,9] is to find all stationary points $y_k(\sigma, \lambda)$ of $L_k(y, \sigma, \lambda)$ such that $\frac{\partial L_k(y_k(\sigma, \lambda), \sigma, \lambda)}{\partial y} = 0$, assuming that $\lambda I + Q$ is invertible. Then the relationship between $y_k(\sigma, \lambda) \in R^n$ and $(\sigma, \lambda) \in R_+^m \times R_+$ is given by

$$y_k(\sigma, \lambda) = -(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma). \quad (5)$$

Substituting (5) into (4) results in the canonical dual function:

$$\begin{aligned} d_k(\sigma, \lambda) = & -\frac{1}{2}(f + Qx_k + A^T \sigma)^T(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma) \\ & - \sigma^T(b - Ax_k) - \lambda \mu_k + p(x_k) \\ = & -\frac{1}{2}\sigma^T A(\lambda I + Q)^{-1} A^T \sigma \\ & - [b - Ax_k + A(\lambda I + Q)^{-1}(f + Qx_k)]^T \sigma \\ & - \frac{1}{2}(f + Qx_k)^T(\lambda I + Q)^{-1}(f + Qx_k) - \lambda \mu_k + p(x_k). \end{aligned} \quad (6)$$

In general, the canonical dual function $d_k(\sigma, \lambda)$ is an “over-estimator” of the Lagrangian dual $\theta_k(\sigma, \lambda) \equiv \min_{y \in R^n} L_k(y, \sigma, \lambda)$. However, the two dual functions d_k and θ_k coincide on the set $\{(\sigma, \lambda) \geq 0 | \lambda I + Q \succ 0\}$, where “ $\succ 0$ ” means “positive definiteness.” More details of recent results along this line can be referred to [17].

Now, let us focus on the following canonical dual of problem (3):

$$\begin{aligned} Z^d(x_k, \mu_k) \equiv \sup \quad & d_k(\sigma, \lambda), \\ \text{s.t.} \quad & \lambda I + Q \succ 0, \\ & \lambda \geq 0, \sigma \geq 0. \end{aligned} \quad (7)$$

We shall refer to this problem as $(CDP)_k$ and denote its “interior solution set” by $\text{int}(\mathcal{F}) = \{(\sigma, \lambda) \succ 0 | \lambda I + Q \succ 0\}$ with a closure $\bar{\mathcal{F}} = \{(\sigma, \lambda) \geq 0 | \lambda I + Q \succeq 0\}$. It is important to notice that the dual feasible domain of $(CDP)_k$ and the sets $\text{int}(\mathcal{F})$ and $\bar{\mathcal{F}}$ are independent of the parameter k . Below is an analytic property of $d_k(\sigma, \lambda)$.

Lemma 2. $d_k(\sigma, \lambda)$ is a continuous concave function over $\text{int}(\mathcal{F})$. For any $(\sigma, \lambda) \in \text{int}(\mathcal{F})$:

$$\frac{\partial d_k(\sigma, \lambda)}{\partial \sigma} = -A(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma) - (b - Ax_k), \quad (8)$$

$$\frac{\partial d_k(\sigma, \lambda)}{\partial \lambda} = \frac{1}{2}(f + Qx_k + A^T \sigma)^T(\lambda I + Q)^{-2}(f + Qx_k + A^T \sigma) - \mu_k, \quad (9)$$

$$\begin{aligned} \frac{\partial^2 d_k(\sigma, \lambda)}{\partial(\sigma, \lambda)^2} = & -[A^T, -(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma)]^T(\lambda I + Q)^{-1} \\ & \times [A^T, -(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma)], \end{aligned} \quad (10)$$

where $\frac{\partial^2 d_k(\sigma, \lambda)}{\partial(\sigma, \lambda)^2}$ represents the Hessian matrix of $d_k(\sigma, \lambda)$.

Proof. Expressions (8) and (9) can be directly verified. To compute the Hessian matrix of $d_k(\sigma, \lambda)$, we have:

$$\frac{\partial^2 d_k(\sigma, \lambda)}{\partial \sigma \partial \sigma} = -A(\lambda I + Q)^{-1} A^T, \quad (m \times m \text{ matrix}),$$

$$\begin{aligned} \frac{\partial^2 d_k(\sigma, \lambda)}{\partial \sigma \partial \lambda} &= \left(\frac{\partial^2 d_k(\sigma, \lambda)}{\partial \lambda \partial \sigma} \right)^T \\ &= A(\lambda I + Q)^{-2}(f + Qx_k + A^T \sigma), \quad (m \times 1 \text{ vector}), \end{aligned}$$

and

$$\frac{\partial^2 d_k(\sigma, \lambda)}{\partial \lambda \partial \lambda} = -(f + Qx_k + A^T \sigma)^T(\lambda I + Q)^{-3}(f + Qx_k + A^T \sigma).$$

Consequently:

$$\begin{aligned} \frac{\partial^2 d_k(\sigma, \lambda)}{\partial(\sigma, \lambda)^2} &= \begin{bmatrix} \frac{\partial^2 d_k(\sigma, \lambda)}{\partial \sigma \partial \sigma} & \frac{\partial^2 d_k(\sigma, \lambda)}{\partial \sigma \partial \lambda} \\ \frac{\partial^2 d_k(\sigma, \lambda)}{\partial \lambda \partial \sigma} & \frac{\partial^2 d_k(\sigma, \lambda)}{\partial \lambda \partial \lambda} \end{bmatrix} \\ &= -[A^T, -(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma)]^T \\ &\quad \times (\lambda I + Q)^{-1}[A^T, -(\lambda I + Q)^{-1}(f + Qx_k + A^T \sigma)]. \end{aligned}$$

Since $(\lambda I + Q)$ is positive definite over $\text{int}(\mathcal{F})$, $\frac{\partial^2 d_k(\sigma, \lambda)}{\partial(\sigma, \lambda)^2}$ must be negative semi-definite. This shows the concavity of $d_k(\sigma, \lambda)$. $\square$

Lemma 2 implies that the canonical dual problem $(CDP)_k$ in (7) is a concave optimization problem. However, the inversion operation of $(\lambda I + Q)^{-1}$ in $d_k(\sigma, \lambda)$ could be numerically unstable, especially when λ approaches the maximum eigenvalue of $-Q$. Let the minimum eigenvalue of the matrix Q be λ_0 . Our strategy is to pre-select a $\bar{\lambda} > \max\{0, -\lambda_0\}$, fix it through all iterations, and

approximate $(CDP)_k$ by the following “restricted” canonical dual problem:

$$Z^{rd}(x_k, \mu_k, \bar{\lambda}) \equiv \sup_{\sigma \geq 0} d_k(\sigma, \bar{\lambda}), \quad (11)$$

We shall refer to this problem as $(RCDP)_k$. Since the choice of $\bar{\lambda} > 0$ may affect the results of $(RCDP)_k$, we shall discuss this issue later in some more details.

Notice that $(RCDP)_k$ is a simple convex quadratic programming problem that can be solved efficiently by many existing methods including the active set method [7] and conic programming interior point algorithms [4,12]. The relation between $(LCQP)_k$ and $(RCDP)_k$ is given below.

Lemma 3 (Weak duality). For any given $x_k \in R^n$, $\mu_k > 0$ being sufficiently large such that $\{x \in R^n | Ax \leq b\} \cap \{x \in R^n | \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\} \neq \emptyset$, and $\bar{\lambda} > \max\{0, -\lambda_0\}$ being fixed, we have $Z^{rd}(x_k, \mu_k, \bar{\lambda}) \leq Z^p(x_k, \mu_k)$ and the problem $(RCDP)_k$ has a finite maximizer.

Proof. Since $\bar{\lambda}I + Q \succ 0$, $y_k(\sigma, \bar{\lambda}) = -(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\sigma)$ is the only global minimizer of the Lagrangian function $L_k(y, \sigma, \bar{\lambda})$. Then:

$$\begin{aligned} d_k(\sigma, \bar{\lambda}) &= L_k(y_k(\sigma, \bar{\lambda}), \sigma, \bar{\lambda}) = \min_{y \in R^n} L_k(y, \sigma, \bar{\lambda}) \\ &\leq \min_{\{y | Ay \leq b - Ax_k - \frac{1}{2}y^T y \leq \mu_k\}} L_k(y, \sigma, \bar{\lambda}) \\ &\leq \min_{\{y | Ay \leq b - Ax_k - \frac{1}{2}y^T y \leq \mu_k\}} \frac{1}{2}y^T Qy + (f + Qx_k)^T y \\ &\quad + p(x_k)Z^p(x_k, \mu_k), \end{aligned} \quad (12)$$

for any $\sigma \geq 0$. Consequently:

$$Z^{rd}(x_k, \mu_k, \bar{\lambda}) = \sup_{\sigma \geq 0} d_k(\sigma, \bar{\lambda}) \leq Z^p(x_k, \mu_k).$$

Noticing that problem $(LCQP)_k$ has a bounded feasible domain, we know $Z^p(x_k, \mu_k) < +\infty$. Hence the concave quadratic problem $(RCDP)_k$ is bounded above with a feasible solution $\sigma = 0$. This further implies that $(RCDP)_k$ has a finite maximizer [7]. $\square$

Let $\bar{\sigma}^k$ be a finite maximizer of $(RCDP)_k$. In the following theorem we show that $\bar{\sigma}^k$ in the canonically dual space can be converted to a good moving direction $\bar{y}_k$ in the primal feasible domain of $(LCQP)$.

Theorem 1. Let $x_k \in R^n$, $\mu_k > 0$ be sufficiently large such that $\{x \in R^n | Ax \leq b\} \cap \{x \in R^n | \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\} \neq \emptyset$, and $\bar{\lambda} > \max\{0, -\lambda_0\}$. If $\bar{\sigma}^k$ is a finite maximizer of $(RCDP)_k$, then:

$$\frac{\partial d_k(\bar{\sigma}^k, \bar{\lambda})}{\partial \sigma_i} \leq 0, \quad (13)$$

and

$$\bar{\sigma}_i^k \frac{\partial d_k(\bar{\sigma}^k, \bar{\lambda})}{\partial \sigma_i} = 0, \quad (14)$$

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

Moreover, the dual-to-primal conversion:

$$\bar{y}_k = -(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\bar{\sigma}^k), \quad (15)$$

leads the vector $x_{k+1} = x_k + \bar{y}_k$ to be a feasible solution of $(LCQP)$.

If $\frac{1}{2}\bar{y}_k^T \bar{y}_k \leq \mu_k$, then we have:

$$\begin{aligned} 0 &\leq p(x_{k+1}) - Z^p(x_k, \mu_k) \leq p(x_{k+1}) - Z^{rd}(x_k, \mu_k, \bar{\lambda}) \\ &= \bar{\lambda} \left(\mu_k - \frac{1}{2}\bar{y}_k^T \bar{y}_k \right). \end{aligned} \quad (16)$$

On the other hand, if $\frac{1}{2}\bar{y}_k^T \bar{y}_k > \mu_k$, then $\bar{y}_k^T$ is a global minimizer of problem $(LCQP)_k$ in which μ_k is replaced by $\bar{\mu}_k = \frac{1}{2}\bar{y}_k^T \bar{y}_k$.

Proof. We first prove expressions (13) and (14) using the KKT condition of $(RCDP)_k$. Note that the associated Lagrangian function is

$$L(\sigma, \tau) = d_k(\sigma, \bar{\lambda}) + \tau^T \sigma,$$

where $\tau \in R_m^+$. Since $\bar{\sigma}^k \geq 0$ is a finite maximizer of $(RCDP)_k$, there associates a multiplier $\bar{\tau} \geq 0$ such that:

$$\frac{\partial d_k(\bar{\sigma}^k, \bar{\lambda})}{\partial \sigma_i} + \bar{\tau}_i = 0 \quad \text{for } i = 1, 2, \dots, m,$$

and

$$\bar{\tau}^T \bar{\sigma}^k = 0.$$

Therefore, we have:

$$\frac{\partial d_k(\bar{\sigma}^k, \bar{\lambda})}{\partial \sigma_i} = -\bar{\tau}_i \leq 0,$$

and

$$\bar{\sigma}_i^k \frac{\partial d_k(\bar{\sigma}^k, \bar{\lambda})}{\partial \sigma_i} = 0,$$

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

With expressions (8) and (13) and the dual-to-primal conversion (15), we can easily see that $A(x_k + \bar{y}_k) = Ax_{k+1} \leq b$, i.e., x_{k+1} is feasible to $(LCQP)$. Also, from (8) and (14), we have:

$$(\bar{\sigma}^k)^T(b - Ax_k) = (\bar{\sigma}^k)^T A \bar{y}_k. \quad (17)$$

Together with (15) and the first expression in (6), we have:

$$\begin{aligned} Z^{rd}(x_k, \mu_k, \bar{\lambda}) &= d_k(\bar{\sigma}^k, \bar{\lambda}) \\ &= \frac{1}{2}(f + Qx_k + A^T\bar{\sigma}^k)^T \bar{y}_k - (\bar{\sigma}^k)^T A \bar{y}_k - \bar{\lambda}\mu_k + p(x_k) \\ &= -\frac{1}{2}(f + Qx_k + A^T\bar{\sigma}^k)^T \bar{y}_k + (f + Qx_k)^T \bar{y}_k - \bar{\lambda}\mu_k + p(x_k) \\ &= \frac{1}{2}\bar{y}_k^T(\bar{\lambda}I + Q)\bar{y}_k + (f + Qx_k)^T \bar{y}_k - \bar{\lambda}\mu_k + p(x_k) \\ &= p(x_k + \bar{y}_k) + \bar{\lambda} \left(\frac{1}{2}\bar{y}_k^T \bar{y}_k - \mu_k \right). \end{aligned} \quad (18)$$

When $\frac{1}{2}\bar{y}_k^T \bar{y}_k < \mu_k$, $\bar{y}_k$ becomes feasible to $(LCQP)_k$. Based on the weak duality in Lemma 3, we have:

$$\begin{aligned} Z^p(x_k, \mu_k) &\geq \sup_{\sigma \geq 0} d_k(\sigma, \bar{\lambda}) = Z^{rd}(x_k, \mu_k, \bar{\lambda}) \\ &= p(x_{k+1}) - \bar{\lambda} \left(\mu_k - \frac{1}{2}\bar{y}_k^T \bar{y}_k \right). \end{aligned}$$

This implies that (i) x_{k+1} is feasible to $Ax \leq b$, (ii) x_{k+1} lies inside the ball of $\{x \in R^n | \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\}$, and (iii) the primal objective value $p(x_{k+1})$ has a positive gap $\bar{\lambda}(\mu_k - \frac{1}{2}\bar{y}_k^T \bar{y}_k)$ with the dual estimation $Z^{rd}(x_k, \mu_k, \bar{\lambda})$.

When $\frac{1}{2}\bar{y}_k^T \bar{y}_k = \mu_k$, the point x_{k+1} is feasible to $Ax \leq b$ and it lies on the boundary of the ball $\{x \in R^n | \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\}$. Moreover, the gap $\bar{\lambda}(\mu_k - \frac{1}{2}\bar{y}_k^T \bar{y}_k)$ vanishes in this case and $\bar{y}_k$ becomes a global minimizer of problem $(LCQP)_k$.

When $\frac{1}{2}\bar{y}_k^T \bar{y}_k > \mu_k$, the point x_{k+1} lies outside of the ball, but it is still feasible to $Ax \leq b$. Since $\bar{y}_k$ is not feasible to $(LCQP)_k$, there is no weak duality between $(LCQP)_k$ and $(RCDP)_k$. Indeed, the optimal

value $Z^{rd}(x_k, \mu_k, \bar{\lambda})$ of the restricted canonical dual problem exceeds $p(x_{k+1})$ by the amount of $\bar{\lambda}(\frac{1}{2}\bar{y}_k^T\bar{y}_k - \mu_k)$. This shows expression (16).

Finally, if we fix x_k and let μ_k vary, then expression (6) indicates that this change only causes a vertical shift in $d_k(\sigma, \bar{\lambda})$, but not affects the essence of $(\text{RCDP})_k$. Therefore, the maximizer $\bar{\sigma}^k$ of $(\text{RCDP})_k$, the moving direction $\bar{y}_k$ and the next iterative x_{k+1} all remain the same as μ_k varies when x_k is fixed. Consequently, if we increase μ_k to match $\frac{1}{2}\bar{y}_k^T\bar{y}_k$ such that $\bar{y}_k$ lies on the boundary of the enlarged ball, then $\bar{y}_k$ becomes a global minimizer of problem $(\text{LCQP})_k$ in which μ_k is replaced by $\bar{\mu}_k = \frac{1}{2}\bar{y}_k^T\bar{y}_k$. $\square$

Notice that Theorem 1 asserts that x_{k+1} is feasible to (LCQP) regardless the status of x_k . Even if x_k is infeasible, the next iterate x_{k+1} must be feasible all the time.

3. An iterative scheme

Once we get a feasible solution of problem (1), the next two results form the basis of our proposed iterative scheme.

Lemma 4. Let x_k be feasible to (LCQP) in form of problem (1) and $\bar{y}_k$ be defined by (15) with $\bar{\sigma}^k$ being a maximizer of problem $(\text{RCDP})_k$. Then, x_k is a KKT point of problem (1) if and only if $\bar{y}_k = 0$.

Proof. Following Theorem 1, we have:

$$\begin{cases} \bar{y}_k = -(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\bar{\sigma}^k), \\ A\bar{y}_k \leq b - Ax_k, \\ (\bar{\sigma}^k)^T(A\bar{y}_k - b + Ax_k) = 0, \\ \bar{\sigma}^k \geq 0. \end{cases} \quad (19)$$

If $\bar{y}_k = 0$, the positive definiteness of $\bar{\lambda}I + Q$ implies that:

$$\begin{cases} f + Qx_k + A^T\bar{\sigma}^k = 0, \\ Ax_k \leq b, \\ (\bar{\sigma}^k)^T(Ax_k - b) = 0, \\ \bar{\sigma}^k \geq 0. \end{cases} \quad (20)$$

Thus x_k is a KKT point of problem (1).

On the other hand, if x_k is a KKT point of problem (1), then there exists a $\sigma' \geq 0$ such that:

$$\begin{cases} f + Qx_k + A^T\sigma' = 0, \\ Ax_k \leq b, \\ \sigma'^T(Ax_k - b) = 0, \\ \sigma' \geq 0. \end{cases} \quad (21)$$

From the canonical dual function defined by the KKT point x_k , we have:

$$\begin{aligned} d_k(\sigma, \lambda) &= -\frac{1}{2}(f + Qx_k + A^T\sigma)^T(\lambda I + Q)^{-1}(f + Qx_k + A^T\sigma) \\ &\quad - \sigma^T(b - Ax_k) - \lambda\mu_k + p(x_k) \\ &\leq -\lambda\mu_k + p(x_k) = d_k(\sigma', \lambda), \end{aligned} \quad (22)$$

where the inequality holds due to the facts that $\frac{1}{2}(f + Qx_k + A^T\sigma)^T(\lambda I + Q)^{-1}(f + Qx_k + A^T\sigma) \geq 0$ and $\sigma^T(b - Ax_k) \geq 0$ while the last equality holds due to the fact that (x_k, σ') is a KKT pair satisfying (21). In other words, for all $\lambda > \max\{0, -\lambda_0\}$, σ' is a common global maximizer of $d_k(\sigma, \lambda)$ over $\{\sigma \geq 0\}$. Notice that the optimal value of $d_k(\sigma, \lambda)$ over $\{\sigma \geq 0\}$ is given by $p(x_k) - \lambda\mu_k$, which is linearly decreasing in λ . In particular, when x_k is a KKT point of (LCQP) in form of problem (1) and $\lambda = \bar{\lambda}$, we have:

$$\begin{aligned} d_k(\sigma', \bar{\lambda}) &= d_k(\bar{\sigma}^k, \bar{\lambda}) \\ &= -\frac{1}{2}(f + Qx_k + A^T\bar{\sigma}^k)^T(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\bar{\sigma}^k) \\ &\quad - (\bar{\sigma}^k)^T(b - Ax_k) - \bar{\lambda}\mu_k + p(x_k) \\ &= -\bar{\lambda}\mu_k + p(x_k). \end{aligned} \quad (23)$$

Since $-\frac{1}{2}(f + Qx_k + A^T\bar{\sigma}^k)^T(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\bar{\sigma}^k) \leq 0$ and $-(\bar{\sigma}^k)^T(b - Ax_k) \leq 0$, expression (23) is made possible only when both terms are 0. This further implies that $f + Qx_k + A^T\bar{\sigma}^k = 0$ and consequently, $\bar{y}_k = 0$. $\square$

Lemma 5. Let x_k be a feasible solution of problem (1), $\bar{y}_k$ be defined by (15), and $x_{k+1} = x_k + \bar{y}_k$. Then:

$$p(x_{k+1}) \leq p(x_k) - \frac{1}{2}\bar{y}_k^T(\bar{\lambda}I + Q)\bar{y}_k - \frac{\bar{\lambda}}{2}\bar{y}_k^T\bar{y}_k. \quad (24)$$

Proof. Noticing that $\bar{\sigma}^k \geq 0$ and $b - Ax_k \geq 0$, from (18) and (6), we have:

$$\begin{aligned} p(x_{k+1}) &= Z^{rd}(x_k, \mu_k, \bar{\lambda}) - \bar{\lambda}\left(\frac{1}{2}\bar{y}_k^T\bar{y}_k - \mu_k\right) \\ &= d_k(\bar{\sigma}^k, \bar{\lambda}) - \bar{\lambda}\left(\frac{1}{2}\bar{y}_k^T\bar{y}_k - \mu_k\right) \\ &= -\frac{1}{2}(f + Qx_k + A^T\bar{\sigma}^k)^T(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\bar{\sigma}^k) \\ &\quad - (\bar{\sigma}^k)^T(b - Ax_k) - \bar{\lambda}\mu_k + p(x_k) - \bar{\lambda}\left(\frac{1}{2}\bar{y}_k^T\bar{y}_k - \mu_k\right) \\ &= -\frac{1}{2}\bar{y}_k^T(\bar{\lambda}I + Q)\bar{y}_k - (\bar{\sigma}^k)^T(b - Ax_k) - \frac{\bar{\lambda}}{2}\bar{y}_k^T\bar{y}_k + p(x_k) \\ &\leq -\frac{1}{2}\bar{y}_k^T(\bar{\lambda}I + Q)\bar{y}_k - \frac{\bar{\lambda}}{2}\bar{y}_k^T\bar{y}_k + p(x_k). \quad \square \end{aligned}$$

The above lemma says that when x_k is feasible to problem (1), then $x_{k+1} = x_k + \bar{y}_k$ is either a KKT point of problem (1) or a feasible solution of problem (1) with a lower objective value.

We now present an iterative canonical-duality-based algorithm (ICDA) as follows.

Step 0: Select $\bar{\lambda} > \max\{0, -\lambda_0\}$, $x_0 \in \mathbb{R}^n$, and a sufficiently large number $\mu_0 > 0$ such that $\{x | Ax \leq b, \frac{1}{2}(x - x_0)^T(x - x_0) \leq \mu_0\} \neq \emptyset$. Set $k = 0$.

Step 1: Solve the concave quadratic maximization problem $(\text{RCDP})_k$ with parameters μ_k , x_k and $\bar{\lambda}$ to find an optimal solution $\bar{\sigma}^k$. Let $\bar{y}_k = -(\bar{\lambda}I + Q)^{-1}(f + Qx_k + A^T\bar{\sigma}^k)$.

Step 2: If $\bar{y}_k = 0$, stop and report x_k as a KKT point of (LCQP) . Otherwise, set $x_{k+1} = x_k + \bar{y}_k$, $\mu_{k+1} = 1$, and update k by $k + 1$. Go to Step 1.

Theoretically speaking, the initial μ_0 should be large enough to include part of the nonempty feasible domain in order to provide a finite upper bound of $(\text{RCDP})_k$. It follows that the weak duality in Lemma 3 holds and $(\text{RCDP})_k$ provides a finite optimal dual solution. In reality, μ_0 in $(\text{RCDP})_k$ is only a constant that does not affect the dual optimal solution. Hence, in practice, the initial μ_0 can be arbitrarily selected. Also notice that Theorem 1 says that x_1 and all points followed are always feasible to problem (1). Therefore, any arbitrarily chosen $\mu_k > 0$ for $k \geq 1$ can guarantee that $\{x | Ax \leq b, \frac{1}{2}(x - x_k)^T(x - x_k) \leq \mu_k\} \neq \emptyset$. Since the choice of μ_k does not affect the next iterative x_{k+1} , for convenience, we may simply choose $\mu_k = 1$ for $k \geq 1$.

In the next theorem, we show that the proposed ICDA algorithm either terminates with a KKT point of problem (1) in a finite number of iterations or generates an infinite sequence of feasible solutions with decreasing objective values that converges to a KKT point of problem (1).

Theorem 2. Assuming that (LCQP) in form of problem (1) has a nonempty feasible domain and its objective function is bounded below, the ICDA algorithm either terminates with a KKT point of problem (1) with $\bar{y}_k = 0$ in finite many iterations, or it generates a sequence of feasible solutions $\{x_k\}_{k=1}^{\infty}$ of problem (1) with strictly decreasing objective values $\{p(x_k)\}$. Moreover, if the dual sequence $\{\bar{\sigma}^k\}$ is bounded, any convergent subsequence of $\{x_k\}_{k=1}^{\infty}$ converges to a KKT point of problem (1).

Proof. Lemma 4 shows that if the ICDA algorithm terminates at an iteration k with $\bar{y}_k = 0$, then x_k is a KKT point of problem (1). On the other hand, if the ICDA algorithm keeps iterating without termination, then $\bar{y}_k \neq 0$ for all $k > 0$. Using expression (24) in Lemma 5 and the assumption that problem (1) is bounded below, we know the sequence of objective values $\{p(x_k)\}$ is strictly decreasing and bounded below. Hence the sequence converges to its greatest lower bound.

Again, from (24) of Lemma 5, for any positive integer r , we have:

$$p(x_1) - p(x_{r+1}) = \sum_{k=1}^r (p(x_k) - p(x_{k+1})) \geq \frac{\bar{\lambda}}{2} \sum_{k=1}^r \|\bar{y}_k\|^2. \quad (25)$$

This implies that $\sum_{k=1}^{\infty} \|\bar{y}_k\|^2$ is a convergent series and $\lim_{k \rightarrow \infty} \bar{y}_k = 0$. Let (x^*, σ^*) be any limit point of $\{(x_k, \bar{\sigma}_k)\}$. Taking the limit over the convergence subsequence on both sides of expression (19) and using k as the index of the subsequence for convenience, we have:

$$\begin{cases} \lim_{k \rightarrow \infty} (f + Qx_k + A^T \bar{\sigma}_k) = (f + Qx^* + A^T \sigma^*) = 0, \\ \lim_{k \rightarrow \infty} (Ax_k) = Ax^* \leq b, \\ \lim_{k \rightarrow \infty} \bar{\sigma}_k^T (A\bar{y}_k - b + Ax_k) = \sigma^* (Ax^* - b) = 0, \\ \lim_{k \rightarrow \infty} \bar{\sigma}_k = \sigma^* \geq 0. \end{cases} \quad (26)$$

This shows that (x^*, σ^*) is indeed a KKT pair of (1). $\square$

For practical implementations, we may terminate the ICDA algorithm when $\frac{1}{2} \bar{y}_k^T \bar{y}_k \leq \epsilon$ for a sufficiently small $\epsilon > 0$. Then the algorithm stops after taking a finite number of iterations. In this case, the reported x_k can be viewed as an approximate KKT point of problem (1).

4. Preliminary numerical experiments

To evaluate the numerical performance of the canonical duality approach, we provide three examples for testing.

Table 1
Numerical results for the 3 examples.

	Parameters	Av.I	Av.T	Av.T/Av.I	SV.T	Best.V	G.found	ATG
E1	$\lambda_1, x_b = 5$	7.2	5.4711	0.7599	3.9699	-8.00	7	10.6362
E1	$\lambda_2, x_b = 5$	11.75	8.9672	0.7632	7.2228	-8.00	5	19.9625
E2	$\lambda_1, x_b = 1$	13	10.6031	0.8156	0.4417	-10.00	20	10.6031
E2	$\lambda_2, x_b = 1$	27	25.4219	0.9416	0.5614	-10.00	20	25.4219
E2	$\lambda_1, x_b = 3$	17.85	14.4187	0.8078	6.3268	-8.95	0	Null
E2	$\lambda_2, x_b = 3$	31.05	29.7508	0.9581	0.7452	-9.00	0	Null
E3	$\lambda_1, x_b = 100$	13.25	44.4570	3.3552	57.3003	-10000.00	1	23.2656
E3	$\lambda_2, x_b = 100$	29.4	97.0078	3.2996	52.7758	-10000.00	4	57.4570

Abbreviations: Av.I, average iterations; SV.T, standard deviation CPU time; G.found, # of times to find the global optimum; Av.T, average CPU time by seconds; Best.V, the best objective value obtained; ATG, average CPU time to find a global optimum.

Example 1. E1

$$\begin{aligned} Z_0 = \min \quad & \frac{1}{2} x_1^2 - \frac{1}{2} x_2^2 - x_1 - x_2, \\ \text{s.t.} \quad & x_1 + x_2 \leq 1, \\ & -x_1 + x_2 \leq 1, \\ & x_2 \geq -5. \end{aligned}$$

In this example, there are three KKT points, namely, $(1, -5)^T$, $(1, -1)^T$ and $(0, 1)^T$. Among them, $(1, -5)^T$ is the unique global minimizer with the optimal value of -8.

Example 2. E2

$$\begin{aligned} Z_0 = \min \quad & -\frac{1}{2} \sum_{i=1}^s (x_i^2 + x_i) + \frac{1}{2} \sum_{i=s+1}^{s+t} x_i^2, \\ \text{s.t.} \quad & -1 \leq x_i \leq 1, \quad i = 1, 2, \dots, s+t. \end{aligned}$$

The objective function has uncoupled variables and thus can be treated independently. For each of the first s variables, the subproblem $\min_{-1 \leq x_i \leq 1} -\frac{1}{2}(x_i^2 + x_i)$, for $i = 1, \dots, s$, associates with the Lagrangian function:

$$L(x_i, \sigma_1, \sigma_2) = -\frac{1}{2}(x_i^2 + x_i) + \sigma_1(x_i - 1) + \sigma_2(-x_i - 1),$$

where $\sigma_1, \sigma_2 \geq 0$. There are three KKT points: (i) $x_i = -\frac{1}{2}$ with Lagrangian multipliers $(0, 0)$, (ii) $x_i = -1$ with multipliers $(0, \frac{1}{2})$ and (iii) $x_i = 1$ with multipliers $(\frac{3}{2}, 0)$, for $i = 1, \dots, s$. In addition, the subproblem $\min_{-1 \leq x_i \leq 1} \frac{1}{2} x_i^2$ is a convex optimization problem and it has only one KKT point at $x_i = 0$, for $i = s+1, \dots, s+t$. Combining all variables x_i , for $i = 1, \dots, s, s+1, \dots, s+t$, we see there are 3^s KKT points and only one global minimizer $(\text{ones}(1, s), \text{zeros}(1, t))^T$ with the objective value $Z_0 = -s$, where $\text{ones}(1, s)$ and $\text{zeros}(1, t)$ represent a row of s ones and a row of t zeros, respectively. In this example, we choose $s = 10$ and $t = 10$ for our computational experiments. The optimal value should be $Z_0 = -10$.

Example 3. E3

$$\begin{aligned} Z_0 = \min \quad & -\sum_{i=1}^n x_i^2, \\ \text{s.t.} \quad & \sum_{i=1}^j x_i \leq j, \quad j = 1, 2, \dots, n, \\ & x_i \geq 0, \quad i = 1, 2, \dots, n. \end{aligned}$$

This example is taken from Strekalovskii [15]. The global minimizer is $x^T = (0, 0, \dots, 0, n)$ and the optimal value $Z_0 = -n^2$. We choose $n = 100$ in our computational experiments. In this case, the optimal value should be -10,000. All numerical experiments are conducted on an Intel T2050 laptop with 1.60 Gigahertz Duo CPU, 2.00 Gigabyte RAM memory and Microsoft Windows XP version 5.1 oper-

ating system. Since $(\text{RCDP})_k$ is a convex quadratic programming problem, we use Matlab (Version 7.7.0.471) and call the solver `cvx` (Version 1.2) developed by Michael Grant and Stephen Boyd [11] to solve it.

For each example, we test two kinds of $\bar{\lambda}$: (i) $\lambda_1 = -\lambda_0 + 0.001$ which is close to $-\lambda_0$ and (ii) $\lambda_2 = -3\lambda_0$ which is a bit far away. The minimum eigenvalues of Q in the three examples are $\lambda_0 = -1, -1, -2$, respectively. The initial point x_0 is selected to be uniformly distributed from an n – dimensional box $(-\frac{1}{2}x_b, \frac{1}{2}x_b)^n$ where x_b is a controllable parameter determining the size of the box. The ICDA algorithm is stopped whenever $\frac{1}{2}\|\bar{y}_k\|^2 < 10^{-8}$. In each example, when $\bar{\lambda}$ and x_b are decided, we run it for 20 times with 20 different randomly-sampled x_0 and report the average results of them.

Table 1 documents the numerical results for examples E1, E2 and E3, respectively. All the abbreviations are defined at the bottom of the table and “Null” means “not available” for the item.

From the two columns reporting ‘Av.I’ and ‘Av.T’, we observe that, in average, ICDA with a smaller λ takes fewer iterations and spends less time to reach a KKT point. The average run time per iteration (‘Av.T/Av.I’) for solving each problem $(\text{RCDP})_k$ is also shorter for a smaller λ , expect for problem E3. For the columns ‘Best.V’ and ‘G.found’, they show that the optimal solutions of E1, E2, and E3 can be found, respectively, in 20 trials of ICDA with properly chosen parameters. For the larger size problem E3, it is more difficult to find global optimal solutions, but using λ_2 seems to be more effective than using λ_1 . Our results indicate the potential of using the canonical duality approach to solve the linearly constrained nonconvex quadratic programming problems.

It is interesting to observe that when a larger box with $x_b = 3$ is used to choose the initial solution, none of λ_1 and λ_2 was able to solve problem E2 which consists of 3^{10} KKT points. On the other hand, when $x_b = 1$, all trials with λ_1 and λ_2 succeeded. Recall that the optimal solution of E2 has the first 10 components being 1 and the rest of them being 0. Some other non-optimal KKT points have their first 10 components either being -1 or $-1/2$ and the rest of them being 0. Selecting an initial solution randomly from the 20-dimensional box $(-\frac{1}{2}, \frac{1}{2})^{20}$ is very likely, with more than 99.99% of chance, to have at least one negative component. For those initial solutions having negative components, -1 or $-1/2$ is a much closer KKT component. Yet, the ICDA algorithm would rather to approach a farther one at 1.

In general, we can say from the preliminary numerical results that smaller λ 's run faster, but do not guarantee a better chance for finding the global optimum. We will provide further discussions on the choice of λ in the next section.

5. Discussion and concluding remarks

Note that $(\text{RCDP})_k$ is a restricted subproblem of $(\text{CDP})_k$ at the k^{th} iteration in which $\bar{\lambda} > 0$ is being fixed with $\bar{\lambda}I + Q \succ 0$. We discuss the choice of $\bar{\lambda} > 0$ in this final section.

Lemma 6. Let $\bar{\lambda} > \max\{0, -\lambda_0\}$ be a fixed number and $\bar{\sigma}^k$ be a finite maximizer of program $(\text{RCDP})_k$. If

$$\frac{1}{2}(f + Qx_k + A^T \bar{\sigma}^k)^T (\bar{\lambda}I + Q)^{-2} (f + Qx_k + A^T \bar{\sigma}^k) \neq \mu_k,$$

then there exists a $\lambda' > \max\{0, -\lambda_0\}$, $\lambda' \neq \bar{\lambda}$ such that:

$$Z^{\text{rd}}(x_k, \mu_k, \lambda') > Z^{\text{rd}}(x_k, \mu_k, \bar{\lambda}).$$

In particular, if μ_k is sufficiently large such that:

$$\frac{1}{2}(f + Qx_k + A^T \bar{\sigma}^k)^T (\bar{\lambda}I + Q)^{-2} (f + Qx_k + A^T \bar{\sigma}^k) < \mu_k,$$

then there exists a $\delta > 0$, $\bar{\lambda} - \delta > \max\{0, -\lambda_0\}$ such that:

$$Z^{\text{rd}}(x_k, \mu_k, \lambda) > Z^{\text{rd}}(x_k, \mu_k, \bar{\lambda}),$$

for any $\lambda \in (\bar{\lambda} - \delta, \bar{\lambda})$.

Proof. Suppose that $\frac{1}{2}(f + Qx_k + A^T \bar{\sigma}^k)^T (\bar{\lambda}I + Q)^{-2} (f + Qx_k + A^T \bar{\sigma}^k) < \mu_k$, then there exists a $\delta > 0$ such that $\bar{\lambda} - \delta > \max\{0, -\lambda_0\}$ and $\frac{1}{2}(f + Qx_k + A^T \bar{\sigma}^k)^T (\lambda I + Q)^{-2} (f + Qx_k + A^T \bar{\sigma}^k) < \mu_k$ for all $\lambda \in (\bar{\lambda} - \delta, \bar{\lambda})$. This implies, from expression (9) of Lemma 2, that:

$$\begin{aligned} \frac{\partial d_k(\bar{\sigma}^k, \lambda)}{\partial \lambda} &= \frac{1}{2}(f + Qx_k + A^T \bar{\sigma}^k)^T (\lambda I + Q)^{-2} (f + Qx_k + A^T \bar{\sigma}^k) - \mu_k \\ &< 0 \forall \lambda \in (\bar{\lambda} - \delta, \bar{\lambda}). \end{aligned}$$

Moreover, $d_k(\bar{\sigma}^k, \lambda)$ is decreasing on $(\bar{\lambda} - \delta, \bar{\lambda})$. Therefore, for $\lambda' \in (\bar{\lambda} - \delta, \bar{\lambda})$:

$$\begin{aligned} Z^{\text{rd}}(x_k, \mu_k, \bar{\lambda}) &= d_k(\bar{\sigma}^k, \bar{\lambda}) < d_k(\bar{\sigma}^k, \lambda') \leq \max_{\sigma \geq 0} d_k(\sigma, \lambda') \\ &= Z^{\text{rd}}(x_k, \mu_k, \lambda'). \end{aligned}$$

The other case with $\frac{1}{2}(f + Qx_k + A^T \bar{\sigma}^k)^T (\bar{\lambda}I + Q)^{-2} (f + Qx_k + A^T \bar{\sigma}^k) > \mu_k$ can be similarly proven for $\lambda' \in (\bar{\lambda}, \bar{\lambda} + \delta)$. $\square$

Lemma 6 shows that when $(\text{LCQP})_k$ is formulated with μ_k and its restricted canonical dual $(\text{RCDP})_k$ is solved with a fixed $\bar{\lambda}$, if $\bar{y}_k$ does not lie on the boundary of the ball of $\{y \in R^n | \frac{1}{2}y^T y \leq \mu_k\}$, then the canonical dual problem $(\text{CDP})_k$ can be better approximated by choosing a different λ' .

Corollary 1. Assume that (LCQP) in form of problem (1) has a bounded feasible domain and $\mu_k > 0$ is sufficiently large such that $\{x | \frac{1}{2}x^T x \leq \mu_k\} \supset \{x | Ax \leq b\}$ and $\{x | \frac{1}{2}x^T x = \mu_k\} \cap \{x | Ax \leq b\} = \emptyset$.

Then the limit $\lim_{\lambda \rightarrow \max\{0, -\lambda_0\}^+} Z^{\text{rd}}(0, \mu_k, \lambda)$ exists and

$$Z^{\text{rd}}(0, \mu_k, \lambda) < \lim_{\lambda \rightarrow \max\{0, -\lambda_0\}^+} Z^{\text{rd}}(0, \mu_k, \lambda) = Z^d(0, \mu_k) \leq Z^p(0, \mu_k) = Z.$$

Proof. Recall what we have shown before that any $\bar{\lambda} > \max\{0, -\lambda_0\}$ can lead to a maximizer $\bar{\sigma}^k$ for $(\text{RCDP})_k$ and a corresponding $\bar{y}_k$ that is feasible to $Ax \leq b$. Since $\{x | \frac{1}{2}x^T x \leq \mu_k\} \supset \{x | Ax \leq b\}$ and $\{x | \frac{1}{2}x^T x = \mu_k\} \cap \{x | Ax \leq b\} = \emptyset$, we must have $\frac{1}{2}\bar{y}_k^T \bar{y}_k < \mu_k$. By Lemma 6, $\bar{\lambda}$ can not be optimal to $(\text{CDP})_k$. Indeed, we have:

$$Z^d(0, \mu_k) = \sup_{\lambda \rightarrow \max\{0, -\lambda_0\}^+} Z^{\text{rd}}(0, \mu_k, \lambda).$$

This proves the lemma. $\square$

Corollary 1 indicates that when (LCQP) has a bounded feasible domain and $\mu_k > 0$ is sufficiently large, problem $(\text{LCQP})_k$ coincides with problem (LCQP) and $\lim_{\lambda \rightarrow \max\{0, -\lambda_0\}^+} Z^{\text{rd}}(0, \mu_k, \lambda)$ offers the best lower bound for problem (LCQP) . From expression (18), this best bound can also be expressed as:

$$\begin{aligned} Z^d(0, \mu_k) &= \sup_{\lambda > \max\{0, -\lambda_0\}} \left\{ \max_{\sigma \geq 0} d_k(\sigma, \lambda) \right\} \\ &= \sup_{\lambda > \max\{0, -\lambda_0\}} \left\{ p(y_k(\lambda)) + \lambda \left(\frac{1}{2}y_k(\lambda)^T y_k(\lambda) - \mu_k \right) \right\}, \end{aligned}$$

where $y_k(\lambda) = -(\lambda I + Q)^{-1}(f + A^T \sigma^k(\lambda))$ and $\sigma^k(\lambda)$ maximizes $d_k(\sigma, \lambda)$ over $\sigma \geq 0$ for each $\lambda > \max\{0, -\lambda_0\}$. The lower bound consists of two terms, namely, $p(y_k(\lambda))$ and a duality gap $\lambda(\frac{1}{2}y_k(\lambda)^T y_k(\lambda) - \mu_k)$ due to the negativity of eigenvalue λ_0 and the involvement of μ_k . We observe that a smaller λ does give a better lower bound, but it may actually result in a larger (worse) primal value $p(y_k(\lambda))$. Fortunately, this undesired situation can never happen when we deal with a linearly constrained convex quadratic program as shown below.

Corollary 2. Let $\epsilon > 0$ be an arbitrary number. If the Q matrix of (LCQP) in form of problem (1) is positive semi-definite, then there exists an $M > 0$ such that $\{x | \frac{1}{2}x^T x \leq M\}$ contains a minimizer of problem (1) and the corresponding (RCDP) $_k$ with parameters $\mu_k = M$, $x_k = 0$ and $\lambda = \epsilon/M$ generates an ϵ – optimal solution of problem (1).

Proof. Under the assumptions that problem (1) has a nonempty feasible domain and its objective function is bounded below over the feasible domain, there exists a finite minimizer for linearly constrained convex quadratic programming problems [7]. Therefore, there exists an $M > 0$ such that the set $\{x \in R^n | \frac{1}{2}x^T x \leq M\}$ contains a minimizer of problem (1) with $Z = Z^p(0, M)$. When Q is positive semi-definite, its minimum eigenvalue $\lambda_0 \geq 0$. Using Theorem 1 with $\bar{\lambda} = \epsilon/M$, we can get a feasible solution $\bar{x}$ of problem (1) such that:

$$p(\bar{x}) - Z^p(0, M) = p(\bar{x}) - Z \leq \frac{\epsilon}{M} \left(M - \frac{1}{2} \bar{x}^T \bar{x} \right) \leq \epsilon.$$

Hence $\bar{x}$ is an ϵ – optimal solution of problem (1). $\square$

Corollary 2 says that an ϵ – optimal solution for linearly constrained convex quadratic programming problems can be obtained by solving problem (RCDP) $_k$ only once with a sufficiently small λ .

In summary, the performance of the proposed ICDA algorithm depends solely on the parameter $\bar{\lambda}$ and the location of the initial point x_0 . Once these two factors are given, the ICDA algorithm at each iteration x_k produces a *large-step steepest descent direction* $\bar{y}_k$ in the sense that $x_{k+1} = x_k + \bar{y}_k$ minimizes $p(x) = \frac{1}{2}x^T Qx + f^T x$ over $Ax \leq b$ and $\frac{1}{2}(x - x_k)^T (x - x_k) \leq \frac{1}{2} \|\bar{y}_k\|^2$. We say that it is a “steepest descent” direction, because $\bar{y}_k$ points to the lowest value in the neighborhood of x_k with a radius of $\|\bar{y}_k\|$. We also say that it is a “large-step”, because $\|\bar{y}_k\|$ is in general not very small unless x_k is close to a KKT point. Although the traditional (negative) gradient direction with a line search is often termed as a “steepest descent method”, it is indeed descent, but the “steepest” property holds only “instantaneously” or “locally”. Notice that, with a larger value of $\|\bar{y}_k\|$, the algorithm is searching for a bigger domain and is considered to have better performance. The performance of the proposed ICDA algorithm depends on the choice of $\bar{\lambda}$, but, unfortunately, we do not know how to set it *a priori*.

Before concluding this paper, we want to mention that, although we have shown that the proposed canonical duality based ICDA algorithm can indeed provide KKT solutions to linearly constrained quadratic programming problems, the quality of the solutions depends on the parameter λ and the selection of an initial point. Our computational results only indicate the possibility of applying ICDA with different starting points to generate a good KKT point that is a (near) global minimizer. Many issues remain to be further addressed. We are particularly interested in studying the following issues: (i) the convergence rate of the proposed ICDA algorithm, (ii) how to choose λ to speed up the convergence rate,

and (iii) the pros and cons of solving problem (CDP) $_k$ directly, instead of solving problem (RCDP) $_k$.

Acknowledgments

Xing's research has been supported by the Key Project No. 108005 of the Chinese Ministry of Education and Grant Nos. 10801087 and 11171177 of the China National Science Foundation. Fang's research has been supported by Grant No. DMI-0553310 of the United States National Science Foundation. Sheu's research has been sponsored by the Grant No. 98-2115-M-006-010-MY2 of the Taiwan National Science Council.

References

- [1] O. Barrientos, R. Correa, An algorithm for global minimization of linearly constrained quadratic functions, *Journal of Global Optimization* 16 (2000) 77–93.
- [2] S.C. Billups, K.G. Murty, Complementarity problems, *Journal of Computational and Applied Mathematics* 124 (2000) 303–318.
- [3] S. Burer, D. Vandenbussche, A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations, *Programming, Series A* 113 (2008) 259–282.
- [4] S.J. Chung, K.G. Murty, Polynomially bounded algorithms for convex quadratic programming, in: O.L. Mangasarian, R. Meyer, S.M. Robinson (Eds.), *Nonlinear Programming*, vol. 4, Academic Press, 1981, pp. 439–485.
- [5] S.-C. Fang, D.Y. Gao, R.L. Sheu, S.Y. Wu, Canonical dual approach to solving 0–1 quadratic programming problems, *Journal of Industrial and Management Optimization* 4 (2008) 125–142.
- [6] S.-C. Fang, D.Y. Gao, R.L. Sheu, W. Xing, Global optimization for a class of fractional programming problems, *Journal of Global Optimization* 45 (2009) 337–353.
- [7] R. Fletcher, A general quadratic programming algorithm, *IMA Journal of Applied Mathematics* 7 (1971) 76–91.
- [8] D.Y. Gao, Canonical dual transformation method and generalized triality theory in nonsmooth global optimization, *Journal of Global Optimization* 17 (2000) 127–160.
- [9] D.Y. Gao, Canonical duality theory and solutions to constrained nonconvex quadratic programming, *Journal of Global Optimization* 29 (2004) 377–399.
- [10] N.I.M. Gould, P.L. Toint, Numerical methods for large-scale non-convex quadratic programming, in: A.H. Siddiqui, M. Kočvara (Eds.), *Trends in Industrial and Applied Mathematics: Proceedings of the 1st International Conference on Industrial and Applied mathematics of the Indian Subcontinent*, 2002, pp. 149–179.
- [11] M. Grant, S. Boyd, *cvx Users' Guide for cvx version 1.2*. <www.stanford.edu/~boyd/cvxbook/>, 2009.
- [12] Y. Nesterov, A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*, Studies in Applied Mathematics (1994) 13.
- [13] P.M. Pardalos, S.A. Vavasis, Quadratic programming with one negative eigenvalue is NP-hard, *Journal of Global Optimization* 1 (1991) 15–22.
- [14] H.D. Sherali, C.H. Tuncbilek, A reformulation-convexification approach for solving nonconvex quadratic programming problems, *Journal of Global Optimization* 7 (1995) 1–31.
- [15] A.S. Strekalovskii, The search for a global maximum of a convex functional on an admissible set, *Computational Mathematics and Mathematical Physics* 33 (1993) 315–328.
- [16] Z. Wang, S.-C. Fang, D.Y. Gao, W. Xing, Global extremal conditions for multi-integer quadratic programming, *Journal of Industrial and Management Optimization* 4 (2008) 213–225.
- [17] W. Xing, S.-C. Fang, R.-L. Sheu, L. Zhang, Canonical dual solutions to the quadratic programming over a quadratic constraint, *Journal of Global Optimization*, accepted for publication.