ON BOX-CONSTRAINED TOTAL LEAST SQUARES PROBLEM

ZHUOYI XU, YONG XIA AND DEREN HAN
LMIB of the Ministry of Education, School of Mathematical Sciences
Beihang University, Beijing, 100191, P. R. China

ABSTRACT. We study box-constrained total least squares problem (BTLS), which minimizes the ratio of two quadratic functions with lower and upper bounded constraints. We first prove that (BTLS) is NP-hard. Then we show that for fixed number of dimension, it is polynomially solvable. When the constraint box is centered at zero, a relative $4/7$-approximate solution can be obtained in polynomial time based on SDP relaxation. For zero-centered and unit-box case, we show that the direct nontrivial least square relaxation could provide an absolute $(n + 1)/2$-approximate solution. In the general case, we propose an enhanced SDP relaxation for (BTLS). Numerical results demonstrate significant improvements of the new relaxation.

1. Introduction. In this paper, we consider the box-constrained total least square problem, which minimizes the ratio of two particular quadratic functions with bounded constraints:

\[(BTLS) \quad \min f(x) = \frac{\|Ax - b\|^2}{x^Tx + 1} \quad (1)\]

\[\text{s.t.} \quad l \leq x \leq u, \quad (2)\]

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\| \cdot \|$ denotes the Euclidean norm, $l$ and $u$ are lower and upper bounds of $x$, respectively, and it is assumed that

\[l_i < u_i, \quad i = 1, \ldots, n.\]

Problem (BTLS) arises from modeling over-determined system of linear equations $Ax \approx b$, where both the data matrix $A$ and the observation vector $b$ are contaminated by noises or affected by various errors. For solving such an over-determined system of linear equations, the total least squares (TLS) method is first introduced in [13], which is to find the minimizer $(E, r)$ of the optimization problem

\[\min_{E, r, x} \{\|E\|^2_F + \|r\|^2 : (A + E)x = b + r\},\]

where $E \in \mathbb{R}^{m \times n}$ and $r \in \mathbb{R}^m$ are unknown perturbations, and $\|(\cdot)\|_F$ denotes the Frobenius norm of matrix $(\cdot)$. (TLS) method has been widely used in many scientific fields such as signal processing, automatic control. Observing that the optimization
problem in terms of $E$ and $r$ is a linear-equality constrained convex quadratic program, one can write the minimizer $(E, r)$ as an explicit formulation with respect to $x$. Consequently, (TLS) can be reformulated as the following quadratic fractional optimization problem \[2\]

\[
\text{(TLS)} \min_{x \in \mathbb{R}^n} \frac{\|Ax - b\|^2}{\|x\|^2 + 1}.
\]

(TLS) can be efficiently solved. Actually, the optimal value of (TLS) is equal to the minimal eigenvalue of the augmented matrix $[A, b]^T [A, b]$, which can be done by a singular-value decomposition of the augmented matrix $[A, b]$ \[13\].

Quadratically constrained (TLS) is studied in \[4\]. In many practical situations, the linear system is ill-conditioned. Regularization is a powerful way to stabilize the solution, which introduces the following single quadratic constraint

\[
\|Lx\|^2 \leq \rho,
\]

where $L \in \mathbb{R}^{d \times n}$ is an instance-dependent regularization matrix and $\rho > 0$ is a parameter. Though being a nonconvex optimization problem, this case can be efficiently and globally solved. Beck, Ben-Tal and Teboulle present a bisection algorithm \[2\] and each subproblem is a generalized trust-region subproblem enjoying hidden convexity \[6, 24\]. Notice that their algorithm also works for the two-sided quadratic constrained case

\[
\eta \leq \|Lx\|^2 \leq \rho.
\]

The convergence of the bisection algorithm is certainly linear. Later, a fixed point iteration method with superlinear global convergence rate is proposed by Beck and Teboulle \[3\]. Actually, this method coincides with the well-known Dinkelbach’s algorithm \[11\] in the area of fractional programming. The superlinear convergence of Dinkelbach’s algorithm is established in \[21\]. Moreover, under some assumption, (TLS) with an ellipsoid constraint can be reformulated as a semidefinite programming (SDP) problem \[3\]. Necessary and sufficient condition for the assumption to guarantee the hidden convexity is presented in \[23, 25\].

(BTLS) is first presented in \[4\] as a special case of the quadratically constrained (TLS), since the constraints (2) can be reformulated as quadratic inequalities

\[
(x_i - u_i)(x_i - l_i) \leq 0, \quad i = 1, \ldots, n.
\]

In some practical applications of (BTLS), box constraints are necessarily required. For example, in the application of image in a BMP format, each component (or pixel) $(x_i)$ lies between 0 and 255. That is, we have $l = 0, \ u = 255$ in (BTLS). Beck and Teboulle \[4\] propose a general SDP relaxation for approximating quadratically constrained (TLS). To the best of our knowledge, (BTLS) has not been carefully studied in literature. We also notice that Dinkelbach’s algorithm may fail to efficiently solve (BTLS), though it works well for the single-ellipsoid constrained (TLS). Let us consider a trivial case of (BTLS) (for example, the solution of the unconstrained (TLS) happens to lie in the box), Dinkelbach’s algorithm needs to iteratively solve difficult nonconvex quadratic programming subproblems.

In this paper, we first prove that (BTLS) is NP-hard. We propose a global optimization algorithm to solve (BTLS). Polynomial-time solvability is shown when the dimension $n$ is fixed.

Our second contribution is in the SDP relaxation for (BTLS). Following Goesman and Williamson’s SDP-based 0.878-approximation for Maxcut problem \[14\], approximating particular nonconvex quadratic optimization problems via SDP relaxation
attracts many researcher’s attention [16]. In particular, Nesterov [19] establishes $2/\pi$-approximation for binary quadratic optimization and then generalizes to quadratic program with homogeneous and diagonal quadratic constraints and gets a $4/7$-approximation, which is independently established in Ye [26]. In this paper, we show that the $4/7$-approximation can be extended to (BTLS) when the boxed feasible region is centered at the origin 0, i.e., $l_i + u_i = 0$ for $i = 1, \ldots, n$. In the general case, we enhance the standard SDP relaxation by adding redundant second-order cone constraints. Significant improvements are observed in limited numerical experiments.

The remainder of this paper is organized as follows. In Section 2, we prove the NP-hardness of (BTLS). In Section 3, we globally solve (BTLS), which implies that (BTLS) is polynomially solvable when $n$ is fixed. In Section 4, we extend the $4/7$-approximation to (BTLS) when $l_i + u_i = 0$ for all $i$. We strengthen the standard SDP relaxation and do numerical experiments in Section 5. We conclude this paper in Section 6.

**Notations.** Throughout this paper, $v(\cdot)$ stands for the optimal value of problem $(\cdot)$. $E(\cdot)$ represents the average value of $(\cdot)$. $A \succ (\succeq) 0$ means that matrix $A$ is positive (semi)definite. $\text{Tr}(A)$ denotes the trace of matrix $A$. $I_n$ denotes the identity matrix of order $n$. Let $e$ be the all-one $n$-dimensional vector and $e_i$ be the $i$-th column of $I_n$. For a vector $a$, $A = \text{Diag}(a)$ returns a diagonal matrix with $A_{ii} = a_i$.

2. NP-hardness. The main result in this section is to prove the NP-hardness of (BTLS).

**Theorem 1.** (BTLS) is NP-hard.

**Proof.** Let $a = (a_1, \ldots, a_n)^T$ be a vector with integer entries, the partition problem (PP) checking whether the following equation

$$a^T x = 0, \ x \in \{-1, 1\}^n$$

has a solution is NP-hard [12]. We show that (PP) can be polynomially reduced to a special case of (BTLS). Let $l = -e$, $u = e$, and

$$A = \begin{pmatrix} a^T \\ 0_n^T \end{pmatrix}, \ b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(BTLS) with the above equipment can be written as follows:

$$\begin{align*}
\text{(C)} & \quad \min \quad \frac{(a^T x)^2 + 1}{x^T x + 1} \\
& \quad \text{s.t.} \quad -e \leq x \leq e.
\end{align*}$$

(3)

(4)

Notice that for any feasible solution of (C) it holds that

$$\begin{align*}
(a^T x)^2 + 1 & \geq 1, \\
x^T x + 1 & \leq n + 1.
\end{align*}$$

(5)

(6)

Therefore, we have

$$v(C) \geq \frac{1}{n + 1}.$$  

(7)

The equality in (7) is attained if and only if there is an $x \in [-e, e]$ such that both inequalities (5) and (6) hold as equalities simultaneously. That is, $v(C) = 1/(n + 1)$ if and only if (PP) has a solution. As we have reduced the NP-hard problem (PP) to (C), a special case of (BTLS), the proof is complete. \qed
3. Polynomial solvable case. In this section, we present a global optimization algorithm to solve (BTLS), which is done in polynomial time if the dimension $n$ is fixed.

According to the generalized Charnes-Cooper transformation [9], we introduce new variables
\[ y = \frac{x}{\sqrt{x^T x + 1}} \in \mathbb{R}^n, \quad z = \frac{1}{\sqrt{x^T x + 1}} \in \mathbb{R}, \]
and then reformulate (BTLS) as the following nonconvex quadratic optimization with one sphere and some other linear constraints:
\begin{align}
\text{(ETRS)} \quad & \min \| Ay - bz \|^2 \\
\text{s.t.} \quad & y^T y + z^2 = 1, \\
& l_i z \leq y_i \leq u_i z, \quad i = 1, \ldots, n. 
\end{align}

The above formulation is known as extended trust-region subproblem in literature. Based on the fact that quadratic minimization problem over a ball or sphere (i.e., the classical trust-region subproblem) has at most one local-non-global minimizer [17], global optimization algorithms are proposed for (ETRS) by enumerating all local minimizers [15, 7]. Recently, as a further improvement, Beck and Pan [5] proposed a branch and bound algorithmic scheme to solve problem of this type.

When applying the above global optimization algorithms to solve (BTLS), we show that the procedures can be simplified. It is based on the following key observation, which is a corollary of Lemma 3.2 in [17]. Here we give a short reproof.

**Lemma 2.** For any $A \in \mathbb{R}^{n \times n}$, the quadratic form $x^T Ax$ has no local-non-global minimizer over the unit sphere $x^T x = 1$.

**Proof.** Since orthogonal transformation does not change the number of local minimizers, without loss of generality, we assume $A$ is a diagonal matrix with eigenvalues $\lambda_1 = \ldots = \lambda_s < \lambda_{s+1} \leq \ldots \leq \lambda_n$. Then the global optimal solution set of the minimum eigenvalue problem
\begin{align}
\text{(E)} \quad & \min \ h(x) = x^T Ax \\
\text{s.t.} \quad & x^T x = 1
\end{align}
is given by
\[ S = \{x : x^T x = 1, \ x_i = 0, \forall i = s+1, \ldots, n\}. \]

Suppose $\bar{x}$ is a local-non-global minimizer of (E), there is a $j \in \{s+1, \ldots, n\}$ such that $\bar{x}_j \neq 0$. For any positive $\epsilon < \bar{x}_j^2$, define $z(\epsilon) \in \mathbb{R}^n$ with elements being
\[ z_i(\epsilon) = \begin{cases} 
\sqrt{\bar{x}_i^2 + \epsilon} & \text{if } i = 1, \\
\sqrt{\bar{x}_i^2 - \epsilon} & \text{if } i = j, \\
\bar{x}_i & \text{otherwise.}
\end{cases} \]

Then, $z(\epsilon)$ is on the unit sphere and
\[ h(z(\epsilon)) = h(\bar{x}) + \epsilon(\lambda_1 - \lambda_j) < h(\bar{x}), \]
which yields a contradiction. The proof is complete. \qed

Lemma 2 admits the following extension.

**Lemma 3.** For any $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times n}$, the problem
\begin{align}
\text{(BE)} \quad & \min \ x^T Ax \\
\text{s.t.} \quad & x^T x = 1, Bx = 0
\end{align}
has no local-non-global minimizer.
Proof. Let the columns of $V \in \mathbb{R}^{n \times k}$ be orthogonal to each other and span the null space of $B$. Then

$$V^TV = I_k, \ \{x : Bx = 0\} = \{Vy : y \in \mathbb{R}^k\},$$

and hence

$$v(\text{BE}) = \min_{y^Ty = 1} y^TV^TAVy.$$

According to Lemma 2, there is no local-non-global minimizer for the reduced problem in terms of $y$. Notice that the mapping from $x$ to $y$ is linear. The proof is complete. \hfill \square

Now, we present a simplified backtracking algorithm to globally solve (ETRS). Let $\Omega_n = \{(y, z) : l_i z \leq y_i \leq u_i z, \ i = 1, \ldots, n\}$. First solve the eigenvalue problem

$$\min_{y^Ty+z^2=1} \|Ay - bz\|^2$$

and get the optimal solution set $S^*$, which is a unit-ball in the eigenspace corresponding to the minimum eigenvalue of $[A-b]T[A-b]$. The set $S^* \cap \Omega_n$, if nonempty, is also the globally optimal solution set of (ETRS). Notice that $S^* \cap \Omega_n \neq \emptyset$ if and only if

$$0 \geq \min_{(y,z) \in S^* \cap \Omega_n} (y_n - l_n z)(y_n - u_n z),$$

which amounts to solve another (ETRS) with $n-1$ linear two-sided constraints. Otherwise, $S^* \cap \Omega_n = \emptyset$. Then (ETRS) reduces to two subproblems with the last two-sided linear constraint being replaced by $y_n - u_n z = 0$ and $y_n - l_n z = 0$, respectively. By suitable variable reduction, these two subproblems are both in the formulation of (ETRS) with $n-1$ linear two-sided constraints. Notice that each subproblem has no local-non-global minimizer according to Lemma 3. The above procedure can be repeated until there is no any constraint. The final unconstrained optimization problem (ETRS) reduces to a minimum eigenvalue problem, which can be efficiently solved in $O(n^3)$ time.

Denote by $C_k$ the complexity of (ETRS) with $k$ linear two-sided constraints. Then we have the recursive relation:

$$C_k = 3C_{k-1}.$$ Consequently, we get an overestimate of the total complexity of the above algorithm.

**Theorem 4.** Problem (BTLS) can be globally solved in $O(3^nn^3)$ time. Thus, when $n$ is fixed, (BTLS) is polynomially solvable.

To our best knowledge, this is the first algorithm to globally solve (BTLS).

General box-constrained nonconvex optimization problem is difficult to solve. As shown by Theorem 1.1.2 in [20], using any derivative-free method to globally find an $\epsilon$-approximation minimizer of a Lipschitz continuous function over a $n$-dimensional box requires at least $O(1/\epsilon^n)$ function evaluations in the worst case. Thus it is time-consuming to get a high-accurate solution even when $n$ is small. As a comparison, our algorithm efficiently solves small dimensional instances of (BTLS) according to Theorem 4.
4. **Approximation.** In this section, we present approximation analysis for (BTLS) based on SDP relaxation and discrete least square relaxation, respectively.

(ETRS) (8)-(10) can be reformulated as homogeneously and quadratically constrained quadratic programming problem:

\[(\text{QCQP}) \quad \min_{y,z} \quad g(y, z) = \|Ay - bz\|^2 \quad \text{(11)} \]
\[\text{s.t.} \quad y^T y + z^2 = 1, \quad \text{(12)} \]
\[\left( e_i^T y - \frac{u_i + l_i}{2} z \right)^2 \leq \left( \frac{u_i - l_i}{2} z \right)^2, \quad i = 1, \ldots, n. \quad \text{(13)} \]

Semidefinite programming (SDP) relaxation is an efficient approach to approximate nonconvex quadratic optimization problems. As a convex relaxation, SDP can be solved in polynomial time \[\text{[18]}\]. By lifting \((y^T, z) \mapsto W \in \mathbb{R}^{(n+1)\times(n+1)}\), Beck and Teboulle \[\text{[4]}\] propose the following SDP relaxation of (QCQP):

\[(\text{SDP}) \quad \min_W \quad \text{Tr}(\mathcal{M}W) \quad \text{(14)} \]
\[\text{s.t.} \quad \text{Tr}(W) = 1, \quad \text{(15)} \]
\[\text{Tr}(D_iW) \leq 0, \quad i = 1, 2, \ldots, n \quad \text{(16)} \]
\[W \succeq 0, \quad \text{(17)} \]

where
\[\mathcal{M} = \begin{pmatrix} A^T A & -A^T b \\ -b^T A & b^T b \end{pmatrix}, \quad \text{and} \quad D_i = \begin{pmatrix} e_i^T e_i & -\frac{l_i + u_i}{2} e_i^T \\ -\frac{l_i + u_i}{2} e_i & u_i l_i \end{pmatrix}. \]

The dual of the problem (SDP) reads as follows

\[(\text{D}) \quad \max \quad -\lambda_0 \quad \text{s.t.} \quad \mathcal{M} + \lambda_0 I_{n+1} + \sum_{i=1}^n \lambda_i D_i \succeq 0 \]
\[\lambda_i \geq 0, \quad i = 1, 2, \ldots, n. \]

There is no duality gap between (SDP) and (D) due to the following result.

**Lemma 5.** Suppose \(l_i < u_i\) for \(i = 1, 2, \ldots, n\), Slater’s condition holds for both (SDP) and (D).

**Proof.** Let \((y_0, z_0)\) be an interior point of (QCQP). Define
\[W(\epsilon) = \frac{\epsilon}{n+1} I_{n+1} + (1 - \epsilon) \begin{pmatrix} y_0y_0^T & z_0 y_0 \\ z_0 y_0^T & z_0^2 \end{pmatrix}. \]
Then, for any sufficiently small \(\epsilon > 0\), one can verify that \(W(\epsilon)\) is a relative interior feasible point of (SDP).

It is not difficult to check Slater’s condition for (D). Actually, for any \(\lambda_i > 0, i = 1, 2, \ldots, n\), any sufficiently large \(\lambda_0\) leads to an interior feasible point of (D). The proof is complete.

Clearly, (SDP) provides a lower bound of (BTLS), i.e.,
\[v(\text{SDP}) \leq v(\text{BTLS}). \]
(SDP) is observed to be exact in many examples. In particular, when \(n = 1\), we can prove its exactness.

**Proposition 1.** When \(n = 1\) and \(l_1 < u_1\), \(v(\text{SDP}) = v(\text{BTLS}).\)

**Proof.** According to Lemma 5, Slater’s condition holds for both (SDP) and (D). Then applying Theorem 2.3 in \[\text{[27]}\] completes the proof.
Even if (SDP) is not exact, it could provide an approximation solution of high quality. Let $W^*$ be an optimal solution of (SDP). Beck and Teboulle present in [4] an approach to generate an approximation solution of (BTLS). They first solve a unit-eigenvector corresponding to the maximum eigenvalue of $W^*$, denoted by $(y^*, z^*)$ where $y^* \in \mathbb{R}^n$, $z^* \in \mathbb{R}$, and then output $y^*/z^*$.

We notice that, when the boxed feasible region is centered at the origin 0, i.e., $l_i + u_i = 0$ for $i = 1, \ldots, n$, (QCQP) (11)-(12) becomes a nonconvex quadratic optimization with simple (i.e., homogeneous and diagonal) quadratic constraints. An $4/7$-approximation analysis has been established in [19, 26] for QCQP with such structure.

We now present the randomized approach [26]. Let $W^* = V^* V^*$ be the Cholesky decomposition of $W^*$. Define $D^* = \text{Diag}(\sqrt{w_{11}^*}, \sqrt{w_{22}^*}, \ldots, \sqrt{w_{n+1,n+1}^*})$.

Let $\sigma(t)$ be a vector with the $j$-th component being $(\sigma(t))_j = \begin{cases} 1, & \text{if } t_j \geq 0, \\ -1, & \text{else.} \end{cases}$ (18)

Let $\eta$ be uniformly and randomly distributed on the $(n+1)$-dimensional unit-sphere. Define $(\hat{y}, \hat{z}) = D^* \sigma(V^* \eta)$, where $\hat{y} \in \mathbb{R}^n$, $\hat{z} \in \mathbb{R}$.

**Theorem 6** ([26]). When $l_i + u_i = 0$ for $i = 1, \ldots, n$, $(\hat{y}, \hat{z})$ is a feasible point of (QCQP)

$$E(g(\hat{y}, \hat{z})) - q \leq \frac{4}{7} q - \overline{q},$$

where $\overline{q} = v(\text{QCQP})$ and $\overline{q}$ is the maximum value of $g(y, z)$ over the constraints (11)-(13).

Now we reconsider (BTLS). Since the above generated $(\hat{y}, \hat{z})$ satisfies (12)-(13), $\hat{z} \neq 0$, and hence we can define

$$\hat{x} = \frac{\hat{y}}{\hat{z}}.$$ 

It follows from the constraint (13) that $\hat{x}$ is a feasible solution of (BTLS). Notice that

$$g(\hat{y}, \hat{z}) = f(\hat{x}).$$

Let $\overline{p} = \max_{x \in [l, u]} f(x)$ and $\bar{p} = v(\text{BTLS})$. Then we have $\overline{p} = \overline{q}$ and $\bar{p} = q$.

As a corollary of Theorem 6, we have the following approximation for (BTLS).

**Corollary 1.** When $l_i + u_i = 0$ for $i = 1, \ldots, n$, $\hat{x}$ is feasible to (BTLS) and satisfy

$$E(f(\hat{x})) - \bar{p} \leq \frac{4}{7} \overline{p} - \bar{p}.$$ 

Finally, we study a special case of (BTLS) with a zero-centered unit-box, i.e., $l_i + u_i = 0$ and $u_i = 1$ for $i = 1, \ldots, n$. In this case, we show that (BTLS) has a new approximation based on a box-constrained least square relaxation:

$$v(\text{BTLS}) \geq v(\text{LS}) = \min_{x \in [-1,1]^n} \frac{||Ax - b||^2}{n + 1}.$$
Theorem 7. (LS) provides an absolute \((n+1)\)-approximation for (BTLS), that is, 
\[ v(\text{LS}) \leq v(\text{BTLS}) \leq (n+1) \cdot v(\text{LS}). \]

Moreover, suppose (LS) provides a tighter lower bound than the naive unconstrained least square relaxation 
\[ \min \|Ax-b\|^2/(n+1), \] 
then we get an \((n+1)/2\)-approximation for (BTLS):
\[ v(\text{LS}) \leq v(\text{BTLS}) \leq \frac{n+1}{2} v(\text{LS}). \]

Proof. Let \(\bar{x}\) be an optimal solution of the box-constrained least square problem (LS). We have

\[ v(\text{BTLS}) \leq \min_{x^T x \geq \bar{x}^T \bar{x}, x \in [-1,1]^n} \frac{\|Ax-b\|^2}{x^T x + 1} \]
\[ \leq \min_{x^T x \geq \bar{x}^T \bar{x}, x \in [-1,1]^n} \frac{\|Ax-b\|^2}{\bar{x}^T \bar{x} + 1} \]
\[ = \min_{x \in [-1,1]^n} \frac{\|Ax-b\|^2}{x^T x + 1} \]  
(19)
\[ = \frac{n+1}{2} v(\text{LS}), \]  
(20)
\[ \leq (n+1) v(\text{LS}), \]

where the first equality (19) holds since \(\bar{x}\) is an optimal solution of (LS) and hence the constraint \(x^T x \geq \bar{x}^T \bar{x}\) is removable.

For the second part of the result, the assumption implies that \(\bar{x}\) is not in the interior point of the unit-box. That is, there is an index \(k\) such that \(\bar{x}_k \in \{-1,1\}\). Therefore, \(\bar{x}^T \bar{x} \geq 1\). The proof of the second part follows from (20).

Remark 1. Theorem 7 implies that the quadratic programming relaxation \(v(\text{LS})\) is exact if \(v(\text{LS}) = 0\). According to (20), when (LS) has an optimal vertex solution, \(v(\text{LS})\) is also exact.

Remark 2. Suppose (LS) is better than the naive unconstrained relaxation, then \(v(\text{LS})\) is exact when \(n = 1\).

5. Improved SDP relaxation. In this section we strengthen the SDP relaxation (14)-(17). Our approach is to add in (QCQP) more redundant constraints before SDP lifting.

We first relax (12) to the following ball constraint
\[ \left\| \begin{pmatrix} y \\ z \end{pmatrix} \right\| \leq 1. \]

Multiplying both sides each linear constraint in (10), we get the following SOC constraints as in [8, 22]:
\[ \left\| \begin{pmatrix} y \\ z \end{pmatrix} (u_i z - y_i) \right\| \leq u_i z - y_i, \left\| \begin{pmatrix} y \\ z \end{pmatrix} (y_i - l_i z) \right\| \leq y_i - l_i z. \]

Second, multiplying the linear constraint in (10) to each other yields
\[ (u_i z - y_i)(y_j - l_i z) \geq 0, \quad i, j = 1, \ldots, n, \]
where the cases \(i = j\) corresponds to the constraints (13). This approach is known as reformulation-linearization-technique (RLT), see for example, [1].
Lifting \((y^T, z, 1)^T(y^T, z, 1)\) to
\[
\begin{pmatrix}
Y_{n+1} \\
Y_{n+2}
\end{pmatrix} \in \mathbb{R}^{(n+2) \times (n+2)}
\]
yields the following improved SDP relaxation of (BTLS):
\[
\text{(ISDP)} \quad \min \quad \text{Tr}(MY_{n+1})
\text{s.t.} \quad \text{Tr}(Y_{n+1}) = 1,
\]
\[
\begin{align*}
u_i Y_{n+1,j} - Y_{i,j} + l_j Y_{j,n+1} - u_i l_j Y_{n+1,n+1} & \geq 0, \quad i, j = 1, \ldots, n, \\
\|u_i Y_{1:n+1,n+1} - Y_{1:n+1}\| & \leq u_i Y_{n+2,n+1} - Y_{n+2,i}, \quad i = 1, \ldots, n,
\|Y_{1:n+1,i} - l_i Y_{1:n+1,n+1}\| & \leq Y_{n+2,i} - l_i Y_{n+2,n+1}, \quad i = 1, \ldots, n,
\end{align*}
\]
\[Y \succeq 0.\]

Clearly, we have
\[v_{(SDP)} \leq v_{(ISDP)} \leq v_{(BTLS)}.\]

It is unknown when the first inequality becomes strict. We do limited numerical experiments to compare \(v_{(SDP)}\) and \(v_{(ISDP)}\).

We generate the input data \(A\) and \(b\) in (BTLS) using the same MATLAB scripts:

\[
\text{rand}('\text{state}', 1); \quad \text{rand}(m, n); \quad \text{rand}(m, 1);
\]

For each \(i = 1, \ldots, n\), we set \(l_i = 0\) and \(u_i = 50, 100, 255\), respectively. We let \((m, n) \in \{(15, 10), (50, 25), (75, 50), (100, 75), (125, 100), (150, 125)\}\).

All the numerical experiments are constructed in MATLAB R2015b and carried out on a laptop computer with 1.8 GHz processor and 8GB RAM. Both (SDP) and (ISDP) are solved by SDPT3 within CVX [10]. Numerical results are reported in Tables 1 and 2. In all our tested examples, (ISDP) strictly improves (SDP). More significant improvements are observed as the dimension increases. As the cost of this improvement, CPU time rises rapidly with the increase of both dimension and the number of constraints.

**Table 1.** Numerical comparison of lower bounds

| \((m, n)\) |  |  |  |
|---|---|---|---|
|  | \(l = 0, u = 50\) | \(l = 0, u = 100\) | \(l = 0, u = 255\) |
| \((15, 10)\) | \(v_{(SDP)}\) | 0.1847 | 0.1742 | 0.1675 |
|  | \(v_{(ISDP)}\) | 0.1884 | 0.1760 | 0.1683 |
| \((50, 25)\) | 0.4530 | 0.4137 | 0.3903 | 0.4011 |
| \((75, 50)\) | 0.4732 | 0.4214 | 0.3803 | 0.4033 |
| \((100, 75)\) | 0.4732 | 0.4214 | 0.3803 | 0.4033 |
| \((125, 100)\) | 0.4034 | 0.4546 | 0.3893 | 0.4011 |
| \((150, 125)\) | 0.3664 | 0.4256 | 0.3903 | 0.4011 |

**Table 2.** Numerical comparison of CPU time

| \((m, n)\) |  |  |  |
|---|---|---|---|
|  | \(l = 0, u = 50\) | \(l = 0, u = 100\) | \(l = 0, u = 255\) |
| \((15, 10)\) | \(t_{(SDP)}\) | 0.69 | 0.41 | 0.41 |
|  | \(t_{(ISDP)}\) | 2.70 | 1.35 | 1.41 |
| \((50, 25)\) | 0.94 | 0.40 | 4.36 | 4.67 |
| \((75, 50)\) | 1.44 | 0.70 | 19.16 | 20.76 |
| \((100, 75)\) | 2.33 | 0.70 | 66.26 | 70.89 |
| \((125, 100)\) | 4.92 | 2.91 | 176.77 | 239.28 |
| \((150, 125)\) | 9.99 | 686.35 | 564.68 | 647.24 |
6. **Conclusion.** We thoroughly study box-constrained total least squares problem (BTLS). We first prove that (BTLS) is NP-hard. A global optimization algorithm is presented, which shows that (BTLS) is polynomially solvable when the number of dimension is fixed. Suppose the box is centered at zero, a $4/7$-approximate solution is obtained in polynomial time based on SDP relaxation. Moreover, in the unit-box case, the direct least square (LS) relaxation gives an absolute $(n+1)$-approximate solution. If the (LS) relaxation does not reduce to the naive unconstrained LS relaxation, the approximate ratio can be improved to $(n+1)/2$. Finally, we enhance the SDP relaxation by adding SOC and RLT-based constraints. Numerical results show significant improvements of the enhanced SDP relaxation. It is unknown whether a solution with better approximation can be obtained via the enhanced SDP. Future researches include more efficiently solving the improved SDP relaxation and further studying the total least squares problem with a simplex constraint
\[
\min \|Ax - b\|^2 \quad \text{s.t.} \quad \|x\|_1 \leq k,
\]

where $k$ is an integer to control the sparsity of the solution. This problem can be regarded as a total LASSO problem.

**Acknowledgements.** This research was supported by NSFC grants 11822103, 11571029, 11771056 and Beijing Natural Science Foundation Z180005.

**REFERENCES**

[1] K. Anstreicher, Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming, *J. Global Optim.*, 43 (2009), 471–484.

[2] A. Beck, A. Ben-Tal and M. Teboulle, Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares, *SIAM J. Matrix Anal. Appl.*, 28 (2006), 425–445.

[3] A. Beck and M. Teboulle, A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid, *Math. Program.*, 118 (2009), 13–35.

[4] A. Beck and M. Teboulle, On minimizing quadratically constrained ratio of two quadratic functions, *J. Convex Anal.*, 17 (2010), 789–804.

[5] A. Beck and D. Pan, A branch and bound algorithm for nonconvex quadratic optimization with ball and linear constraints, *J. Global Optim.*, 69 (2017), 309–342.

[6] A. Ben-Tal and M. Teboulle, Hidden convexity in some nonconvex quadratically constrained quadratic programming, *Math. Program.*, 72 (1996), 51–63.

[7] D. Bienstock and A. Michalka, Polynomial solvability of variants of the trust-region subproblem, in *Proceedings of the 2014 Annual ACM-SIAM Symposium on Discrete Algorithms* (eds. C. Chekuri), SODA, (2014), 380–390.

[8] S. Burer and K. Anstreicher, Second-order-cone constraints for extended trust-region problems, *SIAM J. Optim.*, 23 (2013), 432–451.

[9] A. Charnes and W. Cooper, Programming with linear fractional functionals, *Nav. Res. Logist. Q.*, 9 (1962), 181–186.

[10] CVX Research, Inc. CVX: Matlab software for disciplined convex programming, version 2.0. [http://cvxr.com/cvx](http://cvxr.com/cvx), April 2011.

[11] W. Dinkelbach, On nonlinear fractional programming, *Manag. Sci.*, 13 (1967), 492–498.

[12] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. Freeman, San Francisco, 1979.

[13] G. Golub and C. Loan, An analysis of the total least-squares problem, *SIAM J. Numer. Anal.*, 17 (1980), 883–893.

[14] M. Goemans and D. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM*, 42 (1995), 1115–1145.
[15] Y. Hsia and R. L. Sheu, Trust region subproblem with a fixed number of additional linear inequality constraints has polynomial complexity, arXiv:1312.1398, 2013.

[16] Z. Luo, W. Ma, A. So, Y. Ye and S. Zhang, Semidefinite relaxation of quadratic optimization problems, IEEE Signal Proc. Mag., 27 (2010), 20–34.

[17] J. Martínez, Local minimizers of quadratic functions on Euclidean balls and spheres, SIAM J. Optim., 4 (1994), 159–176.

[18] Y. Nesterov and A. Nemirovskii, Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms, SIAM Publications, SIAM, Philadelphia, 1993.

[19] Y. Nesterov, Semidefinite relaxation and nonconvex quadratic optimization, Optim. Method Softw., 9 (1998), 141–160.

[20] Y. Nesterov, Introductory Lectures on Convex Optimization: A Basic Course, Kluwer Academic Publishers, 2003.

[21] P. Pardalos and A. Phillips, Global optimization of fractional programs, J. Global Optim., 1 (1991), 173–182.

[22] J. Sturm and S. Zhang, On cones of nonnegative quadratic functions, Math. Oper. Res., 28 (2003), 246–267.

[23] Y. Xia, On minimizing the ratio of quadratic functions over an ellipsoid, Optimization, 64 (2015), 1097–1106.

[24] Y. Xia, A survey of hidden convex optimization, J. Oper. Res. Soc. China., 8 (2020), 1–28.

[25] M. Yang and Y. Xia, On Lagrangian duality gap of quadratic fractional programming with a two-sided quadratic constraint, Optim. Lett., 14 (2020), 569–578.

[26] Y. Ye, Approximating quadratic programming with bound and quadratic constraints, Math. Program., 84 (1999), 219–226.

[27] Y. Ye and S. Zhang, New results on quadratic minimization, SIAM J. Optim., 14 (2003), 245–267.

Received May 2020; 1st revision August 2020; Final revision September 2020.

E-mail address: zhiyuanling@buaa.edu.cn
E-mail address: yxia@buaa.edu.cn
E-mail address: handr@buaa.edu.cn