CrossMark

Simplified full Nesterov–Todd step infeasible interior-point algorithm for semidefinite optimization based on a kernel function

Weiwei Wang[1,2] · **Hongwei Liu**[1] · **Hongmei Bi**[3]

Abstract In this paper, we propose a new complexity analysis of the full Nesterov–Todd step infeasible interior-point algorithm for semidefinite optimization. With a specific feasibility step and the centering step induced by a well-known kernel function, the property of exponential convexity of the kernel function underlying the matrix barrier function is crucial in the analysis and enable us easily to estimate the proximity of iterates to center path. The analysis of the algorithm is simplified and the iteration bound obtained coincides with the currently best iteration bound for infeasible interior-point algorithm.

✉ Weiwei Wang
weiwei.wang2007@163.com

Hongwei Liu
hwliu@mail.xidian.edu.cn

Hongmei Bi
bihongmei2007@163.com

[1] School of Mathematics and Statistics, Xidian University, Xi'an, China

[2] School of Science, Xi'an Technological University, Xi'an, China

[3] School of Science, Air Force Engineering University, Xi'an, China

Springer

1 Introduction

Semidefinite optimization (SDO) problems are convex optimization problems over the intersection of an affine set and the cone of positive semidefinite matrices. They have received considerable attention and become one of the most active research areas in mathematical programming. Due to the success of interior-point methods (IPMs) in solving linear optimization (LO), primal-dual IPMs were extended to SDO, which was an important contribution made by Nesterov and Todd [16,17]. One may distinguish different IPMs by whether they are feasible or infeasible. Feasible IPMs start with a strictly feasible point and maintain feasibility during the solution process. Different from feasible IPMs, interior-point algorithms that accept infeasible initial points are often referred to IIPMs, see more in [9,12,22].

Earlier, all primal-dual IPMs used the Newton direction as the search direction. Peng et al. [19] presented a new class of feasible IPMs. They used a direction determined by a so-called self-regular barrier function. Bai et al. [4] introduced a new class of kernel function. Based on the properties of these kernel functions, they derived many new and tight estimates that greatly simplified the analysis of feasible IPMs. By using an algebraic equivalent transformation of the central path, Darvay [5] explored a new technique in finding a class of search directions. Pan et al. [20] found the connection between Darvay's approach and the method based on kernel functions.

Roos [21] designed a new IIPM for LO based on using the perturbed problems. Since the algorithm uses only full-Newton steps, it has the advantage that no line searches are needed. Later on, Mansouri and Roos [14] extended the proposed algorithm [21] to SDO. Using a new proximity measure to overcome the difficulty of obtaining an upper bound of updated proximity after one full-Newton step, Zhang et al. [30] presented a simplified analysis of the full-Newton step IIPM for SDO. By another definition for feasibility step, Kheirfam [8] generalized the simplified full-Newton step IIPM for LO [13] to SDO.

The aim of this paper is to investigate the kernel function, which has a finite value at the boundary of feasible region:

$$\psi(t) = (t - 1)^2. \tag{1}$$

The search direction determined by the kernel function in (1) is the same as the one introduced by Darvay [5]. Liu and Sun [11] proposed a full Nesterov–Todd(NT) step IIPM for SDO, where the feasibility step was induced by a specific kernel function. The kernel function in (1) is only up to a constant comparing with the one in [11]. It is also discussed by Wang and Bai [26] for feasible IPM, and used by Zhang and Xu [28] for LO to determine the search directions and measure the proximity of iterates to center path. However, the properties of the kernel function, such as exponential convexity, are not discussed in these papers.

Recently, Wang et al. [27] used the property of exponential convexity of the kernel function to simplify the analysis of full-Newton step feasible IPMs for LO problems. In this paper, based on using Nesterov–Todd search direction, the definition of feasibility step by Khierfam [8], and motivated by Wang et al. [27], we not only use the kernel function to determine the centering step, but also apply the property of exponential

convexity of the kernel function to derive new estimates, which simplify the analysis of a full NT-step IIPM for SDO.

The article is organized as follows. In Sect. 2, we introduce the concept of some special matrix functions and basic results of feasible IPM for SDO. In Sect. 3, we present the search directions and describe our algorithm in detail. In Sect. 4, we introduce the kernel function and study the properties of barrier function. In Sect. 5, we analyze the algorithm and derive the complexity bound. In Sect. 6, we give a simple numerical example for the algorithm. Finally, some concluding remarks can be found in Sect. 7.

The following notations are used in the paper. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. The set of all $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m \times n}$. A^T denotes the transpose of $A \in \mathbb{R}^{m \times n}$. The set of all symmetric $n \times n$ real matrices is denoted by $\mathcal{S}^n$. $\mathcal{S}^n_{++}$ ($\mathcal{S}^n_+$) denotes the set of all matrices in $\mathcal{S}^n$ which are positive definite(positive semidefinite). I denotes $n \times n$ identity matrix. $A \succ 0$ ($A \succeq 0$) means that A is positive definite (positive semidefinite). For any $V \in \mathcal{S}^n$, we denote by $\lambda(V)$ the vector of eigenvalues of V arranged in non-increasing order, that is, $\lambda_{\max}(V) = \lambda_1(V) \geq \lambda_2(V) \geq \cdots \geq \lambda_n(V) = \lambda_{\min}(V)$. For any matrix M, we denote by $\sigma_1(M) \geq \sigma_2(M) \geq \cdots \geq \sigma_n(M)$ the singular values of M. For $Q \in \mathcal{S}^n_{++}$, we use $Q^{\frac{1}{2}}$ to denote the symmetric square root of Q. The sign $\sim$ denotes similarity of two matrices. Given $G, H \in \mathbb{R}^{m \times n}$, the inner product between them is defined as $\langle G, H \rangle := \text{Tr}(G^T H)$, the trace of the matrix $G^T H$. The Frobenius norm of $M \in \mathbb{R}^{n \times n}$ is $\|M\| := \langle M, M \rangle^{1/2}$.

2 Preliminaries

2.1 Special matrix functions

Since we need to use some special matrix functions that are useful in the analysis of the algorithm, we review some facts.

Definition 2.1 [25, Definition 2.2] Let $V \in S^n_{++}$ and

$$V = Q^T(\lambda(V))Q,$$

where Q is an arbitrary orthogonal matrix that diagonalizes V, and let $\psi(t)$ be the function in (1). The matrix function $\psi(V) : S^n_{++} \to S^n$ is defined by

$$\psi(V) = Q^T diag(\psi(\lambda_1(V)), \psi(\lambda_2(V)), \ldots, \psi(\lambda_n(V)))Q. \tag{2}$$

Definition 2.2 [25, Definition 2.3] The real value matrix function $\Psi(V) : S^n_{++} \to R_+$ is defined by

$$\Psi(V) = Tr(\psi(V)) = \sum_{i=1}^{n} \psi(\lambda_i(V)). \tag{3}$$

Replacing $\psi(\lambda_i(V))$ in (2) by $\psi'(\lambda_i(V))$, we obtain the matrix function $\psi'(V)$. We call $\Psi(V)$ matrix barrier function and $\psi(t)$ the kernel function for it.

Lemma 2.1 $\Psi(V)$ *is strictly convex with respect to* $V \succ 0$ *and vanishes at its global minimal point* $V = I$, *i.e.,* $\psi(I) = \psi'(I) = 0_{n \times n}$. *Moreover,* $\Psi(I) = 0$.

Proof Since $\psi(t) = (t-1)^2$, we have $\psi'(t) = 2(t-1)$, $\psi''(t) = 2 > 0$. We can see that $\psi(t)$ is strictly convex with respect to $t > 0$ and vanishes at its global minimal point $t = 1$, i.e., $\psi(1) = \psi'(1) = 0$. By using the strictly convex of $\psi(t)$, definitions 2.1 and 2.2, the rest proof of the lemma is similar to the proof of Proposition 5.2.6 (i) in [19]. □

2.2 The central path for SDO

We consider the standard form of the SDO problem:

$$(P) \quad \min\langle C, X \rangle, \quad s.t. \ \langle A_i, X \rangle = b_i, \ i = 1, 2, \ldots, m, \ X \succeq 0, \qquad (4)$$

where $C, X \in \mathcal{S}^n$, $b \in \mathbb{R}^m$, and $A_i \in \mathcal{S}^n$, $i = 1, 2, \ldots, m$. We call problem (P) the primal form of SDO problem, and X is the primal variable.

The dual problem of (P) is given by

$$(D) \quad \max b^T y, \quad s.t. \ \sum_{i=1}^{m} y_i A_i + S = C, \ S \succeq 0, \qquad (5)$$

where $y \in \mathbb{R}^m$ and $S \in \mathcal{S}^n$ are the dual variables.

The set of primal-dual feasible solutions is denoted by

$$\mathcal{F} := \{(X, y, S) \in \mathcal{S}_+^n \times \mathbb{R}^m \times \mathcal{S}_+^n : \langle A_i, X \rangle = b_i, \ i = 1, 2, \ldots, m,$$
$$\sum_{i=1}^{m} y_i A_i + S = C\},$$

and the relative interior of the primal-dual feasible set is

$$\mathcal{F}^0 := \{(X, y, S) \in \mathcal{S}_{++}^n \times \mathbb{R}^m \times \mathcal{S}_{++}^n : \langle A_i, X \rangle = b_i, \ i = 1, 2, \ldots, m,$$
$$\sum_{i=1}^{m} y_i A_i + S = C\}.$$

We assume that both (P) and (D) satisfy the interior point condition (IPC), i.e., $\mathcal{F}^0$ is nonempty and the matrices A_i, $i = 1, 2, \ldots, m$, are linearly independent. It is well known [6] under the IPC the optimality conditions for (P) and (D) can be written as follows:

$$\langle A_i, X \rangle = b_i, i = 1, 2, \ldots, m, X \succeq 0,$$
$$\sum_{i=1}^{m} y_i A_i + S = C, S \succeq 0, \qquad (6)$$
$$XS = 0,$$

where the last equality is called the complementarity equation. The basic idea of primal-dual IPMs is to replace the complementarity condition $XS = 0$ by the parameterized equation $XS = \mu I$ with $\mu > 0$. Then we get the perturbed system

$$
\begin{aligned}
\langle A_i, X \rangle &= b_i, i = 1, 2, \ldots, m, X \succeq 0, \\
\sum_{i=1}^{m} y_i A_i + S &= C, S \succeq 0, \\
XS &= \mu I.
\end{aligned}
\tag{7}
$$

It is proved in [10,15] that there is a unique solution $(X(\mu), y(\mu), S(\mu))$ to the central path equations (7) for any barrier parameter $\mu > 0$. Moreover, the limit point (X^*, y^*, S^*), as μ goes to 0, is a primal-dual optimal solution of the corresponding SDO problem.

2.3 The classic NT search direction

We consider the symmerization scheme yielding the Nesterov–Todd (NT) direction [17], or simply the Newton-direction [23]. The direction is determined by the following system:

$$
\begin{aligned}
\langle A_i, \triangle X \rangle &= 0, i = 1, 2, \ldots, m, \\
\sum_{i=1}^{m} \triangle y_i A_i + \triangle S &= 0, \\
\triangle X + P \triangle S P^T &= \mu S^{-1} - X,
\end{aligned}
\tag{8}
$$

where $(\triangle X, \triangle y, \triangle S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ is the search direction, and the choice of P is

$$
P = X^{1/2}(X^{1/2} S X^{1/2})^{-1/2} X^{1/2} = S^{-1/2}(S^{1/2} X S^{1/2})^{1/2} S^{-1/2}.
\tag{9}
$$

Let $D = P^{1/2}$. Then D can be used to scale X and S to the same matrix V defined by

$$
V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} DSD.
\tag{10}
$$

Obviously

$$
V^2 = \frac{1}{\mu} D^{-1} X S D.
\tag{11}
$$

After we define

$$
\begin{aligned}
\bar{A}_i &:= D A_i D, i = 1, 2, \ldots, m, \\
D_X &:= \frac{1}{\sqrt{\mu}} D^{-1} \triangle X D^{-1}, \\
D_S &:= \frac{1}{\sqrt{\mu}} D \triangle S D,
\end{aligned}
\tag{12}
$$

the system (8) is reduced to

$$\langle \bar{A}_i, D_X \rangle = 0, i = 1, 2, \ldots, m,$$
$$\sum_{i=1}^{m} \triangle y_i \bar{A}_i + D_S = 0, \tag{13}$$
$$D_X + D_S = V^{-1} - V,$$

which gives the usual scaled NT search direction.

3 Infeasible full NT step IPM

We call the triple (X, y, S) an ε-optimal solution of (P) and (D) if the norms of the residuals

$$b_i - Tr(A_i X), i = 1, \ldots, m$$

and

$$C - \sum_{i=1}^{m} y_i A_i - S$$

do not exceed ε and the duality satisfies $Tr(XS) \leq \varepsilon$.

As usual, we assume (P) and (D) have an optimal solution (X^*, y^*, S^*) with $Tr(X^*S^*) = 0$, and we choose the initial iterate (X^0, y^0, S^0) with

$$X^0 = S^0 = \zeta I, y^0 = 0, \mu^0 = \zeta^2, \tag{14}$$

where ζ is a positive number such that

$$X^* + S^* \preceq \zeta I. \tag{15}$$

3.1 The perturbed problems

We denote the initial values of the primal and dual residuals by r_b^0 and R_C^0, respectively:

$$(r_b^0)_i = b_i - Tr(A_i X^0), i = 1, \ldots, m,$$
$$R_C^0 = C - \sum_{i=1}^{m} y_i^0 A_i - S^0. \tag{16}$$

For any ν with $0 < \nu \leq 1$, we consider the perturbed problem (P_ν)

$$min\{\langle C - \nu R_C^0, X \rangle : \langle A_i, X \rangle = b_i - \nu(r_b^0)_i, X \succeq 0\},$$

and its dual problem (D_ν)

$$max \left\{ \sum_{i=1}^{m} (b_i - \nu(r_b^0)_i) y_i : \sum_{i=1}^{m} y_i A_i + S = C - \nu R_C^0, S \succeq 0 \right\}.$$

Note that if $\nu = 1$, then $X = X^0$ yields a strictly feasible solution of (P_ν), and $(y, S) = (y^0, S^0)$ yields strictly feasible solution of (D_ν). We conclude that if $\nu = 1$, then (P_ν) and (D_ν) are strictly feasible. Generally, we have the following lemma.

Lemma 3.1 [14, Lemma 4.1] *Let the original problems (P) and (D) be feasible. Then for each ν satisfying $0 < \nu \le 1$, the perturbed problems (P_ν) and (D_ν) are strictly feasible.*

Assuming that (P) and (D) are both feasible, for each ν such that $0 < \nu \le 1$, Lemma 3.1 implies that the problems (P_ν) and (D_ν) satisfy IPC, then their central paths exist. This means that the system

$$\langle A_i, X \rangle = b_i - \nu(r_b^0)_i, i = 1, 2, \ldots, m, X \succeq 0,$$
$$\sum_{i=1}^{m} y_i A_i + S = C - \nu R_C^0, S \succeq 0, \qquad (17)$$
$$XS = \mu I$$

has a unique solution for every $\mu > 0$. We denote it by $(X(\mu, \nu), y(\mu, \nu), S(\mu, \nu))$, which is the μ-center of (P_ν) and (D_ν). Next, we will always have $\mu = \nu \zeta^2$.

3.2 Search directions

3.2.1 Search directions for centering steps

We define

$$D_X^c := \frac{1}{\sqrt{\mu}} D^{-1} \Delta^c X D^{-1}, D_S^c := \frac{1}{\sqrt{\mu}} D \Delta^c S D. \qquad (18)$$

For centering steps, by the Definitions 2.1 and 2.2, following [19] we replace the right-hand side $V^{-1} - V$ in the third equation in system (13) by $-\psi'(V)$, then the scaled search direction $(D_X^c, \Delta^c y, D_S^c)$ is given by the system

$$\langle \bar{A}_i, D_X^c \rangle = 0, i = 1, 2, \ldots, m,$$
$$\sum_{i=1}^{m} \Delta^c y_i \bar{A}_i + D_S^c = 0, \qquad (19)$$
$$D_X^c + D_S^c = D_V,$$

with $D_V = -\psi'(V)$. The search directions $\Delta^c X$ and $\Delta^c S$ are computed by (18).

3.2.2 Search directions for feasibility step

Following [8] the search direction $(\Delta^f X, \Delta^f y, \Delta^f S)$ is given by the following system:

$$\langle A_i, \Delta^f X \rangle = \theta \nu (r_b^0)_i, \quad i = 1, 2, \ldots, m,$$

$$\sum_{i=1}^{m} \Delta^f y_i A_i + \Delta^f S = \theta \nu R_C^0, \tag{20}$$

$$\Delta^f X + P \Delta^f S P^T = 0,$$

where θ is a fixed barrier update parameter.

We define

$$D_X^f := \frac{1}{\sqrt{\mu}} D^{-1} \Delta^f X D^{-1}, \quad D_S^f := \frac{1}{\sqrt{\mu}} D \Delta^f S D. \tag{21}$$

Then the system (20) can be written as

$$\langle \bar{A}_i, D_X^f \rangle = \frac{1}{\sqrt{\mu}} \theta \nu (r_b^0)_i, \quad i = 1, 2, \ldots, m,$$

$$\sum_{i=1}^{m} \frac{1}{\sqrt{\mu}} \Delta^f y_i \bar{A}_i + D_S^f = \frac{1}{\sqrt{\mu}} \theta \nu D R_C^0 D, \tag{22}$$

$$D_X^f + D_S^f = 0.$$

3.3 An iteration of the algorithm

Initially $X^0 = S^0 = \zeta I$ and $\mu^0 = \zeta^2$, whence $V^0 = I$ and $\Psi(V^0) = 0$. We assume that at the start of each iteration, just before a μ-update, $\Psi(V)$ is smaller than or equal to a (small) threshold value $\tau > 0$. This obviously holds at the start of the first iteration.

Now we describe a main iteration of our algorithm. Suppose that for some $\mu \in (0, 1]$ we have (X, y, S) is strictly feasible for the perturbed problems (P_ν) and (D_ν), where $\mu = \nu \zeta^2$, and such that $\Psi(V) \leq \tau$. We first perform a feasibility step to generate iterates (X^f, y^f, S^f) which is strictly feasible for the perturbed problems $(P_{\nu+})$ and $(D_{\nu+})$, where $\nu^+ = (1 - \theta)\nu$ with $\theta \in (0, 1)$, and close to the μ^+-center of $(P_{\nu+})$ and $(D_{\nu+})$ with $\mu^+ = \nu^+ \zeta^2$, i.e., $\Psi(V^f) = \Psi(X^f, S^f, \mu^+) \leq \tau^f$ for some suitable value of τ^f. After the feasibility step, we perform some centering steps to get a strictly feasible triple (X^+, y^+, S^+) of $(P_{\nu+})$ and $(D_{\nu+})$ such that $\Psi(X^+, S^+, \mu^+) \leq \tau$. The algorithm stops if the norms of the residuals and the duality gap are less than the accuracy parameter ε.

We now summarize the generic full NT-step IIPM for SDO in Fig. 1.

4 The properties of the barrier function

By $\psi'(t) = 2(t - 1)$ we get $D_V = -\psi'(V) = 2(I - V)$. We also use the norm-based proximity measure $\delta(V)$ defined by

$$\delta(V) := \frac{1}{2} \|\psi'(V)\| = \frac{1}{2} \sqrt{\sum_{i=1}^{n} (\psi'(\lambda_i(V)))^2}. \tag{23}$$

Algorithm 1 Full NT-step IIPM

Input:
  An accuracy parameter $\varepsilon > 0$;
  a fixed barrier update parameter $\theta, 0 < \theta < 1$;
  a threshold parameter $\tau > 0$;
  an initialization parameter $\zeta > 0$.
begin:
$X := \zeta I; y = 0; S := \zeta I; \mu := \zeta^2; \nu = 1;$
while $max\{\langle X, S \rangle, \|r_b\|, \|R_C\|\} \geq \varepsilon$ **do**
    feasibility step:
    $(X, y, S) := (X, y, S) + (\Delta^f X, \Delta^f y, \Delta^f S);$
    update of μ and ν:
    $\mu := (1 - \theta)\mu;$
    $\nu := (1 - \theta)\nu;$
    centering steps:
    while $\Psi(V) > \tau$ **do**
      $(X, y, S) := (X, y, S) + (\Delta^c X, \Delta^c y, \Delta^c S);$
    end while
end while
end

Fig. 1 The algorithm

By Lemma 2.1, we have

$$\Psi(V) = 0 \Leftrightarrow \delta(V) = 0 \Leftrightarrow V = I. \tag{24}$$

Lemma 4.1 $\psi(t) = \psi'(t)^2/4$, and $\Psi(V) = \delta(V)^2$ with $\delta(V) = \|D_V\|/2$.

Proof Since $\psi(t) = (t - 1)^2$, it is straightforward that $\psi(t) = \psi'(t)^2/4$. By $\psi(t) = \psi'(t)^2/4$, the definition 2.2 and (23), we obtain

$$\Psi(V) = Tr(\psi(V)) = \sum_{i=1}^{n} \psi(\lambda_i(V)) = \frac{1}{4} \sum_{i=1}^{n} \left(\psi'(\lambda_i(V))\right)^2 = \delta(V)^2.$$

$\square$

In order to discuss the property of exponential convexity, we need the following lemma. The lemma is a direct consequence of conclusion (d) of Theorem 3.3.14 in [7, pp.176–177]. We cite it here without proof, see [7] for more details.

Lemma 4.2 *Let* $M, N \in S^n$ *be two nonsingular matrices and* $f(t)$ *a real-valued function such that* $f(e^t)$ *is a convex function. Then*

$$\sum_{i=1}^{n} f(\sigma_i(MN)) \leq \sum_{i=1}^{n} f(\sigma_i(M)\sigma_i(N)),$$

where $\sigma_i(M), \sigma_i(N)$ *and* $\sigma_i(MN), i = 1, 2, \ldots, n,$ *denote the singular values of* M, N *and* MN *arranged in non-increasing order, respectively.*

Lemma 4.3 *Let $t_1 \geq 1/2$ and $t_2 \geq 1/2$. Then*

$$\psi(\sqrt{t_1 t_2}) \leq (\psi(t_1) + \psi(t_2))/2.$$

Proof By some simple calculation, the above inequality is equivalent to $\sqrt{t_1} + \sqrt{t_2} \geq \sqrt{2}$. The expression holds if $t_1, t_2 \geq 1/2$. □

Similarly as in [3,4], we say that $\psi(t)$ is e-convex whenever $t \geq 1/2$. It is easily verified that $\psi(e^t)$ is a convex function, so we can apply the Lemma 4.2 to $\psi(t)$.

Lemma 4.4 *Suppose that matrices V_1, V_2 are symmetric positive definite. Let $\lambda_{min}(V_1) \geq 1/2$ and $\lambda_{min}(V_2) \geq 1/2$, then*

$$\Psi\left([V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}}]^{\frac{1}{2}}\right) \leq \frac{1}{2}\left(\Psi(V_1) + \Psi(V_2)\right).$$

Proof By the definition of the singular values of a matrix, for any nonsingular matrix $V \in S^n$,

$$\sigma_i(V) = \left(\lambda_i(V^T V)\right)^{\frac{1}{2}} = \left(\lambda_i(V V^T)\right)^{\frac{1}{2}}, i = 1, 2, \ldots, n.$$

Hence,

$$\sigma_i\left(V_1^{\frac{1}{2}} V_2^{\frac{1}{2}}\right) = \left(\lambda_i(V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}})\right)^{\frac{1}{2}} = \lambda_i\left([V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}}]^{\frac{1}{2}}\right), i = 1, 2, \ldots, n.$$

Since V_1, V_2 are symmetric positive definite, we have

$$\sigma_i(V_1) = \lambda_i(V_1), \sigma_i(V_2) = \lambda_i(V_2), i = 1, 2, \ldots, n.$$

By the Definition 2.2, Lemmas 4.2 and 4.3, we obtain

$$\Psi\left([V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}}]^{\frac{1}{2}}\right) = \sum_{i=1}^{n} \psi\left(\sigma_i(V_1^{\frac{1}{2}} V_2^{\frac{1}{2}})\right)$$

$$\leq \sum_{i=1}^{n} \psi\left(\sigma_i(V_1^{\frac{1}{2}})\sigma_i(V_2^{\frac{1}{2}})\right)$$

$$\leq \frac{1}{2}\sum_{i=1}^{n}\left(\psi\left(\sigma_i^2(V_1^{\frac{1}{2}})\right) + \psi\left(\sigma_i^2(V_2^{\frac{1}{2}})\right)\right)$$

$$= \frac{1}{2}\sum_{i=1}^{n}(\psi(\sigma_i(V_1)) + \psi(\sigma_i(V_2)))$$

$$= \frac{1}{2}\left(\Psi(V_1) + \Psi(V_2)\right).$$

□

Lemma 4.5 $\|V\| \leq \sqrt{n} + \delta(V)$.

Proof It follows from $\|V\| = \|I - D_V/2\| \leq \|I\| + \|D_V\|/2$. □

Lemma 4.6 $1 - \delta(V) \leq \lambda_i(V) \leq 1 + \delta(V), 1 \leq i \leq n$.

Proof By Lemma 4.1 we have $\delta(V) = \sqrt{\Psi(V)} = \sqrt{\sum_{i=1}^n (1 - \lambda_i(V))^2}$, then $|1 - \lambda_i(V)| \leq \delta(V)$, so the lemma follows. □

The following lemma shows the effect of a μ-update on the value of $\Psi(V)$.

Lemma 4.7 *Let* $0 < \theta < 1$, *for any positive definite matrix* V:

$$\Psi\left(\frac{V}{\sqrt{1-\theta}}\right) \leq \Psi(V) + \frac{\theta}{1-\theta}\|V\|^2. \tag{25}$$

Proof Let $\beta \geq 1$, define $v_i := \lambda_i(V), 1 \leq i \leq n$, then $v > 0$. From the definition of $\Psi(V)$, we have

$$\Psi(\beta V) = \sum_{i=1}^n \psi(\beta \lambda_i(V)) = \sum_{i=1}^n \psi(\beta v_i) = \Psi(\beta v).$$

By the definition of $\psi(t)$ in (1),

$$\psi(\beta t) = (t-1)^2 + (\beta^2 - 1)t^2 + 2(1-\beta)t \leq \psi(t) + (\beta^2 - 1)t^2. \tag{26}$$

Let $\beta = 1/\sqrt{1-\theta}$, then

$$\begin{aligned}
\Psi\left(\frac{V}{\sqrt{1-\theta}}\right) = \Psi\left(\frac{v}{\sqrt{1-\theta}}\right) &\leq \Psi(v) + \frac{\theta}{1-\theta}\|v\|^2 \\
&= \sum_{i=1}^n \psi(v_i) + \frac{\theta}{1-\theta}\sum_{i=1}^n v_i^2 \\
&= \sum_{i=1}^n \psi(\lambda_i(V)) + \frac{\theta}{1-\theta}\sum_{i=1}^n (\lambda_i(V))^2 \\
&= \Psi(V) + \frac{\theta}{1-\theta}\|V\|^2.
\end{aligned}$$

□

5 Analysis of the algorithm

5.1 Analysis of the centering step

After a centering step, the new iterates are given by

$$X^+ := X + \Delta^c X, \ y^+ := y + \Delta^c y, \ S^+ := S + \Delta^c S.$$

The new V-matrix is

$$(V^+)^2 = \frac{1}{\mu} D^{-1} X^+ S^+ D. \tag{27}$$

The following results are crucial in the analysis of the algorithm. We list them without proof.

Lemma 5.1 [29, Lemma 4.2] *The new iterates (X^+, y^+, S^+) are strictly feasible if $\Psi(V) < 1$.*

Theorem 5.1 [29, Theorem 4.5] *If $\Psi(V) < 1$, then $\Psi(V^+) < \Psi(V)^2$.*

Theorem 5.1 implies that the centering step is quadratically convergent.

5.2 Analysis of the feasibility step

After a feasibility step, the new iterates are given by

$$X^f := X + \triangle^f X, \, y^f := y + \triangle^f y, \, S^f := S + \triangle^f S. \tag{28}$$

By (20) and (28), it is easy to show that the iterates are feasible for the new perturbed problems. A key point in the analysis is to show that X^f and S^f are positive definite and satisfy $\Psi(V^f) \le \tau^f (\tau^f < 1)$.

By using (21), we obtain

$$\triangle^f X = \sqrt{\mu} D D_X^f D, \, \triangle^f S = \sqrt{\mu} D^{-1} D_S^f D^{-1}.$$

It follows from (28) that

$$\begin{aligned}
X^f &= X + \triangle^f X = \sqrt{\mu} D \left(V + D_X^f \right) D, \\
S^f &= S + \triangle^f S = \sqrt{\mu} D^{-1} \left(V + D_S^f \right) D^{-1}.
\end{aligned} \tag{29}$$

Therefore

$$X^f S^f = \mu D \left(V + D_X^f \right) \left(V + D_S^f \right) D^{-1}, \tag{30}$$

which implies that

$$X^f S^f \sim \mu \left(V + D_X^f \right) \left(V + D_S^f \right). \tag{31}$$

Let

$$\begin{aligned}
D_{XS}^f &:= \tfrac{1}{2} \left(D_X^f D_S^f + D_S^f D_X^f \right), \\
M &:= \left(D_X^f V - V D_X^f \right) + \tfrac{1}{2} \left(D_X^f D_S^f - D_S^f D_X^f \right).
\end{aligned} \tag{32}$$

Note that D_{XS}^f is a symmetric matrix and M a skew-symmetric matrix. From the third equation in system (22), by multiplying both sides from the left with V, we get

$$V D_X^f + V D_S^f = 0. \tag{33}$$

Therefore

$$
\begin{aligned}
\frac{X^f S^f}{\mu} &\sim \left(V + D_X^f\right)\left(V + D_S^f\right) \\
&= V^2 + V D_S^f + D_X^f V + D_X^f D_S^f \\
&= V^2 + \frac{1}{2}\left(D_X^f D_S^f + D_S^f D_X^f\right) + \left(D_X^f V - V D_X^f\right) \\
&\quad + \frac{1}{2}\left(D_X^f D_S^f - D_S^f D_X^f\right) \\
&= V^2 + D_{XS}^f + M.
\end{aligned}
\tag{34}
$$

The following lemma shows the sufficient condition of the strict feasibility of the new iterates (X^f, y^f, S^f).

Lemma 5.2 *Let $X \succ 0$ and $S \succ 0$. Then the iterates (X^f, y^f, S^f) are strictly feasible if*

$$
\lambda_{min}(V) > \left|\lambda_i\left(D_X^f\right)\right|, \ 1 \leq i \leq n.
$$

Proof By Lemma 10 in [8], (X^f, y^f, S^f) are strictly feasible if $V^2 + D_{XS}^f \succ 0$. Using $D_S^f = -D_X^f$ we have $\lambda_i\left(D_{XS}^f\right) = -\lambda_i^2\left(D_X^f\right)$, therefore $V^2 + D_{XS}^f \succ 0$ if $\lambda_{min}(V) > \left|\lambda_i\left(D_X^f\right)\right|, \ 1 \leq i \leq n$. This completes the proof. □

Before μ-update, by (10), the new V-matrix is given by

$$
\hat{V} = \frac{1}{\sqrt{\mu}}\left(D^{-1} X^f S^f D\right)^{\frac{1}{2}}.
\tag{35}
$$

Now we want to find an upper bound for $\Psi(\hat{V})$ by applying exponential convexity, so we assume the eigenvalues of the matrices $V + D_X^f$ and $V + D_S^f$ are at least $1/2$, i.e.

$$
\lambda_i\left(V + D_X^f\right) \geq \frac{1}{2} \ and \ \lambda_i\left(V + D_S^f\right) \geq \frac{1}{2}, \ 1 \leq i \leq n.
\tag{36}
$$

It is obvious that (36) holds if

$$
\lambda_{min}(V) - \left|\lambda_i\left(D_X^f\right)\right| \geq \frac{1}{2}, \ 1 \leq i \leq n.
\tag{37}
$$

By (35) we have $\hat{V}^2 \sim \frac{1}{\mu} X^f S^f \sim \frac{1}{\mu}\left(X^f\right)^{\frac{1}{2}} S^f \left(X^f\right)^{\frac{1}{2}}$, from (31),

$$
\hat{V}^2 \sim \bar{V}^2 := \left(V + D_X^f\right)^{\frac{1}{2}}\left(V + D_S^f\right)\left(V + D_X^f\right)^{\frac{1}{2}}.
\tag{38}
$$

Lemma 5.3 $\Psi(\hat{V}) \leq \Psi(V) + \left\| D_X^f \right\|^2$.

Proof Using (38), the definition of $\psi(V)$ and $\Psi(V)$, Lemma 4.4 and $D_S^f = -D_X^f$, we have

$$
\begin{aligned}
\Psi(\hat{V}) = \Psi(\bar{V}) &= \Psi\left(\left((V + D_X^f)^{\frac{1}{2}}(V + D_S^f)(V + D_X^f)^{\frac{1}{2}} \right)^{\frac{1}{2}} \right) \\
&\leq \frac{1}{2}\left(\Psi(V + D_X^f) + \Psi(V + D_S^f) \right) \\
&= \frac{1}{2}\left(Tr\left(\psi(V + D_X^f) \right) + Tr\left(\psi(V + D_S^f) \right) \right) \\
&= \frac{1}{2}Tr\left(\psi(V + D_X^f) + \psi(V + D_S^f) \right) \\
&= \frac{1}{2}Tr\left(2(V - I)^2 + (D_X^f + D_S^f)(V - I) \right. \\
&\quad + (V - I)(D_X^f + D_S^f) + (D_X^f)^2 + (D_S^f)^2 \Big) \\
&= \frac{1}{2}Tr\left(2(V - I)^2 + 2(D_X^f)^2 \right) \\
&= Tr\left((V - I)^2 \right) + Tr\left((D_X^f)^2 \right) \\
&= Tr(\psi(V)) + Tr\left((D_X^f)^2 \right) \\
&= \Psi(V) + \left\| D_X^f \right\|^2.
\end{aligned}
$$

$\square$

5.3 Upper bound for $\left\| D_X^f \right\|$

By a similar argument given in [8, Section 5], we have

$$
\left\| D_X^f \right\| \leq \frac{\theta}{\zeta \lambda_{min}(V)} Tr(X + S). \tag{39}
$$

Next, we denote $\delta(V)$ simply as δ.

Lemma 5.4
$$
\left\| D_X^f \right\| \leq \frac{(\sqrt{n} + \delta)^2 + n}{1 - \delta}\theta. \tag{40}
$$

Proof In [14, Lemma 5.15], it is shown that

$$
v\zeta Tr(X + S) \leq \langle X, S \rangle + vn\zeta^2. \tag{41}
$$

By $XS \sim \mu V^2$ and Lemma 4.5 we have

$$
Tr(XS) = \mu\|V\|^2 \leq v\zeta^2(\sqrt{n} + \delta)^2. \tag{42}
$$

The lemma follows from (39) and Lemma 4.6.   □

By (37) and Lemma 4.6 we obtain that the e-convex holds if $\left|\lambda_i (D_X^f)\right| \leq \left\|D_X^f\right\| \leq 1/2 - \delta, 1 \leq i \leq n$, then using Lemma 5.4 we have

$$\theta \leq \frac{(1/2 - \delta)(1 - \delta)}{(\sqrt{n} + \delta)^2 + n}. \tag{43}$$

Note that the right-hand side of the above inequality is monotonically decreasing with respect to δ.

5.4 The choice of τ, θ and τ^f

Theorem 5.2 *Let* $\mu^+ = (1 - \theta)\mu$, *where* $0 < \theta < 1$ *and* $V^f = \frac{\hat{V}}{\sqrt{1-\theta}}$, *then*

$$\Psi(V^f) \leq \frac{1}{1 - \theta}\left(\Psi(V) + 2\theta\sqrt{n\Psi(V)} + \theta n + (1 - 2\theta)/4\right). \tag{44}$$

Particularly, if $\Psi(V) \leq \tau = \frac{1}{16}$ *and* $\theta = \frac{1}{13n}$, *then*

$$\Psi(V^f) \leq \tau^f = \frac{1}{2}.$$

Proof Before μ-update, by Lemma 5.3 we have $\Psi(\hat{V}) \leq \Psi(V) + \left\|D_X^f\right\|^2$. From (34) and (35),

$$\hat{V}^2 \sim \left(V^2 + D_{XS}^f + M\right). \tag{45}$$

Then

$$\|\hat{V}\|^2 = Tr\left(\hat{V}^2\right) = Tr\left(V^2 + D_{XS}^f + M\right). \tag{46}$$

Since M is skew-symmetric, and $D_S^f = -D_X^f$, we obtain

$$\|\hat{V}\|^2 = Tr\left(V^2\right) + Tr\left(D_{XS}^f\right) = \|V\|^2 - Tr\left((D_X^f)^2\right) = \|V\|^2 - \left\|D_X^f\right\|^2. \tag{47}$$

After μ-update, by Lemma 4.7 we get

$$\begin{aligned}\Psi(V^f) &\leq \Psi(\hat{V}) + \frac{\theta}{1 - \theta}\|\hat{V}\|^2 \\ &\leq \Psi(V) + \left\|D_X^f\right\|^2 + \frac{\theta}{1 - \theta}\left(\|V\|^2 - \left\|D_X^f\right\|^2\right).\end{aligned} \tag{48}$$

Since $\left\|D_X^f\right\|^2 \leq (1/2 - \delta)^2 \leq 1/4$, by Lemmas 4.1 and 4.5, the inequality in (44) follows. If $\Psi(V) \leq \tau = 1/16$, by Lemma 4.1 $\delta(V) \leq 1/4$, hence if we take $\theta =$

$1/(13n)$, the inequality (43) holds. Then $\Psi(V^f) \leq \tau^f = 1/2$ follows from (44) and simple calculation. □

5.5 Complexity analysis

After the feasibility step, we have derived an upper bound for the $\Psi(V^f)$, i.e. $\Psi(V^f) \leq 1/2$. By the quadratic convergence property of the centering step (Theorem 5.1), at most

$$\log_2(\log_2 1/\tau) = \log_2(\log_2 16) = 2$$

centering steps suffice to get iterates (X^+, y^+, S^+) that satisfy $\Psi(X^+, S^+; \mu^+) \leq 1/16$. So each main iteration consists of at most 3 so-called inner iterations, in each of which we need to compute a new search direction.

In each main iteration both the duality gap and the norms of the residuals are reduced by the factor $1 - \theta$. Hence, using $Tr(X^0 S^0) = n\zeta^2$, the total number of main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{max\{n\zeta^2, \|r_b^0\|, \|R_C^0\|\}}{\varepsilon}.$$

Since

$$\theta = \frac{1}{13n},$$

the total number of inner iterations is bounded above by

$$39n \log \frac{max\{n\zeta^2, \|r_b^0\|, \|R_C^0\|\}}{\varepsilon}.$$

Now we state our main result without further proof.

Theorem 5.3 *If (P) and (D) have an optimal solution (X^*, y^*, S^*) such that $X^* + S^* \leq \zeta I$, then after at most*

$$39n \log \frac{max\{n\zeta^2, \|r_b^0\|, \|R_C^0\|\}}{\varepsilon}$$

iterations, the algorithm finds an ε-solution of (P) and (D).

6 Numerical results

In this section we consider an application example and give some numerical results.

Maximum eigenvalue minimization Suppose the symmetric matrix $A(x)$ depends affinely on $x \in \mathbb{R}^k : A(x) = A_0 + x_1 A_1 + \cdots + x_k A_k$, where $A_i \in S^n$. The problem of

minimizing the maximum eigenvalue of the matrix $A(x)$ can be cast as the semidefinite program

$$\min \quad t, \quad s.t. \quad tI - A(x) \geq 0, \tag{49}$$

with variables $x \in \mathbb{R}^k$ and $t \in \mathbb{R}$. Problems of this type arise in control theory, structural optimization, graph theory and combinatorial optimization, and other fields (see [24] for details).

Example 1 [1, Example 14.1] Find scalars y_1, y_2, and y_3 such that the maximum eigenvalue of $F = A_0 + y_1 A_1 + y_2 A_2 + y_3 A_3$ with

$$A_0 = \begin{bmatrix} 2 & -0.5 & -0.6 \\ -0.5 & 2 & 0.4 \\ -0.6 & 0.4 & 3 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

is minimized.

This problem can be formulated as the standard form of the SDO problem

$$\max b^T y, \quad s.t. \sum_{i=1}^{4} y_i A_i + S = C, \ S \succeq 0, \tag{50}$$

where $b = [0, 0, 0, 1]^T$, $y = [y_1, y_2, y_3, y_4]^T$, $C = -A_0$, $A_4 = I$, by (49) we observe that $-y_4$ is the maximum eigenvalue of matrix F.

The initial point is $(X^0, y^0, S^0) = (\zeta I, 0, \zeta I)$ with $\zeta = 1$. We set $\tau = \frac{1}{16}$, $\theta = \frac{1}{13n}$ and $\varepsilon = 10^{-3}$. A primal-dual optimal solution is given by

$$X^* = \begin{bmatrix} 0.0009 & -0.0000 & -0.0000 \\ -0.0000 & 0.0009 & 0.0000 \\ -0.0000 & 0.0000 & 1.0000 \end{bmatrix}, y^* = \begin{bmatrix} 0.4996 \\ 0.5995 \\ -0.3996 \\ -2.9973 \end{bmatrix},$$

$$S^* = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & -0.0000 \\ 0.0000 & -0.0000 & 0.0009 \end{bmatrix}.$$

The algorithm reaches this solution in 271 iterations, taking 2.0836 seconds. It might be emphasized that the default value $\theta = \frac{1}{13n}$ is theoretically justified in a worst-case. If a larger θ is taken, for example $\theta = \frac{1}{4n}$, the algorithm need just 80 iterations, taking 0.5817 seconds.

By using the first three components of y^*, that is y_1, y_2, and y_3, the maximum eigenvalue of $F = A_0 + y_1 A_1 + y_2 A_2 + y_3 A_3$ is found to be 2.9973.

7 Concluding remarks

Based on the exponential convexity of a simple kernel function, we propose a new complexity analysis of a full NT-step infeasible interior-point algorithm for SDO. The iteration bound of the algorithm coincides with the best known iteration bound for IIPMs. Moreover, the resulting analysis is relatively simple comparing to that in [8,11]. However, from a practical perspective, the full NT-step IIPM may be not so efficient. Recently, by applying the properties of the kernel function-based barrier function discussed in [18], Asadi and Roos [2] attempted to design a large-update version of Roos' algorithm [21] for LO problems. Next, we may focus on designing a large-update IIPM involving the kernel function for SDO problems.

References

1. Antoniou, A., Lu, W.S.: Practical Optimization: Algorithms and Engineering Applications. Springer, Boston (2007)
2. Asadi, A., Roos, C.: Infeasible interior-point methods for linear optimization based on large neighborhood. J. Optim. Theory Appl. **170**(2), 562–590 (2016)
3. Bai, Y.Q., El Ghami, M., Roos, C.: A new efficient large-update primal-dual interior-point method based on a finite barrier. SIAM J. Optim. **13**(3), 766–782 (2003)
4. Bai, Y.Q., El Ghami, M., Roos, C.: A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. SIAM J. Optim. **15**(1), 101–128 (2004)
5. Darvay, Zs: New interior-point algorithms in linear programming. Adv. Model. Optim. **5**(1), 51–92 (2003)
6. De Klerk, E.: Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Kluwer, Dordrecht (2002)
7. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991/1994)
8. Kheirfam, B.: Simplified infeasible interior-point algorithm for SDO using full Nesterov–Todd step. Numer. Algorithms **59**(4), 589–606 (2012)
9. Kojima, M., Megiddo, N., Mizuno, S.: A primal-dual infeasible-interior-point algorithm for linear programming. Math. Program. **61**, 263–280 (1993)
10. Kojima, M., Shindoh, S., Hara, S.: Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. SIAM J. Optim. **7**, 86–125 (1997)
11. Liu, Z., Sun, W.: A full-NT-step infeasible interior-point algorithm for SDP based on kernel functions. Appl. Math. Comput. **217**, 4990–4999 (2011)
12. Lustig, I.J.: Feasible issues in a primal-dual interior point method for linear programming. Math. Program. **49**, 145–162 (1990/1991)
13. Mansouri, H., Roos, C.: Simplified $O(nL)$ infeasible interior-point algorithm for linear optimization using full-Newton steps. Optim. Methods Softw. **22**(3), 519–530 (2007)
14. Mansouri, H., Roos, C.: A new full-Newton step $O(n)$ infeasible interior-point algorithm for semidefinite optimization. Numer. Algorithms **52**, 225–255 (2009)
15. Nesterov, Y.E., Nemirovsky, A.S.: Interior Point Methods in Convex Programming: Theory and Applications. SIAM, Philadelphia (1994)
16. Nesterov, Y.E., Todd, M.J.: Self-scaled barriers and interior-point mehods for convex program-ming. Math. Oper. Res. **22**, 1–42 (1997)
17. Nesterov, Y.E., Todd, M.J.: Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**, 324–364 (1998)
18. Peng, J., Roos, C., Terlaky, T.: A new and efficient large-update interior-point method for linear optimization. J. Comput. Tech. **6**, 61–80 (2001)
19. Peng, J., Roos, C., Terlaky, T.: Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms. Princeton University Press, Princeton (2002)

20. Pan, S., Li, X., He, S.: An infeasible primal-dual interior-point algorithm for linear programs based on logarithmic equivalent transformation. J. Math. Anal. Appl. **314**(2), 644–660 (2006)
21. Roos, C.: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. **16**(4), 1110–1136 (2006)
22. Tanabe, K.: Centered Newton method for linear programming: interior and exterior point method (in Janpanese). In: Tone, K. (ed.) New Methods for Linear Programming, vol. 3, pp. 98–100. The Institute of Statistical Mathematics, Tokyo (1990)
23. Todd, M.J., Toh, K.C., Tütüncü, R.H.: On the Nesterov–Todd direction in semidefinite programming. SIAM J. Optim. **8**(3), 769–796 (1998)
24. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Rev. **38**(1), 49–95 (1996)
25. Wang, G.Q., Bai, Y.Q., Roos, C.: Primal-dual interior-point algorithm for semidefinite optimization based on a simple kernel function. J. Math. Model. Algorithms **4**, 409–433 (2005)
26. Wang, G.Q., Bai, Y.Q.: A new primal-dual path-following interior-point algorithm for semidefinite optimization. J. Math. Anal. Appl. **353**(1), 339–349 (2009)
27. Wang, W., Bi, H., Liu, H.: A full-Newton step interior-point algorithm for linear optimization based on a finite barrier. Oper. Res. Lett. **44**(6), 750–753 (2016)
28. Zhang, L.P., Xu, Y.H.: A new infeasible interior-point algorithm with full step for linear optimization based on a simple function. Int. J. Comput. Math. **88**(15), 3163–3185 (2011)
29. Zhang, L.P.: Full-Newton Step Infeasible Interior-point Algorithm for Conic Programming. Ph.D. Thesis. Shanghai University, Shanghai (2011)
30. Zhang, L.P., Sun, L.M., Xu, Y.H.: Simplified analysis for full-Newton step infeasible interior-point algorithm for semidefinite programming. Optim. **62**(2), 169–191 (2013)