On Duality Theory for Non-Convex Semidefinite Programming *

Wenyu Sun[†] Chengjin Li[‡] and Raimundo J.B. Sampaio[§]

Abstract

In this paper, with the help of convex-like function, we discuss the duality theory for nonconvex semidefinite programming. Our contributions are: duality theory for the general nonconvex semidefinite programming when Slater's condition holds; perfect duality for a special case of the nonconvex semidefinite programming for which Slater's condition fails. We point out that the results of [2] can be regarded as a special case of our result.

Mathematics Subject Classifications: 65K05, 90C22, 90C26,
Key words. semidefinite programming, nonconvex semidefinite programming, duality, convex-like function, invex function

1. Introduction

In this paper, we use X, S^n, S^n_+ to denote the finite dimensional space, the space of symmetric $n \times n$ matrices and the cone of all symmetric positive semidefinite matrices respectively, and denote $C \subset X$ a subspace. We consider the following nonconvex semidefinite programming problem

$$\begin{aligned} \min_{x\in C\subset X} \quad & f(x) \\ \text{s.t.} \quad & g(x) \le 0, \\ & G(x) \preceq 0, \end{aligned} \tag{1.1}$$

*This work was supported by National Natural Science Foundation of China (grant No. 10871098), the Natural Science Fund of Jiangsu Province (grant No. BK2009397), and CNPq of Brazil.

†Corresponding author. School of Mathematical Sciences, Nanjing Normal University, Nanjing 210046, China. Email: wysun@njnu.edu.cn

‡School of Mathematical Sciences, Nanjing Normal University, Nanjing 210046, China; School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China. Email: chengjin98298@sina.com

§Pontifical Catholic University of Panara (PUCPR), Graduate Program in Production and Systems Engineering (PPGEPS), CEP 81611-970 Curitiba, Parana, Brazil. Email: rsampaio@ppgia.pucpr.br

where $f : C \to R$, $g : C \to R^m$ and $G : C \to S^n$ may not be convex functions, and $G(x) \preceq 0$ if and only if $-G(x) \in S^n_+$. Several numerical methods for nonlinear and nonconvex semidefinite programming are presented (see [4, 9, 6, 11, 7, 8]). In this paper our concern is on duality theory for nonconvex semidefinite programming (1.1).

There are two special cases of problem (1.1). When the functions f, g and G in (1.1) are convex, $C = X = S^n$ and $G(x) \equiv -x$, then the optimization problem (1.1) becomes the problem (NSDP) in [2] as follows:

$$\begin{array}{ll} \min_{x \in S^n} & f(x) \\ \text{s.t.} & -g(x) \geq 0, \\ & x \succeq 0. \end{array}$$

In this paper, under the conditions that are different from and weaker than [2], we prove that the properties about (NSDP) in [2] are also satisfied by the nonconvex semidefinite programming (1.1). So, the conclusions obtained in this paper are both a generalization and enhancement of the results of [2].

Another special case of (1.1) is

$$\begin{array}{ll} \min_{x \in X} & f(x) \\ \text{s.t.} & h(x) = 0, \\ & H(x) \preceq 0. \end{array} \tag{1.2}$$

We can reformulate the constraint $h(x) = 0$ in (1.2) as $h(x) \leq 0$ and $-h(x) \leq 0$, so there are two ways to transform the problem (1.2) into (1.1):

Case I), let $g(x) = (h(x)^\top, -h(x)^\top)^\top$ and $G(x) = H(x)$;

Case II), let $g(x) = h(x)$ and

$$G(x) = \begin{pmatrix} -\Lambda_{h(x)} & 0 \\ 0 & H(x) \end{pmatrix},$$

where $\Lambda_{h(x)}$ is a diagonal matrix with its i-th diagonal element $h_i(x)$ which is the i-th entry of the vector $h(x)$.

Optimization problem (1.2) is an important case of semidefinite programming (see [10, 11]). Unfortunately, the Slater's condition does not hold in problem (1.2) when we consider it as a special case of problem (1.1). In this case we employ the new way of [1] to get the perfect duality of nonconvex semidefinite programming (1.2) when the Slater's condition fails.

The organization of this paper is as follows. In Section 2, we give some notation and preliminaries. The duality theory for general nonconvex semidefinite programming (1.1) is given in Section 3 when the Slater's condition holds. In Section 4, we establish the perfect duality for the special case (1.2) of nonconvex semidefinite programming when the Slater's condition fails. Finally, we give a conclusion in Section 5.

2. Notation and Preliminaries

In this section, we introduce some notation and definitions. It is natural to equip S^n with the inner product of the vectors representing the matrices:

$$A \bullet B = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij} = Tr(AB), \quad \forall A, B \in S^n,$$

where Tr stands for trace. By using the definition of the inner product, we can define the F(robenius)-norm of a matrix:

$$\| A \|_F = \sqrt{A \bullet A} = \sqrt{Tr(AA)}, \quad \forall A \in S^n.$$

The *generalized inner product* of points $P_1 = (y_1, A_1)$ and $P_2 = (y_2, A_2) \in R^m \times S^n$ is defined as $\langle P_1, P_2 \rangle = y_1^\top y_2 + A_1 \bullet A_2$ and the norm of point $P = (y, A) \in R^m \times S^n$ is defined as

$$\| P \| = \langle P, P \rangle^{\frac{1}{2}} = (\| y \|_2^2 + \| A \|_F^2)^{\frac{1}{2}}.$$

We reserve Z to represent S^i, R^j and $R^j \times S^i$, where i and j are arbitrary positive integers, and when Z represents S^i ($R^j, R^j \times S^i$, respectively), then Z_+ represents S^i_+ ($R^j_+, R^j_+ \times S^i_+$, respectively).

Under the assumptions above, it is natural to define that, for all $A, B \in Z$,

$$A \geq B \Longleftrightarrow A - B \in Z_+,$$

$$A > B \Longleftrightarrow A - B \in intZ_+.$$

Next, we introduce two kinds of nonconvex functions in [3, 12, 13, 14] as follows.

Definition 2.1 *A function $T : C \to Z$ is said to be convex-like function if for all $x, y \in C, \mu \in [0, 1]$, there exists $z \in C$ such that*

$$T(z) \leq \mu T(x) + (1 - \mu) T(y).$$

Definition 2.2 *A set $C \subseteq X$ is said to be invex if there exists a vector function $\eta : X \times X \to X$ such that*

$$x, y \in C, \mu \in [0, 1] \Longrightarrow y + \mu \eta(x, y) \in C.$$

Definition 2.3 *Let $C \subseteq X$ be an invex set with respect to $\eta : X \times X \to X$ and let $T : C \to Z$. We say that T is a pre-invex function relative to η if*

$$T(y + \mu \eta(x, y)) \leq \mu T(x) + (1 - \mu) T(y), \quad \forall x, y \in C, \ \mu \in [0, 1].$$

It is obvious that every pre-invex function is convex-like function, and every convex function is pre-invex function (i.e., let $\eta(x,y) = x - y$), but the reverse is not true. For an example in [13], we consider the function $f : R \to R$ defined by $f(x) = -|x|$, then f, instead of a convex function, is a pre-invex function with η given by

$$\eta(x,y) = \begin{cases} x - y, & \text{if } x \le 0, y \le 0, \\ x - y, & \text{if } x \ge 0, y \ge 0, \\ y - x, & \text{Otherwise.} \end{cases}$$

The following two kinds of Farkas lemma can be found in [5].

Theorem 2.4 *Let $T(x) = (g(x), G(x)) : C \to R^m \times S^n$ be a convex-like function on C, then the following system*

$$\begin{cases} g(x) < 0, \\ G(x) < 0, \\ x \in C \end{cases} \tag{2.1}$$

has no solution if and only if there exists $(\lambda, U) \in R^m \times S^n$ with $\lambda \ge 0, U \ge 0$ and $(\lambda, U) \ne (0,0)$ such that

$$\lambda^\top g(x) + U \bullet G(x) \ge 0, \quad \forall x \in C. \tag{2.2}$$

Theorem 2.5 *Let C be an invex set (convex set, respectively) with respect to η, and let $g(x), G(x)$ be pre-invex functions (convex functions, respectively) with respect to η on C, then the system (2.1) has no solution if and only if there exists $(\lambda, U) \ge 0$ with $(\lambda, U) \ne 0$ such that (2.2) holds.*

3. Duality Theory

In this section, we discuss the duality theory for general nonconvex semidefinite programming (1.1).

We use Ω to denote the set of feasible points of the non-convex optimization problem (1.1) as

$$\Omega := \{x \in C \mid g(x) \le 0, G(x) \le 0\}, \tag{3.1}$$

and we define the Lagrange function as

$$L(x, \lambda, U) = f(x) + \lambda^\top g(x) + U \bullet G(x). \tag{3.2}$$

We call

$$p(\lambda, U) = \min_{x \in C} L(x, \lambda, U)$$

the dual objective function of $L(x, \lambda, U)$, where $(\lambda, U) \ge 0$. Hence the dual problem to (1.1) can be written as

$$\max_{(\lambda, U) \ge 0} \quad p(\lambda, U). \tag{3.3}$$

Now we study the relationship between (1.1) and (3.3). Similar to Theorem 3.1 in [2], we have that the following theorem holds.

Theorem 3.1 *Suppose that Ω is nonempty, then*

$$f(x) \geq p(\lambda, U), \quad \forall x \in \Omega, \ (\lambda, U) \geq 0.$$

Proof. For all $x' \in \Omega$ and $(\lambda, U) \geq 0$, the inequalities $\lambda^\top g(x') \leq 0$ and $U \bullet G(x') \leq 0$ hold from the definition of Ω. With the definition of $p(\lambda, U)$ we have

$$p(\lambda, U) = \min_{x \in C} L(x, \lambda, U) \leq L(x', \lambda, U) = f(x') + \lambda^\top g(x') + U \bullet G(x') \leq f(x').$$

We complete the proof. □

Corollary 3.2 *Suppose that Ω is nonempty. If $\overline{x} \in \Omega, (\overline{\lambda}, \overline{U}) \geq 0$ and $f(\overline{x}) = p(\overline{\lambda}, \overline{U})$ are satisfied, then $\overline{x}$ and $(\overline{\lambda}, \overline{U})$ are optimal solutions of the primal and the dual, respectively.*

Next, we discuss the saddle point problem of (1.1). The saddle point problem is to find a pair $(\overline{x}, \overline{\lambda}, \overline{U})$ with $\overline{x} \in C$ and $(\overline{\lambda}, \overline{U}) \geq 0$ such that

$$L(\overline{x}, \lambda, U) \leq L(\overline{x}, \overline{\lambda}, \overline{U}) \leq L(x, \overline{\lambda}, \overline{U}) \tag{3.4}$$

holds for all $x \in C$ and $(\lambda, U) \geq 0$, we call $(\overline{x}, \overline{\lambda}, \overline{U})$ a saddle point of $L(x, \lambda, U)$. The relation between the solution of (3.4) with the optimal solutions of primal problem (1.1) and dual problem (3.3) is shown in the following result.

Theorem 3.3 *(Duality theorem) Suppose that Ω is nonempty, then $(\overline{x}, \overline{\lambda}, \overline{U})$ with $\overline{x} \in \Omega$ and $(\overline{\lambda}, \overline{U}) \geq 0$ is the solution of (3.4) if and only if*

(a) $\overline{x}$ is the optimal solution of (1.1);
(b) $(\overline{\lambda}, \overline{U})$ is the optimal solution of (3.3);
(c) $f(\overline{x}) = p(\overline{\lambda}, \overline{U})$.

Proof. (Necessary condition) Since $(\overline{x}, \overline{\lambda}, \overline{U})$ is the solution of (3.4), it follows from inequality (3.4) and the definition of $L(x, \lambda, U)$ that

$$\begin{aligned} f(\overline{x}) + \lambda^\top g(\overline{x}) + U \bullet G(\overline{x}) &\leq f(\overline{x}) + \overline{\lambda}^\top g(\overline{x}) + \overline{U} \bullet G(\overline{x}) \\ &\leq f(x) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x) \end{aligned} \tag{3.5}$$

holds for all $x \in C$ and $(\lambda, U) \geq 0$.

Let $\lambda = \overline{\lambda}$ in the inequality above, then

$$U \bullet G(\overline{x}) \leq \overline{U} \bullet G(\overline{x}), \quad \forall\, U \geq 0.$$

This together with $\overline{U} \geq 0$ implies

$$\overline{U} \bullet G(\overline{x}) = 0. \tag{3.6}$$

Similarly, let $U = \overline{U}$, we have

$$\overline{\lambda}^\top g(\overline{x}) = 0. \tag{3.7}$$

Combining (3.6), (3.7) and the right inequality of (3.5), we have

$$f(\overline{x}) \leq f(x) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x) \leq f(x)$$

for all $x \in \Omega$, which means that $\overline{x}$ is the optimal solution of (1.1).

We now show that $(\overline{\lambda}, \overline{U})$ is the optimal solution of (3.3). From the definition of $p(\overline{\lambda}, \overline{U})$ and the right inequality of (3.5), we know that $\overline{x}$ minimizes $L(x, \overline{\lambda}, \overline{U})$ over all $x \in C$, that is

$$p(\overline{\lambda}, \overline{U}) = L(\overline{x}, \overline{\lambda}, \overline{U}) = f(\overline{x}) + \overline{\lambda}^\top g(\overline{x}) + \overline{U} \bullet G(\overline{x}),$$

then it follows from (3.6) and (3.7) that

$$p(\overline{\lambda}, \overline{U}) = f(\overline{x}),$$

which gives result (c). Applying Corollary 3.2, we get result (b).

(Sufficient condition) Since

$$f(\overline{x}) = p(\overline{\lambda}, \overline{U}) = \min_{x \in C} L(x, \overline{\lambda}, \overline{U}) = \min_{x \in C} f(x) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x),$$

we have

$$f(\overline{x}) \leq f(x) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x) \tag{3.8}$$

for all $x \in C$.

Let $x = \overline{x}$, it follows from $(\overline{\lambda}, \overline{U}) \geq 0$, $g(\overline{x}) \leq 0$ and $G(\overline{x}) \leq 0$ that

$$\overline{\lambda}^\top g(\overline{x}) = \overline{U} \bullet G(\overline{x}) = 0. \tag{3.9}$$

Combining (3.8) and (3.9), we have that $(\overline{x}, \overline{\lambda}, \overline{U})$ satisfies the right inequality of (3.5). The left inequality also follows immediately from the fact that $\lambda^\top g(\overline{x}) \leq 0$ and $U \bullet G(\overline{x}) \leq 0$ for all $(\lambda, U) \geq 0$. The proof is completed. □

To study the relationship between the solution of (1.1) and (3.4), we make the following assumption.

Standard Assumption The Slater Condition (SC) holds for problem (1.1), i.e., there exists $\tilde{x} \in C$ such that $g(\tilde{x}) < 0$ and $G(\tilde{x}) < 0$.

Theorem 3.4 *Assume that* $T(x) = (f(x), g(x), G(x)) : C \to R \times R^m \times S^n$ *is a convex-like function and that the standard assumption holds. If* $\overline{x}$ *is the optimal solution of (1.1), then there exists* $(\overline{\lambda}, \overline{U}) \in R^m_+ \times S^n_+$ *such that* $(\overline{x}, \overline{\lambda}, \overline{U})$ *is the solution of (3.4).*

Proof. Since $\overline{x}$ is the optimal solution of (1.1), the following system

$$\begin{cases} f(x) - f(\overline{x}) < 0, \\ g(x) < 0, \\ G(x) < 0, \\ x \in C \end{cases}$$

has no solution, so it follows from Theorem 2.4 that there exists $(\widetilde{\lambda}, \widetilde{U}) \in R_+^{m+1} \times S_+^n$ with $\widetilde{\lambda} = (\widetilde{\lambda}_0, \widetilde{\lambda}_1^\top)^\top$ and $(\widetilde{\lambda}, \widetilde{U}) \neq 0$ such that

$$\widetilde{\lambda}_0(f(x) - f(\overline{x})) + \widetilde{\lambda}_1^\top g(x) + \widetilde{U} \bullet G(x) \geq 0, \quad \forall x \in C. \tag{3.10}$$

First, we show that $\widetilde{\lambda}_0 \neq 0$. Suppose $\widetilde{\lambda}_0 = 0$, then we have $(\widetilde{\lambda}_1, \widetilde{U}) \neq 0$ and

$$\widetilde{\lambda}_1^\top g(x) + \widetilde{U} \bullet G(x) \geq 0, \quad \forall x \in C,$$

which implies $\widetilde{\lambda}_1^\top g(\widetilde{x}) + \widetilde{U} \bullet G(\widetilde{x}) \geq 0$, where $\widetilde{x} \in C$ is the point with $g(\widetilde{x}) < 0$ and $G(\widetilde{x}) < 0$. Hence it follows from $\widetilde{\lambda}_1 \geq 0$, $\widetilde{U} \geq 0$, $g(\widetilde{x}) < 0$ and $G(\widetilde{x}) < 0$ that

$$\widetilde{\lambda}_1^\top g(\widetilde{x}) = \widetilde{U} \bullet G(\widetilde{x}) = 0,$$

so we have $\widetilde{\lambda}_1 = 0$ and $\widetilde{U} = 0$, which lead to a contradiction. Therefore we have $\widetilde{\lambda}_0 \neq 0$.

Let $\overline{\lambda}_0 = \widetilde{\lambda}_0/\widetilde{\lambda}_0 = 1, \overline{\lambda} = \widetilde{\lambda}_1/\widetilde{\lambda}_0$ and $\overline{U} = \widetilde{U}/\widetilde{\lambda}_0$, then $(\overline{\lambda}, \overline{U}) \geq 0$ and (3.10) can be rewritten as

$$f(x) - f(\overline{x}) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x) \geq 0, \quad \forall x \in C. \tag{3.11}$$

Let $x = \overline{x}$, we have $\overline{\lambda}^\top g(\overline{x}) + \overline{U} \bullet G(\overline{x}) \geq 0$, so it follows from $(g(\overline{x}), G(\overline{x})) \leq 0$ and $(\overline{\lambda}, \overline{U}) \geq 0$ that

$$\overline{\lambda}^\top g(\overline{x}) = \overline{U} \bullet G(\overline{x}) = 0. \tag{3.12}$$

Now we show that $(\overline{x}, \overline{\lambda}, \overline{U})$ is the solution of the saddle point problem (3.4). By (3.11) and (3.12), we have

$$f(\overline{x}) + \overline{\lambda}^\top g(\overline{x}) + \overline{U} \bullet G(\overline{x}) \leq f(x) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x), \quad \forall x \in C,$$

which means that the right inequality of (3.5) holds. On the other hand, $(g(\overline{x}), G(\overline{x})) \leq 0$ and $(\lambda, U) \geq 0$ imply that $\lambda^\top g(\overline{x}) \leq 0$ and $U \bullet G(\overline{x}) \leq 0$, hence the left inequality of (3.5) also holds. Therefore, $(\overline{x}, \overline{\lambda}, \overline{U})$ is the solution of (3.4). We complete the proof. □

Corollary 3.5 *Let C be an invex set (convex set, respectively) with respect to η, and assume that f, g and G are all pre-invex functions (convex functions, respectively) with respect to η on C, and that the standard assumption holds. If $\bar{x}$ is the optimal solution of (1.1), then there exists $(\overline{\lambda}, \overline{U}) \in R_+^m \times S_+^n$, such that $(\overline{x}, \overline{\lambda}, \overline{U})$ is the solution of (3.4).*

Proof. The corollary follows immediately from Theorem 3.4 and Theorem 2.5. □

It is easy to see that Theorem 3.5 about (NSDP) in [2] is a special case of Corollary 3.5.

Corollary 3.6 *Let $T(x) = (f(x), g(x), G(x))$ be convex-like function, and let the standard assumption hold. If problem (1.1) has an optimal solution $\overline{x}$, then the dual problem (3.3) has an optimal solution $(\overline{\lambda}, \overline{U})$. Moreover, we have $f(\overline{x}) = p(\overline{\lambda}, \overline{U})$.*

Proof. The corollary follows immediately from Theorem 3.4 and Theorem 3.3. □

Corollary 3.7 *Let C be an invex set (convex set, respectively) with respect to η, and assume that f, g and G are all pre-invex functions (convex functions, respectively) with respect to η on C, and that the standard assumption holds. If (1.1) has an optimal solution $\overline{x}$, then the dual problem (3.3) has an optimal solution $(\overline{\lambda}, \overline{U})$. Moreover, we have $f(\overline{x}) = p(\overline{\lambda}, \overline{U})$.*

Proof. The corollary follows immediately from Theorem 3.3 and Corollary 3.5. □

4. Perfect Duality of Nonconvex Semidefinite Programming

We reformulate (1.2) as

$$\begin{array}{ll} \min_{x\in X} & f(x) \\ \text{s.t.} & h(x) \leq 0, \\ & -h(x) \leq 0, \\ & H(x) \preceq 0 \ \ (or - H(x) \in S^q_+). \end{array} \tag{4.1}$$

Then the Lagrange function of (4.1) is

$$\overline{L}(x, \lambda, U) = f(x) + \lambda^\top h(x) + U \bullet H(x),$$

and the corresponding dual objective function is

$$\overline{p}(\lambda, U) = \min_{x\in X} \overline{L}(x, \lambda, U), \quad (\lambda, U) \in R^m \times S^q_+.$$

Hence the dual problem to (1.2) is written as

$$\max_{(\lambda,U)\in R^m\times S^q_+} \overline{p}(\lambda, U). \tag{4.2}$$

Moveover, the saddle point of problem (1.2) is the point $(\overline{x}, \overline{\lambda}, \overline{U}) \in X \times R^m \times S^q_+$, which satisfies the following inequality

$$\overline{L}(\overline{x}, \lambda, U) \leq \overline{L}(\overline{x}, \overline{\lambda}, \overline{U}) \leq \overline{L}(x, \overline{\lambda}, \overline{U}) \tag{4.3}$$

for all $x \in X$ and $(\lambda, U) \in R^m \times S^q_+$.

From the reformulated problem (4.1) and Theorem 3.3, we have the following corollary.

Corollary 4.1 *Suppose that* $\Omega' = \{x \in X \mid h(x) = 0, H(x) \leq 0\} \neq \phi$ *, then* $(\overline{x}, \overline{\lambda}, \overline{U})$ *with* $\overline{x} \in \Omega'$ *and* $(\overline{\lambda}, \overline{U}) \in R^m \times S^q_+$ *is the solution of (4.3) if and only if*

(a′) $\overline{x}$ *is the optimal solution of (1.2);*

(b′) $(\overline{\lambda}, \overline{U})$ *is the optimal solution of (4.2);*

(c′) $f(\overline{x}) = \overline{p}(\overline{\lambda}, \overline{U})$ *(here the function* f *in the corollary is the objective function of (1.2)).*

As mentioned in Section 1, *Standard Assumption* is not satisfied by programming (1.2) when it is reformulated into (4.1). Fortunately, thanks to Dinh, Jeyakumar and Lee [1], who bring us a new way to analyze the duality property between problem (1.2) and its dual problem (4.2). Enlightened by Theorem 3.1 in [1], we obtain the following lemma.

Lemma 4.2 *Assume that* $T(x) = (f(x), g(x), G(x))$ *is convex-like function and that the standard assumption holds, then there exists* $(\overline{\lambda}, \overline{U}) \geq 0$ *such that*

$$\inf_{x \in C} \sup_{(\lambda,U) \geq 0} L(x, \lambda, U) = \sup_{(\lambda,U) \geq 0} \inf_{x \in C} L(x, \lambda, U) = \inf_{x \in C} L(x, \overline{\lambda}, \overline{U}) = \inf(P),$$

where $\inf(P)$ *is the optimal value of problem (1.1).*

Proof. It is easy to see that

$$\inf_{x \in C} L(x, \overline{\lambda}, \overline{U}) \leq \sup_{(\lambda,U) \geq 0} \inf_{x \in C} L(x, \lambda, U) \leq \inf_{x \in C} \sup_{(\lambda,U) \geq 0} L(x, \lambda, U) \leq \inf(P). \tag{4.4}$$

The last inequality of (4.4) comes from the fact that

$$\begin{aligned} \inf_{x \in C} \sup_{(\lambda,U) \geq 0} L(x, \lambda, U) &\leq \inf_{x \in C, g(x) \leq 0, G(x) \leq 0} \sup_{(\lambda,U) \geq 0} L(x, \lambda, U) \\ &\leq \inf_{x \in C, g(x) \leq 0, G(x) \leq 0} f(x) = \inf(P). \end{aligned}$$

From (4.4), we can complete the proof if $\inf(P) = -\infty$. Hence, we assume that $\inf(P) \neq -\infty$, then the following system

$$\begin{cases} f(x) - \inf(P) < 0, \\ g(x) < 0, \\ G(x) < 0, \\ x \in C \end{cases}$$

has no solution, and according to the proof of Theorem 3.4 we know that there exists $(\overline{\lambda}, \overline{U}) \geq 0$ such that

$$f(x) - \inf(P) + \overline{\lambda}^\top g(x) + \overline{U} \bullet G(x) \geq 0, \quad \forall x \in C,$$

then

$$\inf_{x \in C} L(x, \overline{\lambda}, \overline{U}) \geq \inf(P). \tag{4.5}$$

Combining (4.4) and (4.5), we complete the proof. □

The perfect duality in nonlinear programming has been in the literatures since at least the 1970s, and also discussed in paper [1]. We recall that a minimization (maximization, respectively) problem is said to be consistent if its value is not equal to $+\infty$ ($-\infty$, respectively). We say that the two problems are in perfect duality if

a) when one problem has finite value, the other has the same value;

b) when both problems are consistent, they share the same values.

Let us consider a series of problems: for all $t > 0$,

$$\begin{array}{cl} \min_{x \in X} & f(x) \\ \text{s.t.} & |h(x)| \leq t, \\ & H(x) \leq tI, \end{array} \tag{4.6}$$

where

$$|h(x)| = \begin{pmatrix} |h_1(x)| \\ \vdots \\ |h_m(x)| \end{pmatrix}.$$

It is easy to see that problem (4.6) is exactly the problem (1.2) if we let t be zero, and that the optimal value of (4.6) depends on the positive real number t. When we denote the optimal value of (4.6) by $\inf(P_t)$ for each $t > 0$, it is obvious that the function $\inf(P_t)$ about t is a non-increasing function on R_+.

Denote by $\langle P \rangle$ ($\langle D \rangle$, respectively) the value of $\lim_{t \downarrow 0} \inf(P_t)$ (the optimal value of the dual problem (4.2), respectively). In order to deduce the perfect duality, we need another simple lemma.

Lemma 4.3

$$\begin{aligned} \sup_{x \in R^m} T(x) &= \sup_{(x_1, x_2) \in R^m_+ \times R^m_+} T(x_1 - x_2), \\ \inf_{x \in R^m} T(x) &= \inf_{(x_1, x_2) \in R^m_+ \times R^m_+} T(x_1 - x_2). \end{aligned}$$

Proof. The inequality $\sup_{x\in R^m} T(x) \geq \sup_{(x_1,x_2)\in R_+^m\times R_+^m} T(x_1-x_2)$ holds obviously. On the other hand, for all $x \in R^m$, there exist

$$x_1 = \max\{0, x\} \geq 0, \quad x_2 = x_1 - x \geq 0,$$

such that $x = x_1 - x_2$, hence $\sup_{x\in R^m} T(x) \leq \sup_{(x_1,x_2)\in R_+^m\times R_+^m} T(x_1-x_2)$.

From the relation $\inf_{x\in Z} T(x) = -\sup_{x\in Z} T(x)$ and the first equality, we complete the proof. □

Theorem 4.4 *Suppose, for all $x, y \in X$ and $\mu \in [0,1]$, that there exists $z \in X$ such that*

$$h(z) = \mu h(x) + (1-\mu)h(y)$$

and

$$(f, H)(z) \leq \mu(f, H)(x) + (1-\mu)(f, H)(y)$$

hold at the same time, then the problem (4.2) and $\lim_{t\downarrow 0}\inf(P_t)$ are in perfect duality.

Proof. First, we show that

$$\langle P\rangle \geq \langle D\rangle. \tag{4.7}$$

In order to show that (4.7) holds, we only need to show for all $U \geq 0$ and $\lambda \in R^m$ that the following inequality

$$\langle P\rangle \geq \inf_{x\in X} f(x) + \lambda^\top h(x) + U \bullet H(x) \tag{4.8}$$

holds.

Suppose that (4.8) does not hold, then there exist $\overline{U} \geq 0$ and $\overline{\lambda} \in R^m$, such that

$$\langle P\rangle < \inf_{x\in X} f(x) + \overline{\lambda}^\top h(x) + \overline{U} \bullet H(x). \tag{4.9}$$

From the definition of $\inf(P_t)$, we have for all $t > 0$ that

$$\begin{aligned}
\inf(P_t) &\geq \inf_{x\in X} \sup_{(\lambda_1,\lambda_2,U)\geq 0} f(x) + \lambda_1^\top(h(x) - te) + \lambda_2^\top(-h(x) - te) \\
&\quad + U \bullet (H(x) - tI) \\
&= \inf_{x\in X} \sup_{(\lambda_1,\lambda_2,U)\geq 0} f(x) + (\lambda_1 - \lambda_2)^\top h(x) + U \bullet H(x) \\
&\quad - t(\sum_{i=1}^{2}\sum_{j=1}^{m} \lambda_i^{(j)} + Tr(U)) \\
&\geq \sup_{(\lambda_1,\lambda_2,U)\geq 0} \inf_{x\in X} f(x) + (\lambda_1 - \lambda_2)^\top h(x) + U \bullet H(x) \\
&\quad - t(\sum_{i=1}^{2}\sum_{j=1}^{m} \lambda_i^{(j)} + Tr(U)),
\end{aligned}$$

where $e = (1, \cdots, 1)^\top \in R^m$, $\lambda_i^{(j)}$ denotes the j-th component of $\lambda_i, (i = 1, 2)$. Then for all $t > 0$ and $(\lambda_1, \lambda_2, U) \geq 0$, we have

$$\begin{aligned} \inf(P_t) &\geq \{\inf_{x \in X} f(x) + (\lambda_1 - \lambda_2)^\top h(x) + U \bullet H(x)\} \\ &- t(\sum_{i=1}^{2} \sum_{j=1}^{m} \lambda_i^{(j)} + Tr(U)). \end{aligned} \tag{4.10}$$

Define $\overline{\lambda}_1 = \max\{0, \overline{\lambda}\} \geq 0$ and $\overline{\lambda}_2 = \max\{0, \overline{\lambda}\} - \overline{\lambda} \geq 0$, let $(\lambda_1, \lambda_2, U) = (\overline{\lambda}_1, \overline{\lambda}_2, \overline{U})$ and $t \to 0$ in (4.10), then we have

$$\langle P \rangle = \lim_{t \downarrow 0} \inf(P_t) \geq \inf_{x \in X} f(x) + \overline{\lambda}^\top h(x) + \overline{U} \bullet H(x),$$

which is a contradiction to (4.9), so (4.7) holds.

By assumption,

$$W(x) = \begin{pmatrix} \Lambda_{h(x)} & 0 & 0 \\ 0 & -\Lambda_{h(x)} & 0 \\ 0 & 0 & H(x) \end{pmatrix}$$

is a convex-like function on X, and the feasible set of (4.6) is

$$\{x \in X \mid |h(x)| \leq t, H(x) \leq tI_q\} = \{x \in X \mid W(x) \leq tI_{2m+q}\}.$$

For simplicity, we omit the subscript $2m + q$ of the identity matrix in the latter part of this paper. Suppose that $\langle D \rangle$ is finite, if there exists $t_0 > 0$ such that

$$\{x \in X \mid W(x) < t_0 I\} = \phi,$$

then we can find a nonzero $(\overline{\lambda}_1, \overline{\lambda}_2, \overline{U}) \geq 0$ such that

$$\inf_{x \in X} \{\overline{\lambda}_1^\top (h(x) - t_0 e) + \overline{\lambda}_2^\top (-h(x) - t_0 e) + \overline{U} \bullet (H(x) - t_0 I)\} \geq 0$$

by using Theorem 2.4. Notice that $(\overline{\lambda}_1, \overline{\lambda}_2, \overline{U}) \geq 0, t_0 > 0$ and $(\overline{\lambda}_1, \overline{\lambda}_2, \overline{U}) \neq 0$, then we have

$$\inf_{x \in X} \{\overline{\lambda}_1^\top h(x) + \overline{\lambda}_2^\top (-h(x)) + \overline{U} \bullet H(x)\} \geq t_0 (\sum_{i=1}^{2} \sum_{j=1}^{m} \overline{\lambda}_i^{(j)} + Tr(\overline{U})) > 0. \tag{4.11}$$

From Lemma 4.3 and (4.11) we have

$$\begin{aligned} \langle D \rangle &= \sup_{U \geq 0, \lambda \in R^m} \inf_{x \in X} \overline{L}(x, \lambda, U) \\ &= \sup_{(\lambda_1, \lambda_2, U) \geq 0} \inf_{x \in X} f(x) + \lambda_1^\top h(x) + \lambda_2^\top (-h(x)) + U \bullet H(x) \\ &\geq \sup_{(\lambda_1, \lambda_2, U) \geq 0} \inf_{x \in X} f(x) + (\lambda_1 + \overline{\lambda}_1)^\top h(x) + (\lambda_2 + \overline{\lambda}_2)^\top (-h(x)) \end{aligned}$$

$$
\begin{aligned}
&\quad +(U+\overline{U})\bullet H(x) \\
&\geq \sup_{U\geq 0,\lambda\in R^m}\inf_{x\in X}\{f(x)+\lambda^\top h(x)+U\bullet H(x)\}+\inf_{x\in X}\{\overline{\lambda}_1^\top h(x) \\
&\quad +\overline{\lambda}_2^\top(-h(x))+\overline{U}\bullet H(x)\} \\
&> \langle D\rangle,
\end{aligned}
$$

which is a contradiction. Therefore, if $\langle D\rangle$ is finite, then $\{x\in X \mid W(x)<tI\}$ is nonempty for all $t>0$.

On the other hand, if $\langle P\rangle$ is finite, then from the definition of $\langle P\rangle$ and the non-increasing of $\inf(P_t)$, we have $\inf(P_t)<+\infty$ for all $t>0$, which means that $\inf(P_t)$ is consistent for all $t>0$, i.e.,

$$\{x\in X \mid W(x)\leq tI\}\neq\Phi,\quad \forall t>0,$$

which is equivalent to

$$\{x\in X \mid W(x)<tI\}\neq\Phi,\quad \forall t>0.$$

Since the constraint mapping of (4.6) is $W(\cdot)-tI$, we can see that, if one of $\langle P\rangle$ and $\langle D\rangle$ has a finite value, (4.6) satisfies the Slater condition for all $t>0$. So, from Lemma 4.2, we have that

$$\inf(P_t)=\sup_{(\lambda_1,\lambda_2,U)\geq 0}\inf_{x\in X} f(x)+(\lambda_1-\lambda_2)^\top h(x)+U\bullet H(x)-t(\sum_{i=1}^{2}\sum_{j=1}^{m}\lambda_i^{(j)}+Tr(U))$$

holds for all $t>0$ under the finiteness condition of $\langle P\rangle$ or $\langle D\rangle$.

By use of $(\lambda_1,\lambda_2,U)\geq 0, t>0$ and Lemma 4.3, we have that

$$
\begin{aligned}
\inf(P_t) &\leq \sup_{(\lambda_1,\lambda_2,U)\geq 0}\inf_{x\in X} f(x)+(\lambda_1-\lambda_2)^\top h(x)+U\bullet H(x) \\
&= \sup_{\lambda\in R^m,U\geq 0}\inf_{x\in X}\overline{L}(x,\lambda,U)=\langle D\rangle
\end{aligned}
$$

holds for all $t>0$, it follows that

$$\langle P\rangle=\lim_{t\downarrow 0}\inf(P_t)\leq\langle D\rangle. \tag{4.12}$$

Combining (4.7) and (4.12) , we have $\langle P\rangle=\langle D\rangle$, which yields conclusion a).

If both problems are consistent, we have

$$+\infty>\langle P\rangle\geq\langle D\rangle>-\infty.$$

So, a) implies b) in this case. Then we complete the proof. □

Corollary 4.5 *Suppose that, for all $x,y\in X$ and $\mu\in[0,1]$, there exists $\eta: X\times X\to X$ such that $h(y+\mu\eta(x,y))=\mu h(x)+(1-\mu)h(y)$ (linear function, respectively) and (f,H) is a pre-invex function (convex function, respectively) with respect to η, then the problem (4.2) and $\lim_{t\downarrow 0}\inf(P_t)$ are in perfect duality.*

5. Conclusion

In this paper, we discuss the duality theory for non-convex semidefinite programming (1.1). Our contributions are as follows: 1. we establish the duality theory for the general nonconvex semidefinite programming (1.1) when Slater's condition holds; 2. we obtain the perfect duality for a special case (1.2) of the nonconvex semidefinite programming for which Slater's condition fails. Furthermore, we point out that the results of [2] can be regarded as a special case of our result.

Acknowledgements. We thank the editors and two anonymous referees for their helpful comments and suggestions which greatly improve our manuscript.

References

[1] N. Dinh, V. Jeyakumar and G.M. Lee, Sequential Lagrangian conditions for convex programs with applications to semidefinite programming, *Journal of Optimization Theory and Applications*, 125 (2005) 85-112.

[2] J.Y. Fan, Duality theories in nonlinear semidefinite programming, *Applied Mathematics Letters*, 18 (2005) 1068-1073.

[3] K. Fan, Minimax theorems, *Proceedings of the National Academy of Sciences of the USA*, vol. 39, No.1, (1953).

[4] C. Kanzow, C. Nagel, H. Kato, M. Fukushima, Successive linearization methods for nonlinear semidefinite programs. *Computational Optimization and Applications*, 31 (2005) 251–273.

[5] C. Li and W. Sun, Generalized Farkas' lemma and optimal conditions for nonconvex semidefinite program, *Numerical Mathematics, A Journal of Chinese Universities*, 30 (2008) 184-192.

[6] C. Li and W. Sun, A class of nonsmooth Newton-type methods for nonlinear semidefinite programming, *Journal of Nanjing Normal University, Natural Science*, 31 (2008) 1-7.

[7] C. Li and W. Sun, On filter-successive linearization methods for nonlinear semidefinite programming, *Science in China Series A*, 52 (2009) 2341-2361.

[8] C. Li, W. Sun and R.J.B. Sampaio, An equivalency condition on nonsigularity in nonlinear semidefinite programming, *Journal of System Sciences and Complexity*, 23 (2010) 822-829.

[9] J. Sun, D. Sun, L. Qi, A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems, *SIAM J. Optimization*, 14 (2004) 783-806.

[10] D. Sun, The strong second order sufficient condition and constraint non-degeneracy in nonlinear semidefinite programming and their implications, *Mathematics of Operations Research*, 31 (2006) 761-776.

[11] D. Sun, J. Sun and L. Zhang, The rate of convergence of the augmented lagrangian method for nonlinear semidefinite programming, *Mathematical Programming*, 114 (2008) 349-391.

[12] W. Sun and Y. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer Optimization and Its Applications (SOIA), Volume 1. Springer, New York, (2006).

[13] T. Weir and V. Jeyakumar, A class of nonconvex functions and Mathematical programming, *Bulletin of the Australian Mathematical Society*, 38 (1988) 177-189.

[14] X.M. Yang, X.Q. Yang and K.L. Teo, Characterizations and applications of prequasi-invex functions, *Journal of Optimization Theory and Applications*, 110 (2001) 645-668.