数理解析研究所講究録
945巻 1996年 57-67

Semidefinite Programming Relaxation for Nonconvex Quadratic Programs

東京工業大学 藤江 哲也 (Tetsuya Fujie)
東京工業大学 小島 政和 (Masakazu Kojima)

Abstract. Any quadratic inequality in the n-dimensional Euclidean space can be relaxed into a linear matrix inequality in $(1+n)\times(1+n)$ symmetric matrices. Based on this principle, we extend the Lovász-Schrijver SDP (semidefinite programming) relaxation developed for a 0-1 integer program to a general nonconvex QP (quadratic program), and present some fundamental characterization of the SDP relaxation including its equivalence to a relaxation using convex-quadratic valid inequalities for the feasible region of the QP.

Key words. Semidefinite Program, Relaxation Method, Interior-Point Method, Linear Matrix Inequality, Nonconvex Quadratic Program

1 Introduction.

We use the symbols $\mathcal{S}(m)$ for the set of $m\times m$ symmetric matrices, and $\mathcal{S}(m)_+$ (or $\mathcal{S}(m)_{++}$) for the cone consisting of $m\times m$ symmetric positive semidefinite (or positive definite, respectively) matrices. We are concerned with a canonical form QP:

$$\text{Minimize } \boldsymbol{c}^T\boldsymbol{y} \text{ subject to } \boldsymbol{y}\in\mathcal{F}. \tag{1}$$

Here

$$\left.\begin{array}{lcl}
\mathcal{F} & \equiv & \left\{\boldsymbol{y}\in R^{1+n} \ : \ y_0=1,\ \boldsymbol{y}^T\boldsymbol{P}_k\boldsymbol{y}\le 0\ (k=1,2,\ldots,m)\right\}, \\
\boldsymbol{c} & \equiv & \begin{pmatrix}\gamma\\ \boldsymbol{d}\end{pmatrix}\in R^{1+n},\ \boldsymbol{y}\equiv\begin{pmatrix}y_0\\ y_1\\ \cdot\\ \cdot\\ y_n\end{pmatrix}\in R^{1+n}, \\
\boldsymbol{P}_k & \equiv & \begin{pmatrix}\pi_k & \boldsymbol{q}_k^T/2\\ \boldsymbol{q}_k/2 & \boldsymbol{Q}_k\end{pmatrix}\in\mathcal{S}(1+n)\ (k=1,2,\ldots,m), \\
\pi_k & \in & R,\ \boldsymbol{q}_k\in R^n,\ \boldsymbol{Q}_k\in\mathcal{S}(n)\ (k=1,2,\ldots,m).
\end{array}\right\} \tag{2}$$

Note that the feasible region $\mathcal{F}$ is contained in the n-dimensional hyperplane $H\equiv\{\boldsymbol{y}\in R^{1+n} : y_0=1\}$, and that the function $H\ni\boldsymbol{y}\to\boldsymbol{y}^T\boldsymbol{P}_k\boldsymbol{y}\in R$ involved in the inequality constraint is convex (or linear) if and only if $\boldsymbol{Q}_k\in\mathcal{S}(n)_+$ (or $\boldsymbol{Q}_k=\boldsymbol{O}$, respectively). $\boldsymbol{Q}_k$ can be indefinite, so that the feasible region $\mathcal{F}$ of the QP (1) is a nonconvex subset of the hyperplane H in general. The canonical form QP (1) covers various mathematical programs such as 0-1 IPs, general nonconvex QPs and linear complementarity problems.

SDP(Semidefinite Program) has been considered as a powerful tool for relaxation of many combinatorial optimization problems, since it yields a tight bound [1, 2, 5, 6, 8, 13, 17] and it is efficiently solvable by interior methods [1, 2, 3, 4, 7, 9, 12, 16]. SDP relaxation is originally proposed by Lovász[10] for stable set problems.

Among many literatures related on SDP relaxation, this paper was motivated by Alizadeh [1] where an elementary outline of the SDP relaxation method proposed by Lovász-Schrijver [11] for 0-1 IPs (integer programs) was presented. The aim of this paper is to present a general method for constructing an SDP which serves as a relaxation of the QP (1) and some fundamental properties on the SDP relaxation. Our SDP relaxation method may be regarded as a straightforward extension of the Lovász-Schrijver SDP relaxation method [11] for 0-1 IPs to the QP (1). It is also characterized in terms of

- a dual of Shor's relaxation method [14] (see also [15]) for general nonconvex QPs (this will be discussed in Section 5), and
- a relaxation using *convex-quadratic valid inequalities* for the feasible region $\mathcal{F}$.

Let

$$\boldsymbol{P} \equiv \begin{pmatrix} \pi & \boldsymbol{q}^T/2 \\ \boldsymbol{q}/2 & \boldsymbol{Q} \end{pmatrix} \in \mathcal{S}(1+n),\ \pi \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} \in \mathcal{S}(n).$$

We say that an inequality $\boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} \leq 0$ is a convex-quadratic (or linear) valid inequality for $\mathcal{F}$ if

$$\boldsymbol{Q} \in \mathcal{S}(n)_+ \text{ (or } \boldsymbol{Q} = \boldsymbol{O}\text{, respectively) and } \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} \leq 0 \text{ for every } \boldsymbol{y} \in \mathcal{F}.$$

Then co $\mathcal{F}$, the convex hull of $\mathcal{F}$ is completely determined by all the convex-quadratic valid inequalities for $\mathcal{F}$;

$$\text{co}\,\mathcal{F} = \bigcap_{\boldsymbol{P}\in\mathcal{V}} \{\boldsymbol{y} \in R^{1+n} \ :\ y_0 = 1,\ \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} \leq 0\},$$

where $\mathcal{V}$ denotes the set of all matrices $\boldsymbol{P} \in \mathcal{S}(1+n)$ that induce convex-quadratic valid inequalities for $\mathcal{F}$. (The identity above is well-known when $\mathcal{V}$ is the set of all matrices $\boldsymbol{P} \in \mathcal{S}(1+n)$ that induce linear valid inequalities for $\mathcal{F}$). The discussion above leads us to a relaxation of the QP (1) using all convex-quadratic valid inequalities for $\mathcal{F}$ that we can generate as a nonnegative combination of the quadratic inequalities of the QP (1):

$$\text{Minimize } \boldsymbol{c}^T\boldsymbol{y} \text{ subject to } \boldsymbol{y} \in \widetilde{\mathcal{F}}, \tag{3}$$

where

$$\left.\begin{array}{rcl}\widetilde{T} & \equiv & \{\boldsymbol{t} \in R^m \ : \ \boldsymbol{t} \geq \boldsymbol{0},\ \sum_{k=1}^{m} t_k \boldsymbol{Q}_k \in \mathcal{S}(n)_+\}, \\ f_{\boldsymbol{t}}(\boldsymbol{y}) & \equiv & \boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k\right) \boldsymbol{y} \ \text{ for every } \boldsymbol{y} \in R^{1+n} \ (\boldsymbol{t} \in \widetilde{T}), \\ \widetilde{\mathcal{F}} & \equiv & \left\{\boldsymbol{y} \in R^{1+n} \ : \ y_0 = 1 \ \text{ and } f_{\boldsymbol{t}}(\boldsymbol{y}) \leq 0 \ (\boldsymbol{t} \in \widetilde{T})\right\}, \\ & = & \left\{\boldsymbol{y} \in R^{1+n} \ : \begin{array}{c} y_0 = 1 \ \text{ and } \boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k\right) \boldsymbol{y} \leq 0 \\ \text{for every } \boldsymbol{t} \geq \boldsymbol{0} \ \text{ such that } \sum_{k=1}^{m} t_k \boldsymbol{Q}_k \in \mathcal{S}(n)_+ \end{array}\right\}. \end{array}\right\} \quad (4)$$

Although the derivation of the relaxation (3) of the QP (1) is simple and straightforward, it seems difficult to implement the relaxation (3) on computer because the set $\widetilde{T}$ over which the index vector $\boldsymbol{t}$ of the convex-quadratic inequality $f_{\boldsymbol{t}}(\boldsymbol{y}) \leq 0$ changes is a continuum, non-polyhedral and convex subset of R^m in general. Under a moderate assumption (Condition 2.2), the main theorem (Theorem 2.3) establishes the equivalence between the SDP relaxation and the relaxation (3) using convex-quadratic valid inequalities. Thus the SDP relaxation may be regarded as an implementable version of the relaxation (3).

We give the main theorem without proof in Section 2. Section 3 states a basic principle which makes it possible for us to extend the Lovász-Schrijver SDP relaxation method for IPs to nonconvex QPs. In Section 4, we present Shor's relaxation method [14], and show some duality relation among the SDP relaxation, Shor's relaxation and the relaxation (3) using convex-quadratic valid inequalities. The discussions in Sections 3 and 4 are not only necessary to prove the main theorem, but also helpful to the readers' deep understanding of the SDP relaxation. Section 5 is devoted to a proof of the main theorem.

2 Main Theorem.

For every $\boldsymbol{A} \in \mathcal{S}(m)$ and $\boldsymbol{B} \in \mathcal{S}(m)$, $\boldsymbol{A} \bullet \boldsymbol{B}$ denotes their inner product, *i.e.*, $\boldsymbol{A} \bullet \boldsymbol{B} \equiv \text{Tr}\, \boldsymbol{A}^T \boldsymbol{B}$ (the trace of $\boldsymbol{A}^T\boldsymbol{B}$). It should be noted that any linear function $g : \mathcal{S}(m) \rightarrow R$ can be written as $g(\boldsymbol{Y}) = \boldsymbol{A} \bullet \boldsymbol{Y}$ for some $\boldsymbol{A} \in \mathcal{S}(m)$. Define

$$\left.\begin{array}{rcl}\boldsymbol{C} & \equiv & \begin{pmatrix} \gamma & \boldsymbol{d}^T/2 \\ \boldsymbol{d}/2 & \boldsymbol{O} \end{pmatrix} \in \mathcal{S}(1+n), \\ \widehat{\mathcal{G}} & \equiv & \{\boldsymbol{Y} \in \mathcal{S}(1+n)_+ \ : \ Y_{00} = 1,\ \boldsymbol{P}_k \bullet \boldsymbol{Y} \leq 0 \ (k = 1, 2, \ldots, m)\}, \\ \widehat{\mathcal{F}} & \equiv & \{\boldsymbol{Y}\boldsymbol{e}_0 \ : \ \boldsymbol{Y} \in \widehat{\mathcal{G}}\},\ \boldsymbol{e}_0 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ 0 \end{pmatrix} \in R^{1+n}. \end{array}\right\} \quad (5)$$

Obviously, $\widehat{\mathcal{G}}$ and $\widehat{\mathcal{F}}$ are convex subsets of $\mathcal{S}(1+n)$ and R^{1+n}, respectively. We now introduce the SDP which will serve as a relaxation of the QP (1):

$$\text{Minimize } \boldsymbol{C} \bullet \boldsymbol{Y} \text{ subject to } \boldsymbol{Y} \in \widehat{\mathcal{G}}; \tag{6}$$

We can rewrite the SDP as a convex minimization problem in the Euclidean space:

$$\text{Minimize } \boldsymbol{c}^T\boldsymbol{y} \text{ subject to } \boldsymbol{y} \in \widehat{\mathcal{F}}. \tag{7}$$

The two problems (6) and (7) above are equivalent in the sense that:

Lemma 2.1

1. *$\boldsymbol{y}$ is a feasible solution of the problem (7) if and only if $\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{e}_0$ for some feasible solution $\boldsymbol{Y}$ of the problem (6).*

2. *$\boldsymbol{y}$ is a minimum solution of the problem (7) if and only if $\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{e}_0$ for some minimum solution $\boldsymbol{Y}$ of the problem (6).*

3. $\inf\{\boldsymbol{C} \bullet \boldsymbol{Y} \ : \ \boldsymbol{Y} \in \widehat{\mathcal{G}}\} = \inf\{\boldsymbol{c}^T\boldsymbol{y} \ : \ \boldsymbol{y} \in \widehat{\mathcal{F}}\}$. ∎

We will be mainly concerned with the convex minimization problem (7) instead of the SDP (6). If we restrict ourselves to QPs derived from 0-1 IPs, our construction of the problem (7) is a special case of the Lovász-Schrijver [11] relaxation method. We impose the following condition on the feasible region $\widehat{\mathcal{G}}$ of the SDP (6) in the main theorem below.

Condition 2.2 There is an interior point $\boldsymbol{Y}$ of the feasible region $\widehat{\mathcal{G}}$ of the SDP (6), a $\boldsymbol{Y} \in \mathcal{S}(1+n)_{++}$ satisfying $Y_{00} = 1$ and $\boldsymbol{P}_k \bullet \boldsymbol{Y} < 0$ $(k = 1, 2, \ldots, m)$. ∎

Now we consider the convex minimization problem (3) introduced in the Introduction as a relaxation of the QP (1) using convex-quadratic valid inequalities for $\mathcal{F}$. If all the extreme points and all the extreme rays of $\widetilde{\mathcal{F}}$ are contained in $\mathcal{F}$ then $\widetilde{\mathcal{F}}$ coincides with co $\mathcal{F}$, the convex hull of $\mathcal{F}$ and the problem (3) gives the best convex relaxation of the QP (1). But $\widetilde{\mathcal{F}} \neq$ co $\mathcal{F}$ in general. We focus our attention to a subset of extreme points of $\widetilde{\mathcal{F}}$ which are shown to be contained in $\mathcal{F}$ ((iii) of Theorem 2.3). We say that a point $\boldsymbol{y} \in \widetilde{\mathcal{F}}$ is *a strictly convex boundary point of* $\widetilde{\mathcal{F}}$ if there exists a $\boldsymbol{t} = (t_1, t_2, \ldots, t_m)^T \geq \boldsymbol{0}$ such that

$$\boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k \right) \boldsymbol{y} = 0 \text{ and } \sum_{k=1}^{m} t_k \boldsymbol{Q}_k \in \mathcal{S}(n)_{++}. \tag{8}$$

It should be noted that the definition of a strictly convex boundary point depends on the algebraic representation of $\mathcal{F}$. That is, a strictly convex boundary point $\boldsymbol{y}$ of $\widetilde{\mathcal{F}}$ of the representation (4) is not necessarily a strictly convex boundary point of $\widetilde{\mathcal{F}}$ of a distinct representation. See section 6.

Now we are ready to state:

Theorem 2.3 (main theorem)

1. $\mathcal{F} \subseteq \widehat{\mathcal{F}} \subseteq \widetilde{\mathcal{F}}$.

2. Suppose that the feasible region $\widehat{\mathcal{G}}$ *satisfies Condition 2.2. Then*

$$\inf\{\boldsymbol{c}^T\boldsymbol{y} \ : \ \boldsymbol{y} \in \widehat{\mathcal{F}}\} = \inf\{\boldsymbol{c}^T\boldsymbol{y} \ : \ \boldsymbol{y} \in \widetilde{\mathcal{F}}\} \tag{9}$$

for every $\boldsymbol{c} \in R^{1+n}$, *and* $\widetilde{\mathcal{F}} = cl\, \widehat{\mathcal{F}}$, *the closure of* $\widehat{\mathcal{F}}$.

3. Every strictly convex boundary point $\boldsymbol{y}$ *of* $\widetilde{\mathcal{F}}$ *belongs to* $\mathcal{F}$. ∎

Proof of the theorem is given in Section 5.

3 A Single Quadratic Inequality.

The most important principle behind the SDP relaxation is: Any quadratic inequality in the n-dimensional Euclidean space can be relaxed into a linear matrix inequality in $(1 + n) \times (1 + n)$ symmetric matrices. We will associate each vector $\boldsymbol{y} = (1, y_1, y_2, \ldots, y_n)^T$ in R^{1+n} with a $(1 + n) \times (1 + n)$ symmetric matrix

$$\boldsymbol{Y} = \boldsymbol{y}\boldsymbol{y}^T = \begin{pmatrix} 1 & y_1 & y_2 & \cdots & y_n \\ y_1 & y_1y_1 & y_1y_2 & \cdots & y_1y_n \\ y_2 & y_2y_1 & y_2y_2 & \cdots & y_2y_n \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ y_n & y_ny_1 & y_ny_2 & \cdots & y_ny_n \end{pmatrix} \in \mathcal{S}(1+n). \tag{10}$$

The matrix $\boldsymbol{Y} \in \mathcal{S}(1 + n)$ contains all the constant, linear and quadratic "atomic" terms, *i.e.*, the nonzero constant term 1, the n linear terms $y_1, y_2, \ldots, y_n$ and the n^2 quadratic terms $y_1y_1, y_1y_2, \ldots, y_ny_n$ in its elements, so that we can represent any function consisting of linear and quadratic forms of $y_1, y_2, \ldots, y_n$ in terms of a linear combination of those terms, *i.e.*, a linear function $\boldsymbol{P} \bullet \boldsymbol{Y}$ of $\boldsymbol{Y}$ for some $\boldsymbol{P} \in \mathcal{S}(1 + n)$.

By the construction, for

$$\boldsymbol{P} \equiv \begin{pmatrix} \pi & \boldsymbol{q}^T/2 \\ \boldsymbol{q}/2 & \boldsymbol{Q} \end{pmatrix} \in \mathcal{S}(1+n), \ \boldsymbol{q} \in R^n, \boldsymbol{Q} \in \mathcal{S}(n), \tag{11}$$

we see that

$$\boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} = \boldsymbol{P} \bullet \boldsymbol{Y} \tag{12}$$

whenever

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{e}_0, \ y_0 = 1 \ \text{ and } \boldsymbol{Y} = \boldsymbol{y}\boldsymbol{y}^T. \tag{13}$$

On the other hand, we know that an $(1+n) \times (1+n)$ matrix $\boldsymbol{Y}$ satisfies (13) for some $\boldsymbol{y} = (y_0, y_1, \ldots, y_n)^T \in R^{1+n}$ if and only if

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{e}_0, \ Y_{00} = 1, \ \boldsymbol{Y} \in \mathcal{S}(1+n)_+ \ \text{ and rank } \boldsymbol{Y} = 1.$$

Hence

$$\boldsymbol{y} \in R^{1+n},\ \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} \leq 0 \text{ and } y_0 = 1$$

if and only if

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{e}_0 \in R^{1+n},\ \boldsymbol{P} \bullet \boldsymbol{Y} \leq 0,\ Y_{00} = 1,\ \boldsymbol{Y} \in \mathcal{S}(1+n)_+ \text{ and rank } \boldsymbol{Y} = 1.$$

Dropping the last rank condition rank $\boldsymbol{Y} = 1$, we obtain:

Lemma 3.1 *Let $\boldsymbol{P} \in \mathcal{S}(1+n)$. If*

$$\boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} \leq 0 \text{ and } y_0 = 1 \tag{14}$$

then

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{e}_0 \in R^{1+n},\ \boldsymbol{P} \bullet \boldsymbol{Y} \leq 0,\ Y_{00} = 1 \text{ and } \boldsymbol{Y} \in \mathcal{S}(1+n)_+ \tag{15}$$

for some $\boldsymbol{Y} \in \mathcal{S}(1+n)$. ∎

(Relaxation by dropping the rank condition as mentioned above has been utilized in many papers [1, 2, 5, 7, 8, 10, 13, 17], etc.).

Lemma 3.2 *Let $\boldsymbol{P}$ be a $(1+n) \times (1+n)$ symmetric matrix of the form (11). Suppose that a $(1+n) \times (1+n)$ matrix $\boldsymbol{Y}$ and $\boldsymbol{y} \in R^{1+n}$ satisfy the relation (15). Let*

$$\boldsymbol{y} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \text{ and } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix}.$$

Then

$$\boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} = \boldsymbol{P} \bullet \boldsymbol{Y} - \boldsymbol{Q} \bullet (\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T) \leq -\boldsymbol{Q} \bullet (\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T).$$

If in addition the $n \times n$ matrix $\boldsymbol{Q}$ is positive semi-definite then $\boldsymbol{y}$ satisfies the relation (14).

Proof: By the definitions of the matrices $\boldsymbol{P}$ and $\boldsymbol{Y}$, we have that

$$\begin{aligned} \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} &= \pi + \boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \\ &= \pi + \boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} - \boldsymbol{Q} \bullet (\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T) \\ &= \boldsymbol{P} \bullet \boldsymbol{Y} - \boldsymbol{Q} \bullet (\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T) \\ &\leq -\boldsymbol{Q} \bullet (\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T). \end{aligned}$$

Thus we have shown the first assertion. It follows from $\boldsymbol{Y} \in \mathcal{S}(1+n)_+$ and $Y_{00} = 1$ that $\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T \in \mathcal{S}(n)_+$. Hence if $\boldsymbol{Q} \in \mathcal{S}(n)_+$ then $\boldsymbol{Q} \bullet (\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T) \geq 0$; hence $\boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} \leq 0$. ∎

4 Duality.

Applying Shor's relaxation method [14] to the QP (1), we obtain an SDP

$$\text{Maximize } t_0 \text{ subject to } \boldsymbol{t} \in T^d, \tag{16}$$

where

$$T^d \equiv \left\{ \boldsymbol{t} = (t_0, t_1, \ldots, t_m)^T \ : \ \begin{array}{l} \boldsymbol{C} - t_0 \boldsymbol{e}_0 \boldsymbol{e}_0^T + \sum_{i=1}^m t_i \boldsymbol{P}_i \in \mathcal{S}(1+n)_+, \\ t_i \geq 0 \ (i = 1, 2, \ldots, m) \end{array} \right\}.$$

Between the two problems (16) and (1), the following relation holds.

Lemma 4.1 ([14], see also [17]) *If $\boldsymbol{t} = (t_0, t_1, \ldots, t_m)^T \in R^{1+m}$ is a feasible solution of the SDP (16) and $\boldsymbol{y} \in R^{1+n}$ a feasible solution of the QP (1), then their objective values t_0 and $\boldsymbol{c}^T \boldsymbol{y}$ satisfies the inequality $t_0 \leq \boldsymbol{c}^T \boldsymbol{y}$; $\sup\{t_0 \ : \ \boldsymbol{t} \in T^d\} \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} \ : \ \boldsymbol{y} \in \mathcal{F}\}$.*

Proof: Assume that $\boldsymbol{t} = (t_0, t_1, \ldots, t_m)^T \in T^d$ and $\boldsymbol{y} \in \mathcal{F}$. Then

$$0 \leq \boldsymbol{y}^T \left(\boldsymbol{C} - t_0 \boldsymbol{e}_0 \boldsymbol{e}_0^T + \sum_{i=1}^m t_i \boldsymbol{P}_i \right) \boldsymbol{y} = \boldsymbol{c}^T \boldsymbol{y} - t_0 + \sum_{i=1}^m t_i \boldsymbol{y}^T \boldsymbol{P}_i \boldsymbol{y} \leq \boldsymbol{c}^T \boldsymbol{y} - t_0.$$

(This proof is essentially due to [17]). ∎

The SDP (16) is corresponding to the Lagrangian dual of the QP (1). See the papers [13, 14, 15] for details.

It is easily verified that the SDPs (6) and (16) are dual to each other. Hence, from the duality theorem (see, for example, Theorem 4.2.1 of [12]) and Lemma 2.1, we obtain:

Lemma 4.2 (Duality between (3) and (16))

1. *If $\boldsymbol{t} = (t_0, t_1, \ldots, t_m)^T \in R^{1+m}$ is a feasible solution of the SDP (16) and $\boldsymbol{y} \in R^{1+n}$ a feasible solution of the problem (3), their objective values t_0 and $\boldsymbol{c}^T \boldsymbol{y}$ satisfy $t_0 \leq \boldsymbol{c}^T \boldsymbol{y}$; $\sup\{t_0 \ : \ \boldsymbol{t} \in T^d\} \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} \ : \ \boldsymbol{y} \in \widehat{\mathcal{F}}\}$.*

2. *Suppose that Condition 2.2 holds and that $-\infty < \hat{g} \equiv \inf\{\boldsymbol{c}^T \boldsymbol{y} \ : \ \boldsymbol{y} \in \widehat{\mathcal{F}}\}$. Then the SDP (16) has a maximum solution $\boldsymbol{t}^* \in R^{1+m}$ with the maximum objective value $t_0^* = \hat{g}$.* ∎

The lemma below establishes a weak duality relation between the convex minimization problem (3) and the SDP (16).

Lemma 4.3 *If $\boldsymbol{t} = (t_0, t_1, \ldots, t_m) \in R^{1+m}$ is a feasible solution of the SDP (16) and $\boldsymbol{y} \in R^{1+n}$ a feasible solution of the problem (3), their objective values t_0 and $\boldsymbol{c}^T \boldsymbol{y}$ satisfy $t_0 \leq \boldsymbol{c}^T \boldsymbol{y}$;*
$\sup\{t_0 \ : \ \boldsymbol{t} \in T^d\} \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} \ : \ \boldsymbol{y} \in \widetilde{\mathcal{F}}\}$.

Proof: Suppose that $\boldsymbol{t} \in T^d$ and $\boldsymbol{y} \in R^{1+n} \in \widetilde{\mathcal{F}}$. Let

$$\boldsymbol{Z} \equiv \begin{pmatrix} \zeta & \boldsymbol{w}^T \\ \boldsymbol{w} & \boldsymbol{Q} \end{pmatrix} = \boldsymbol{C} - t_0 \boldsymbol{e}_0 \boldsymbol{e}_0^T + \sum_{k=1}^{m} t_k \boldsymbol{P}_k,$$

where $\zeta \in R$, $\boldsymbol{w} \in R^n$ and $\boldsymbol{Q} \in \mathcal{S}(n)$. We see by the definitions of the matrices $\boldsymbol{C}$, $\boldsymbol{e}_0\boldsymbol{e}_0^T$, $\boldsymbol{P}_k \in \mathcal{S}(1+n)$ $(k = 1, 2, \ldots, m)$ that $\boldsymbol{Q} = \sum_{k=1}^{m} t_k \boldsymbol{Q}_k$. On the other hand, it follows from $\boldsymbol{Z} \in \mathcal{S}(1+n)_+$ that $\boldsymbol{Q} = \sum_{k=1}^{m} t_k \boldsymbol{Q}_k \in \mathcal{S}(n)_+$. Hence we obtain from $\boldsymbol{y} \in \widetilde{\mathcal{F}}$ that $\boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k \right) \boldsymbol{y} \leq 0$. Consequently,

$$0 \leq \boldsymbol{y}^T \boldsymbol{Z} \boldsymbol{y} = \boldsymbol{y}^T \boldsymbol{C} \boldsymbol{y} - t_0 \boldsymbol{y}^T \boldsymbol{e}_0 \boldsymbol{e}_0^T \boldsymbol{y} + \boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k \right) \boldsymbol{y} \leq \boldsymbol{c}^T \boldsymbol{y} - t_0.$$

∎

5 Proof of the Main Theorem.

(i) The first inclusion relation $\mathcal{F} \subseteq \widehat{\mathcal{F}}$ follows from Lemma 3.1. To prove the second inclusion relation $\widehat{\mathcal{F}} \subseteq \widetilde{\mathcal{F}}$, assume that $\boldsymbol{y} \in \widehat{\mathcal{F}}$. Then there exists a $\boldsymbol{Y} \in \widehat{\mathcal{G}}$ such that $\boldsymbol{y} = \boldsymbol{Y} \boldsymbol{e}_0$; specifically $\boldsymbol{Y}$ satisfies $0 \geq \boldsymbol{P}_k \bullet \boldsymbol{Y}$ $(k = 1, 2, \ldots, m)$. Hence

$$\left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k \right) \bullet \boldsymbol{Y} \leq 0 \text{ for every } \boldsymbol{t} = (t_1, t_2, \ldots, t_m)^T \geq \boldsymbol{0}.$$

By Lemma 3.2, we see that

$$\boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k \right) \boldsymbol{y} \leq 0 \text{ whenever } \sum_{k=1}^{m} t_k \boldsymbol{Q}_k \in \mathcal{S}(n)_+.$$

This implies $\boldsymbol{y} \in \widetilde{\mathcal{F}}$. Thus we have shown that $\widehat{\mathcal{F}} \subseteq \widetilde{\mathcal{F}}$.

(ii) Since $\widehat{\mathcal{F}} \subseteq \widetilde{\mathcal{F}}$, we know that

$$\inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widetilde{\mathcal{F}}\} \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widehat{\mathcal{F}}\} \tag{17}$$

for every $\boldsymbol{c} \in R^{1+n}$. Let $\boldsymbol{c} \in R^{1+n}$ be fixed arbitrarily. If $\inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widehat{\mathcal{F}}\} = -\infty$ then $\inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widetilde{\mathcal{F}}\} = -\infty$ by (17). Hence we obtain the equality (9). So assume that $\hat{g} \equiv \inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widehat{\mathcal{F}}\} > -\infty$. By Lemma 4.2, there exists a maximum solution $\boldsymbol{t}^* = (t_0^*, t_1^*, \ldots, t_m^*)^T \in R^{1+m}$ of the SDP (16) with the objective value $t_0^* = \hat{g}$. We also see by Lemma 4.3 that $t_0^* \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widetilde{\mathcal{F}}\}$. Therefore

$$\inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widetilde{\mathcal{F}}\} \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widehat{\mathcal{F}}\} = \hat{g} = t_0^* \leq \inf\{\boldsymbol{c}^T \boldsymbol{y} : \boldsymbol{y} \in \widetilde{\mathcal{F}}\}.$$

Thus we have shown the equality (9). By the construction, $\widetilde{\mathcal{F}}$ is a closed convex subset of R^{1+n} and $\widehat{\mathcal{F}}$ is a convex subset of R^{1+n}. Hence the identity (9) for every $\boldsymbol{c} \in R^{1+n}$

implies that $\tilde{\mathcal{F}} = \mathrm{cl}\ \hat{\mathcal{F}}$.

(iii) Assume on the contrary that $\boldsymbol{y} \notin \mathcal{F}$ for some strictly convex boundary point $\boldsymbol{y}$ of $\tilde{\mathcal{F}}$. It follows from $\boldsymbol{y} \notin \mathcal{F}$ that $\boldsymbol{y}^T \boldsymbol{P}_j \boldsymbol{y} > 0$ for some $j \in \{1, 2, \ldots, m\}$. Since $\boldsymbol{y}$ is a strictly convex boundary point of $\tilde{\mathcal{F}}$, there exists some $\boldsymbol{t} = (t_1, t_2, \ldots, t_m)^T \geq 0$ for which

$$y_0 = 1,\ \boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k \right) \boldsymbol{y} = 0 \ \text{ and } \ \sum_{k=1}^{m} t_k \boldsymbol{Q}_k \in \mathcal{S}(n)_{++}$$

holds. Hence if $\epsilon > 0$ is sufficiently small, we obtain

$$y_0 = 1,\ \boldsymbol{y}^T \left(\sum_{k=1}^{m} t_k \boldsymbol{P}_k + \epsilon \boldsymbol{P}_j \right) \boldsymbol{y} > 0 \ \text{ and } \ \sum_{k=1}^{m} t_k \boldsymbol{Q}_k + \epsilon \boldsymbol{Q}_j \in \mathcal{S}(n)_{++},$$

which is a contradiction to the assumption that $\boldsymbol{y} \in \tilde{\mathcal{F}}$. This completes the proof of the main theorem.

6 Concluding Discussion.

The effectiveness of the SDP relaxation for a nonconvex QP (or a 0-1 IP) depends on the representation of its feasible region using linear and/or quadratic inequalities. Suppose that the feasible region $\mathcal{F}$ of the canonical form QP (1) is bounded and involves some linear inequality constraints

$$\boldsymbol{y}^T \boldsymbol{P}_k \boldsymbol{y} \equiv \pi_k + \boldsymbol{q}_k^T \boldsymbol{x} \leq 0 \ (k \in K),$$

where $\boldsymbol{y} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}$, and $K \subseteq \{1, 2, \ldots, m\}$. Let S denote the polyhedral region determined by these linear inequalities;

$$S = \left\{ \boldsymbol{y} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \ : \ \pi_k + \boldsymbol{q}_k^T \boldsymbol{x} \leq 0 \ (k \in K) \right\}.$$

We want to cut off all the vertices of S that do not lie in $\mathcal{F}$ when we apply the SDP relaxation. As we will see below, this is always possible if we replace those linear inequality constraints by convex-quadratic inequality constraints

$$\boldsymbol{y}^T \boldsymbol{P}'_k \boldsymbol{y} \equiv (\pi_k + \boldsymbol{q}_k^T \boldsymbol{x})(\pi'_k + \boldsymbol{q}_k^T \boldsymbol{x}) \leq 0 \ (k \in K),$$

where

$$\boldsymbol{P}'_k \equiv \begin{pmatrix} \pi'_k \pi_k & (\pi'_k + \pi_k) \boldsymbol{q}_k^T / 2 \\ (\pi'_k + \pi_k) \boldsymbol{q}_k / 2 & \boldsymbol{Q}'_k \end{pmatrix} \in \mathcal{S}(1+n),\ \boldsymbol{Q}'_k \equiv \boldsymbol{q}_k \boldsymbol{q}_k^T \in \mathcal{S}(n)_+ \ (k \in K),$$

and π'_k $(k \in K)$ are sufficiently large numbers such that

$$\pi_k + \boldsymbol{q}_k^T \boldsymbol{x} \leq 0 \leq \pi'_k + \boldsymbol{q}_k^T \boldsymbol{x} \ (k \in K) \ \text{ for every } \boldsymbol{y} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \mathcal{F}. \tag{18}$$

Let

$$\mathcal{F}' = \left\{ \boldsymbol{y} \in R^{1+n} \ : \begin{array}{l} y_0 = 1, \\ \boldsymbol{y}^T \boldsymbol{P}_i \boldsymbol{y} \leq 0 \ (i \in \{1,2,\ldots,m\} \backslash K), \\ \boldsymbol{y}^T \boldsymbol{P}'_k \boldsymbol{y} \leq 0 \ (k \in K) \end{array} \right\}$$

The condition (18) above on π'_k $(k \in K)$ ensures that $\mathcal{F}' = \mathcal{F}$. Suppose that $\boldsymbol{y} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}$ is a vertex of S. Then there exists a subset K' of K such that

$$\boldsymbol{y}^T \boldsymbol{P}_k \boldsymbol{y} = \pi_k + \boldsymbol{q}_k^T \boldsymbol{x} = 0 \ (k \in K') \ \text{ and } \{\boldsymbol{q}_k \in R^n \ : \ k \in K'\} \quad \text{forms a basis of } R^n.$$

Hence

$$\sum_{k \in K'} \boldsymbol{y}^T \boldsymbol{P}'_k \boldsymbol{y} = 0 \ \text{ and } \sum_{k \in K'} \boldsymbol{Q}'_k \in \mathcal{S}(n)_{++}.$$

If in addition $\boldsymbol{y} \in \widetilde{\mathcal{F}}'$ then $\boldsymbol{y}$ is a strictly convex boundary point of $\widetilde{\mathcal{F}}'$; hence $\boldsymbol{y} \in \mathcal{F}'$ by (iii) of Theorem 2.3. Therefore we can conclude that every vertex $\boldsymbol{y}$ of S belongs to $\widetilde{\mathcal{F}}'$ if and only if $\boldsymbol{y} \in \mathcal{F}'$.

Acknowledgment. The authors are grateful to Professors Stephen Boyd, Farid Alizadeh, Henry Wolkowicz for their helpful comments and references. In particular, the paper [14] was brought to the authors by Stephen Boyd.

References

[1] W. F. Alizadeh, "Interior point methods in semidefinite programming with application to combinatorial optimization," *SIAM Journal on Optimization* **5** (1995) 13–51.

[2] W. F. Alizadeh, J. -P. A. Haeberly and M. L. Overton, "Primal-dual interior-point methods for semidefinite programming," 1994.

[3] S. Boyd, L. E. Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory,* (SIAM, Philadelphia, 1994).

[4] R. M. Freund, "Complexity of an algorithm for finding an approximate solution of a semidefinite program with no regularity assumption," Technical report OR 302-94, Operations Research Center, MIT, 1994.

[5] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of Assoc. Comput. Mach.* to appear. . A preliminary version appeared in *Proceedings of the 26th Annual ACM Symposium on Theory of Computing* (1994) 422–431.

[6] C. Helmberg, S. Poljak, F. Rendl and H. Wolkowicz, "Combining semidefinite and polyhedral relaxation for integer programs," Lecture note in Computer Science 538 (1995) 124–134.

[7] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz, "An interior-point method for semidefinite programming," SIAM Journal on Optimization, to appear.

[8] D. E. Knuth, "The sandwich theorem," *Electronic Journal of Combinatorics* 1 (1995) 1–48.

[9] M. Kojima, S. Shindoh and S. Hara, "Interior-point methods for the monotone semidefinite linear complementarity problems," Research Report #282, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, April 1994, Revised April 1995.

[10] L. Lovász, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory* 25 (1979) 1–7.

[11] L. Lovász and A. Schrijver, "Cones of matrices and set functions and 0-1 optimization," *SIAM Journal on Optimization* 1 (1991) 166–190.

[12] Ju. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications* (SIAM, Philadelphia, 1993).

[13] S. Poljak, F. Rendl and H. Wolkowicz, "A recipe for semidefinite relaxation for (0,1)-quadratic programming," *Journal of Global Optimization* 7 (1995) 51–73.

[14] N. Z. Shor, "Quadratic optimization problems," *Soviet Journal of Computer and Systems Sciences* 25 (1987) 1–11.

[15] N. Z. Shor, "Dual quadratic estimates in polynomial and boolean programming," *Annals of Operations Research* 25 (1990) 163–168.

[16] L. Vandenberghe and S. Boyd, "A primal-dual potential reduction method for problems involving matrix inequalities," *Mathematical Programming* 69 (1995) 205–236.

[17] L. Vandenberghe and S. Boyd, " Semidefinite Programming," Informations Systems Laboratory, Stanford University, 1994.