

Global Quadratic Optimization via Conic Relaxation

Yu. Nesterov *

February 21, 1998

Abstract

We present a convex conic relaxation for a problem of maximizing an indefinite quadratic form over a set of convex constraints on the squared variables. We show that for all these problems we get at least $\frac{12}{37}$ -relative accuracy of the approximation. In the second part of the paper we derive the conic relaxation by another approach based on the second order optimality conditions. We show that for l_p -balls, $p \geq 2$, intersected by a linear subspace, it is possible to guarantee $(1 - \frac{2}{p})$ -relative accuracy of the solution. As a consequence, we prove $(1 - \frac{1}{e \ln n})$ -relative accuracy of the conic relaxation for an indefinite quadratic maximization problem over an n -dimensional unit box with homogeneous linear equality constraints. We discuss the implications of the results for the discussion around the question $P = NP$.

*CORE, Catholic University of Louvain, 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium;
e-mail: nesterov@core.ucl.ac.be

1 Introduction

Starting from the pioneering paper [2], there were obtained several results [4, 7, 5], which show that a solution of an indefinite quadratic maximization problem with some linear constraints on the squared variables can be approximated with a constant relative accuracy. In this paper we present some improvements and extensions of the results [5].

In Section 2 we consider a problem of maximizing an indefinite quadratic form subject to arbitrary convex constraints on the squared variables. For convenience of the dual description we use a conic representation of these constraints. We introduce a convex conic relaxation for that problem and prove that it provides us at least with an approximation of $\frac{\pi-2}{6-\pi}$ relative accuracy. In Section 3 we show how to improve this approximation using the diagonal elements of the quadratic objective function. The relative accuracy, which we can get in this case, is $\frac{12}{37}$. In Section 4 we extend the results of Section 2, 3 onto the case of general convex constraints on squared variables. We conclude the first part of the paper with a discussion of the difficulties which arise in the problems with linear equality constraints on the initial variables (Section 5).

In the second part of the paper, which starts from Section 6, we study another way of deriving the conic relaxation. This approach can be applied only to a small number of sets (l_p -balls, $p \geq 2$), but it allows to treat also the linear equations. We prove that for such problems the conic relaxation gives $(1 - \frac{2}{p})$ relative accuracy. In Section 7 we apply these results to a problem of maximizing a quadratic function over a unit box subject to a system of homogeneous linear equalities. We show that it is possible to compute in polynomial time a $(1 - \frac{1}{e \ln n})$ -solution of that problem. We conclude the paper with a discussion of the results.

In this paper we use the following notation. For two vectors $x, y \in R^n$ we denote $\langle x, y \rangle$ the standard inner product:

$$\langle x, y \rangle = \sum_{i=1}^n x^{(i)} y^{(i)}.$$

Then $\|x\| = \langle x, x \rangle^{1/2}$. Since we work in several finite-dimensional spaces, the meaning of this notation is defined by the spaces of the arguments. For example, for two $(m \times n)$ -matrices X and Y we have

$$\langle X, Y \rangle = \sum_{i=1}^m \sum_{j=1}^n X_{ij} Y_{ij}.$$

We use the standard notation for l_p -norms:

$$\|x\|_p = \left[\sum_{i=1}^n |x^{(i)}|^p \right]^{1/p}, \quad x \in R^n, \quad p \geq 1.$$

Again, the meaning of the notation depends on the dimension of space of the argument. Recall, that for $p = \infty$ we have $\|x\|_\infty = \max_{1 \leq i \leq n} |x^{(i)}|$. The norm dual to $\|\cdot\|_p$ is $\|\cdot\|_{p^*}$ with $p^* = \frac{p}{p-1}$:

$$\|y\|_{p^*} = \max\{\langle y, x \rangle : \|x\|_p \leq 1\}.$$

For a symmetric matrix A we write $A \succeq 0$ if A is positive semidefinite. Notation $B \succeq A$ means that $B - A \succeq 0$. For $x \in R^n$ we denote $D(x)$ the diagonal $(n \times n)$ -matrix with

diagonal entries $x^{(i)}$. Conversely, $d(X) \in R^n$ denotes the diagonal of an $(n \times n)$ -matrix X . In particular, $\langle D(x), Y \rangle = \langle x, d(Y) \rangle$. Notation e_i is used for the i th coordinate vector of R^n and $1_n \in R^n$ stands for the vector of all ones. Thus, $I_n = D(1_n)$ is a unit matrix. Notation 0_n is used for the zero vector in R^n .

We use square brackets in order to indicate the component-wise operations with the vectors. For example, notation $[x \cdot y]$ stays for the vector with components $x^{(i)}y^{(i)}$, $x, y \in R^n$. Notation $[x]^2$ is used for the vector with the components $(x^{(i)})^2$. If $f(\tau)$ is a univariate function, we denote $f[x]$ the vector with the components $f(x^{(i)})$. In order to indicate the partial ordering in R^n we use the usual inequality signs. Thus, $x \geq y$ for x and y from R^n means that $x^{(i)} \geq y^{(i)}$, $i = 1, \dots, n$.

Finally, $[\alpha, \beta]^n$ denotes a continuous box in R^n , that is $\{x \in R^n : \alpha 1_n \leq x \leq \beta 1_n\}$. For a Boolean box $\{x \in R^n : x^{(i)} = (\alpha \text{ or } \beta)\}$ we use notation $\{\alpha, \beta\}^n$.

2 Convex conic constraints on squared variables

Let Q be an arbitrary symmetric $(n \times n)$ -matrix. Consider the following pair of optimization problems:

$$\begin{aligned} \text{find } \phi^* = \max \{ \langle Qx, x \rangle : [x]^2 \in \mathcal{F} \}, \\ \text{find } \phi_* = \min \{ \langle Qx, x \rangle : [x]^2 \in \mathcal{F} \}. \end{aligned} \quad (2.1)$$

where $\mathcal{F}$ is a closed convex set. Our main assumption on the problem (2.1) is as follows.

Assumption 2.1 1). The set $\mathcal{F}$ is bounded. 2). There exists a strictly positive $v \in \mathcal{F}$.

In order to simplify the dual analysis, in this section we assume that the feasible set $\mathcal{F}$ is presented in a conic form:

$$\mathcal{F} = \{v \in K : Av = b\}, \quad (2.2)$$

where K is a convex closed pointed cone in R^n with non-empty interior, A is an $(m \times n)$ -matrix and $b \neq 0_m$. Our additional assumption on the set $\mathcal{F}$ is as follows.

Assumption 2.2 $\{v \in \text{int } K : Av = b\} \neq \emptyset$.

Note that the form (2.2) is quite general, since any bounded convex set can be written in this way (see [6] for details). At the same time, in Section 4 we will show how to transform our result on the case of a general convex feasible set $\mathcal{F}$, not necessarily presented in a conic form.

Using the same technique as in [5], we can rewrite the pair of problems (2.1) in a trigonometric form.

Lemma 2.1

$$\begin{aligned} \phi^* = \max_{\substack{X \succeq 0, d(X) = 1_n, \\ d \geq 0, [d]^2 \in \mathcal{F}}} \frac{2}{\pi} \langle Q, D(d) \arcsin[X] D(d) \rangle, \\ \phi_* = \min_{\substack{X \succeq 0, d(X) = 1_n, \\ d \geq 0, [d]^2 \in \mathcal{F}}} \frac{2}{\pi} \langle Q, D(d) \arcsin[X] D(d) \rangle, \end{aligned} \quad (2.3)$$

Proof:

Indeed, let us represent a vector $x \in R^n$ as follows:

$$x = [d \cdot \sigma], \quad d \geq 0 \in R^n, \quad \sigma \in \{-1, 1\}^n.$$

Note that $[x]^2 = [d]^2$. Therefore $\phi^* = \max_d \{\Phi(d) : d \geq 0, [d]^2 \in \mathcal{F}\}$ with

$$\Phi(d) = \max\{\langle D(d)QD(d)\sigma, \sigma \rangle : \sigma \in \{-1, 1\}^n\}.$$

Using Theorem 2.3 [5], we can represent $\Phi(d)$ in the following form:

$$\Phi(d) = \max\left\{\frac{2}{\pi} \langle D(d)QD(d), \arcsin[X] \rangle : X \succeq 0, d(X) = 1_n\right\}.$$

Inserting this representation in the above expression for ϕ^* we get the first statement of the lemma. The second one can be obtained in a similar way. $\square$

Note that in general none of the problems (2.1) is convex in x . Therefore, in order to estimate their optimal values, we need to use a kind of convex relaxation. Let us define the *conic relaxation* of problems (2.1):

$$\begin{aligned} \psi^* &= \max\{\langle Q, X \rangle : d(X) \in \mathcal{F}, X \succeq 0\}, \\ \psi_* &= \min\{\langle Q, X \rangle : d(X) \in \mathcal{F}, X \succeq 0\}. \end{aligned} \tag{2.4}$$

The goal of this section is to relate these values with the extremal values of the problem (2.1).

Sometimes it is convenient to use a dual form of these relaxations. Recall that for a convex cone $K \subseteq R^n$ the dual cone K^* is defined as follows:

$$K^* = \{u \in R^n : \langle u, v \rangle \geq 0, \forall v \in K\}.$$

Lemma 2.2

$$\begin{aligned} \psi^* &= \min_{y \in R^m, u \in R^n} \{\langle b, y \rangle : Q + D(u) \preceq D(A^T y), u \in K^*\}, \\ \psi_* &= \max_{y \in R^m, u \in R^n} \{\langle b, y \rangle : Q \succeq D(u) + D(A^T y), u \in K^*\}. \end{aligned} \tag{2.5}$$

Proof:

In view of Assumptions 2.1, 2.2, we can get a dual representation of the upper relaxation ψ^* as follows:

$$\begin{aligned} \psi^* &= \max_{X, v} \{\langle Q, X \rangle : Ad(X) = b, d(X) = v, X \succeq 0, v \in K\} \\ &= \max_{X \succeq 0, v \in K} \min_{y \in R^m, u \in R^n} \{\langle Q, X \rangle + \langle y, b - Ad(X) \rangle + \langle u, d(X) - v \rangle\} \\ &= \min_{y \in R^m, u \in R^n} \max_{X, v} \{\langle Q + D(u - A^T y), X \rangle + \langle b, y \rangle - \langle u, v \rangle : X \succeq 0, v \in K\} \\ &= \min_{y \in R^m, u \in R^n} \{\langle b, y \rangle : Q + D(u) \preceq D(A^T y), u \in K^*\}. \end{aligned}$$

Similarly, for the lower relaxation we get the following:

$$\begin{aligned}
\psi_* &= \min_{X,v} \{ \langle Q, X \rangle : Ad(X) = b, d(X) = v, X \succeq 0, v \in K \} \\
&= \min_{X \succeq 0, v \in K} \max_{y \in R^m, u \in R^n} \{ \langle Q, X \rangle + \langle y, b - Ad(X) \rangle + \langle u, v - d(X) \rangle \} \\
&= \max_{y \in R^m, u \in R^n} \min_{X,v} \{ \langle Q - D(u + A^T y), X \rangle + \langle b, y \rangle + \langle u, v \rangle : X \succeq 0, v \in K \} \\
&= \max_{y \in R^m, u \in R^n} \{ \langle b, y \rangle : Q \succeq D(u) + D(A^T y), u \in K^* \}.
\end{aligned}$$

□

Let us establish some relations between the relaxations (2.4) and the optimal values of the problems (2.1). Denote

$$\psi(\alpha) = \alpha\psi^* + (1 - \alpha)\psi_*. \quad (2.6)$$

The proof of the following theorem is similar to that of Theorem 3.3 [5].

Theorem 2.1

$$\psi_* \leq \phi_* \leq \psi(1 - \frac{2}{\pi}) \leq \psi(\frac{2}{\pi}) \leq \phi^* \leq \psi^*. \quad (2.7)$$

Proof:

Note that $\psi_* \leq \psi^*$ by definition. So, the middle inequality in (2.7) is correct. Further, if $[x]^2 \in \mathcal{F}$ then the matrix $X = xx^T$ is feasible for both relaxation problems (2.4) since $d(X) = [x]^2$. Moreover, $\langle Q, X \rangle = \langle Qx, x \rangle$. Thus, both bounding inequalities in the chain (2.7) are valid. Let us prove now two remaining inequalities.

Let us choose arbitrary $u \in K$ and $y \in R^m$, which satisfy the constraints of the dual form (2.5) of the lower relaxation ψ_* :

$$(u, y) \in \mathcal{F}_d = \{(u, y) \in K^* \times R^m : Q \succeq D(u) + D(A^T y)\}. \quad (2.8)$$

Consider a pair (X, d) , which satisfies the constraints of the trigonometric representation (2.3) for ϕ^* :

$$X \succeq 0, \quad d(X) = 1_n, \quad d \geq 0, \quad A[d]^2 = b, \quad [d]^2 \in K. \quad (2.9)$$

Since $X \succeq 0$ and $|X_{ij}| \leq 1$ we have $\arcsin[X] \succeq X$ in view of Corollary 3.2 [5]. Therefore, using Lemma 2.1 we get the following:

$$\begin{aligned}
\phi^* &\geq \frac{2}{\pi} \langle D(d)QD(d), \arcsin[X] \rangle \\
&= \frac{2}{\pi} \langle D(d)(Q - D(u) - D(A^T y))D(d), \arcsin[X] \rangle + \langle u + A^T y, [d]^2 \rangle \\
&\geq \frac{2}{\pi} \langle D(d)(Q - D(u) - D(A^T y))D(d), X \rangle + \langle u + A^T y, [d]^2 \rangle \\
&= \frac{2}{\pi} \langle Q, D(d)XD(d) \rangle + (1 - \frac{2}{\pi}) \langle u + A^T y, [d]^2 \rangle.
\end{aligned}$$

Note that $u \in K^*$ and $[d]^2 \in K$. Therefore $\langle u, [d]^2 \rangle \geq 0$. In view of (2.9) we have

$$\langle A^T y, [d]^2 \rangle = \langle A[d]^2, y \rangle = \langle b, y \rangle.$$

Finally, for any pair (X, d) , which satisfy (2.9) we have $Y = D(d)XD(d)$ feasible for the primal relaxation problems:

$$Y \in \mathcal{F}_p = \{Y : Y \succeq 0, Ad(Y) = b, d(Y) \in K\}.$$

On the other hand, any $Y \in \mathcal{F}_p$ can be represented as $Y = D(d)XD(d)$ with X and d , which satisfy (2.9). Therefore, we conclude that

$$\phi^* \geq \frac{2}{\pi} \langle Q, Y \rangle + (1 - \frac{2}{\pi}) \langle b, y \rangle, \quad \forall Y \in \mathcal{F}_p, (u, y) \in \mathcal{F}_d.$$

This proves the forth inequality in the chain (2.7). The remaining inequality can be proved in a similar way. $\square$

Definition 2.1 *We say that the value ψ approximates ϕ^* with a relative accuracy $\mu \in [0, 1]$ if $|\psi - \phi^*| \leq \mu(\phi^* - \phi_*)$. We call this approximation implementable if $\psi \leq \phi^*$.*

Corollary 2.1 *1. Let $\alpha = \frac{2}{\pi}$. Then the value $\psi(\alpha)$ is an implementable approximation of ϕ^* with the relative accuracy $\mu = \frac{\pi}{2} - 1 < \frac{4}{7}$.*

2. Let $\beta = \frac{(1+\alpha)^2-2}{3\alpha-1}$. Then the value $\psi(\beta)$ approximates ϕ^ with the relative accuracy $\mu = \frac{\pi-2}{6-\pi} < \frac{2}{5}$.*

The proof of that statement is exactly the same as that of Corollary 3.4 in [5].

3 Using additional information

In this section it is shown how to improve the quality of our bounds by taking into account some additional information. Define

$$\begin{aligned} \tau^* &= \max\{\langle d(Q), v \rangle : v \geq 0, v \in \mathcal{F}\}, \\ \tau_* &= \min\{\langle d(Q), v \rangle : v \geq 0, v \in \mathcal{F}\}. \end{aligned} \tag{3.1}$$

Note that these values are computable in polynomial time. In view of Lemma 2.1 with $X = I$ and $[d]^2 = v$ we have

$$\phi_* \leq \tau_* \leq \tau^* \leq \phi^*. \tag{3.2}$$

Hence,

$$\beta^* \equiv \frac{\psi^* - \tau^*}{\psi^* - \psi_*} \in [0, 1], \quad \beta_* \equiv \frac{\tau_* - \psi_*}{\psi^* - \psi_*} \in [0, 1].$$

Using these values we can express τ^* and τ_* as follows:

$$\begin{aligned} \tau^* &= \psi^* - \beta^*(\psi^* - \psi_*) = \psi(1 - \beta^*), \\ \tau_* &= \psi_* + \beta_*(\psi^* - \psi_*) = \psi(\beta_*). \end{aligned}$$

Denote $\omega(\beta) = \beta \arcsin(\beta) + \sqrt{1-\beta^2} \equiv 1 + \int_0^\beta \arcsin(\tau) d\tau$, $\beta \in [0, 1]$. This function is increasing and convex with $\omega(0) = 1$ and $\omega(1) = \frac{\pi}{2}$. In what follows we denote $\bar{\beta}$ the unique root of the following equation:

$$\frac{2}{\pi}\omega(\beta) = 1 - \beta, \quad \beta \in [0, 1].$$

It can be shown that $\frac{23}{70} < \bar{\beta} < \frac{24}{73}$.

Theorem 3.1 1. Denote

$$\alpha^* = \max\{\frac{2}{\pi}\omega(\beta_*), 1 - \beta^*\},$$

$$\alpha_* = \min\{1 - \frac{2}{\pi}\omega(\beta^*), \beta_*\}.$$

The optimal values of the problems (2.1) satisfy the following relations:

$$\psi^* \geq \phi^* \geq \psi(\alpha^*), \quad (3.3)$$

$$\psi_* \leq \phi_* \leq \psi(\alpha_*). \quad (3.4)$$

2. The value $\psi(\alpha^*)$ is an implementable approximation of ϕ^* with relative accuracy

$$\mu = \frac{1-\alpha^*}{1-\alpha_*} \leq \frac{\bar{\beta}}{1-\beta} < \frac{24}{49}.$$

3. Denote $\bar{\alpha} = \frac{\alpha^*(2-\alpha_*)-\alpha_*}{1+\alpha^*-2\alpha_*}$. The value $\psi(\bar{\alpha})$ is a μ -approximation of ϕ^* with

$$\mu = \frac{1-\alpha^*}{1+\alpha^*-2\alpha_*} \leq \frac{\bar{\beta}}{2-3\beta} < \frac{12}{37}.$$

In Items 2 and 3 the upper bounds are achieved for $\beta^* = \beta_* = \bar{\beta}$.

Proof:

Let $X \succeq 0$ and $d \geq 0$ be feasible for the trigonometric form of the upper relaxation (2.3):

$$d(X) = 1_n, \quad A[d]^2 = b, \quad [d]^2 \in K.$$

Consider the matrices $X_\gamma = \gamma X + (1 - \gamma)I_n$, $\gamma \in [0, 1]$. Then

$$\arcsin[X_\gamma] = \arcsin[\gamma X] + (\frac{\pi}{2} - \arcsin(\gamma))I_n.$$

Therefore

$$\begin{aligned} \phi^* &\geq \frac{2}{\pi} \langle Q, D(d) \arcsin[X_\gamma] D(d) \rangle \\ &= \frac{2}{\pi} \langle Q, D(d) \arcsin[\gamma X] D(d) \rangle + \left(1 - \frac{2}{\pi} \arcsin(\gamma)\right) \langle d(Q), [d]^2 \rangle. \end{aligned} \quad (3.5)$$

Let us choose now arbitrary $u \in K$ and $y \in R^m$ which satisfy the constraints of the dual form (2.5) of the lower relaxation ψ_* :

$$(u, y) \in \mathcal{F}_d = \{(u, y) \in K^* \times R^m : Q \succeq D(u) + D(A^T y)\}.$$

Then, in view of Corollary 3.2 [5] we have:

$$\begin{aligned}
\langle Q, D(d) \arcsin[\gamma X] D(d) \rangle &= \langle Q - D(u) - D(A^T y), D(d) \arcsin[\gamma X] D(d) \rangle \\
&\quad + \langle D(u + A^T y), D(d) \arcsin[\gamma X] D(d) \rangle \\
&\geq \gamma \langle Q - D(u) - D(A^T y), D(d) X D(d) \rangle \\
&\quad + \arcsin(\gamma) \langle u + A^T y, [d]^2 \rangle \\
&= \gamma \langle Q, D(d) X D(d) \rangle + (\arcsin(\gamma) - \gamma) \langle u + A^T y, [d]^2 \rangle.
\end{aligned}$$

Note that $\arcsin(\gamma) \geq \gamma$ for $\gamma \in [0, 1]$. At the same time $u \in K^*$ and $[d]^2 \in K$. Therefore $\langle u, [d]^2 \rangle \geq 0$. Finally, $\langle A^T y, [d]^2 \rangle = \langle A[d]^2, y \rangle = \langle b, y \rangle$. Thus,

$$\langle Q, D(d) \arcsin[\gamma X] D(d) \rangle \geq \gamma \langle Q, D(d) X D(d) \rangle + (\arcsin(\gamma) - \gamma) \langle b, y \rangle.$$

Substituting this inequality in (3.5) we get the following:

$$\begin{aligned}
\phi^* &\geq \frac{2}{\pi} (\gamma \langle Q, D(d) X D(d) \rangle + (\arcsin(\gamma) - \gamma) \langle b, y \rangle) + \left(1 - \frac{2}{\pi} \arcsin(\gamma)\right) \langle d(Q), [d]^2 \rangle \\
&\geq \frac{2}{\pi} (\gamma \langle Q, D(d) X D(d) \rangle + (\arcsin(\gamma) - \gamma) \langle b, y \rangle) + \left(1 - \frac{2}{\pi} \arcsin(\gamma)\right) \tau_*.
\end{aligned}$$

Using the same reasoning as in Theorem 2.1, we conclude that

$$\begin{aligned}
\phi^* &\geq \frac{2}{\pi} \gamma \psi^* + \frac{2}{\pi} (\arcsin(\gamma) - \gamma) \psi_* + \left(1 - \frac{2}{\pi} \arcsin(\gamma)\right) \tau_* \\
&= \frac{2}{\pi} \arcsin(\gamma) \psi \left(\frac{\gamma}{\arcsin(\gamma)} \right) + \left(1 - \frac{2}{\pi} \arcsin(\gamma)\right) \psi(\beta_*) \\
&= \psi \left(\frac{2\gamma}{\pi} + \left(1 - \frac{2}{\pi} \arcsin(\gamma)\right) \beta_* \right).
\end{aligned}$$

The right-hand side of the above inequality is maximal for $\gamma^* = \sqrt{1 - \beta_*^2}$. Then

$$\frac{2\gamma^*}{\pi} + \left(1 - \frac{2}{\pi} \arcsin(\gamma^*)\right) \beta_* = \frac{2}{\pi} \left(\sqrt{1 - \beta_*^2} + \beta_* \arccos(\gamma^*) \right) = \frac{2}{\pi} \omega(\beta_*).$$

Thus, $\phi^* \geq \psi(\frac{2}{\pi} \omega(\beta_*))$. Combining this inequality with (3.2), we get the lower bound in (3.3). The relations (3.4) for ψ_* can be obtained in a similar way.

Let us prove now Item 2 of the theorem. In view of (3.3) and (3.4) we have

$$0 \leq \frac{\phi^* - \psi(\alpha^*)}{\phi^* - \phi_*} \leq \frac{\psi^* - \psi(\alpha^*)}{\psi^* - \phi_*} \leq \frac{\psi^* - \psi(\alpha^*)}{\psi^* - \psi(\alpha_*)} = \frac{\psi(1) - \psi(\alpha^*)}{\psi(1) - \psi(\alpha_*)} = \frac{1 - \alpha^*}{1 - \alpha_*}. \quad (3.6)$$

Note that

$$1 - \alpha^* = 1 - \max\left\{\frac{2}{\pi} \omega(\beta_*), 1 - \beta^*\right\} = \min\left\{1 - \frac{2}{\pi} \omega(\beta_*), \beta^*\right\},$$

$$1 - \alpha_* = 1 - \min\left\{1 - \frac{2}{\pi} \omega(\beta^*), \beta_*\right\} = \max\left\{\frac{2}{\pi} \omega(\beta^*), 1 - \beta_*\right\}.$$

Thus, we need to find an upper bound for the ratio

$$\rho(\beta_1, \beta_2) = \frac{\min\{1 - \frac{2}{\pi} \omega(\beta_2), \beta_1\}}{\max\{\frac{2}{\pi} \omega(\beta_1), 1 - \beta_2\}}, \quad 0 \leq \beta_1, \beta_2 \leq 1.$$

Lemma 3.1

$$\max\{\rho(\beta_1, \beta_2) : 0 \leq \beta_1, \beta_2 \leq 1\} = \frac{\bar{\beta}}{1 - \bar{\beta}}.$$

Proof:

We need to prove that

$$(1 - \bar{\beta}) \min\{1 - \frac{2}{\pi}\omega(\beta_2), \beta_1\} \leq \bar{\beta} \max\{\frac{2}{\pi}\omega(\beta_1), 1 - \beta_2\}, \quad 0 \leq \beta_1, \beta_2 \leq 1.$$

This is equivalent to the statement that the convex function

$$g(\beta_1, \beta_2) = \bar{\beta} \max\{\frac{2}{\pi}\omega(\beta_1), 1 - \beta_2\} + (1 - \bar{\beta}) \max\{\frac{2}{\pi}\omega(\beta_2) - 1, -\beta_1\}$$

is non-negative for $0 \leq \beta_1, \beta_2 \leq 1$.

Note that $g(\bar{\beta}, \bar{\beta}) = 0$ in view of the definition of $\bar{\beta}$. The subdifferential of the function $g(\cdot)$ at this point is as follows:

$$\partial g(\bar{\beta}, \bar{\beta}) = \bar{\beta} \text{Conv}\{(\frac{2}{\pi}\omega'(\bar{\beta}), 0); (0, -1)\} + (1 - \bar{\beta}) \text{Conv}\{(0, \frac{2}{\pi}\omega'(\bar{\beta})); (-1, 0)\}.$$

Thus, this set contains the following points:

$$(\frac{2}{\pi}\bar{\beta}\omega'(\bar{\beta}) - 1 + \bar{\beta}, 0), \quad (0, \frac{2}{\pi}(1 - \bar{\beta})\omega'(\bar{\beta}) - \bar{\beta}), \quad (\frac{2}{\pi}\bar{\beta}\omega'(\bar{\beta}), \frac{2}{\pi}(1 - \bar{\beta})\omega'(\bar{\beta})).$$

Note that

$$\frac{2}{\pi}\omega'(\bar{\beta}) < \omega'(\bar{\beta}) < \frac{\bar{\beta}}{1 - \bar{\beta}} < \frac{1 - \bar{\beta}}{\bar{\beta}}.$$

Therefore the first coordinate of the first point and the second coordinate of the second point are negative. Since both coordinates of the third point are positive, we conclude that $0 \in \text{int } \partial g(\bar{\beta}, \bar{\beta})$. $\square$

Applying Lemma 3.1 to (3.6) we prove the statement of Item 2.

In order to prove Item 3 note that in view of inequalities (3.3) and (3.4), for any $\alpha \in [0, 1]$ we have

$$\begin{aligned} \frac{|\psi(\alpha) - \phi^*|}{\phi^* - \phi_*} &\leq \frac{|\psi(\alpha) - \phi^*|}{\phi^* - \psi(\alpha_*)} = \max\left\{\frac{\psi(\alpha) - \phi^*}{\phi^* - \psi(\alpha_*)}, \frac{\phi^* - \psi(\alpha)}{\phi^* - \psi(\alpha_*)}\right\} \\ &\leq \max\left\{\frac{\psi(\alpha) - \psi(\alpha^*)}{\psi(\alpha^*) - \psi(\alpha_*)}, \frac{\psi^* - \psi(\alpha)}{\psi^* - \psi(\alpha_*)}\right\} = \max\left\{\frac{\alpha - \alpha^*}{\alpha^* - \alpha_*}, \frac{1 - \alpha}{1 - \alpha_*}\right\} \equiv r(\alpha). \end{aligned}$$

The minimum $\bar{\alpha}$ of the function $r(\alpha)$ is a solution of the following equation:

$$(\alpha - \alpha^*)(1 - \alpha_*) = (1 - \alpha)(\alpha^* - \alpha_*).$$

That is $\bar{\alpha} = \frac{\alpha^*(2 - \alpha_*) - \alpha_*}{1 + \alpha^* - 2\alpha_*}$. Using Lemma 3.1 we can estimate the optimal value $r(\bar{\alpha})$ as follows:

$$r(\bar{\alpha}) = \frac{1}{1 - \alpha_*} \left(1 - \frac{\alpha^*(2 - \alpha_*) - \alpha_*}{1 + \alpha^* - 2\alpha_*}\right) = \frac{1 - \alpha^*}{1 + \alpha^* - 2\alpha_*} = \frac{\rho(\beta^*, \beta_*)}{2 - \rho(\beta^*, \beta_*)} \leq \frac{\bar{\beta}}{2 - 3\bar{\beta}}.$$

$\square$

4 General constraints on squared variables

Let us consider now the quadratic optimization problems in the following form:

$$\begin{aligned} \text{find } \phi^* &= \max\{\langle Qx, x \rangle : [x]^2 \in \mathcal{F}\}, \\ \text{find } \phi_* &= \min\{\langle Qx, x \rangle : [x]^2 \in \mathcal{F}\}. \end{aligned} \quad (4.1)$$

where $\mathcal{F}$ is a closed convex set, which satisfies Assumption 2.1. Let us show that all results of Sections 2, 3 can be easily applied to the problem (4.1). Denote by $\xi(u)$ the support function of the set $\mathcal{F}$:

$$\xi(u) = \max\{\langle u, v \rangle : v \in \mathcal{F}\}.$$

Theorem 4.1 *The statements of Theorems 2.1, 3.1 are valid for the problem (4.1) with the relaxation values ψ^* , ψ_* and τ^* , τ_* defined as follows:*

$$\begin{aligned} \psi^* &= \min_u \{\xi(u) : D(u) \succeq Q\}, \\ \psi_* &= \max_u \{-\xi(u) : Q + D(u) \succeq 0\}, \\ \tau^* &= \xi(d(Q)), \quad \tau_* = -\xi(-d(Q)). \end{aligned} \quad (4.2)$$

Proof:

In order to prove the theorem we need to rewrite the problem (4.1) in a conic form. Note that in view of Assumption 2.1 the set $\mathcal{F}$ can be represented in the following form:

$$\mathcal{F} = \{v \in S : Bv = d\},$$

where S is a bounded convex set with non-empty interior, B is a non-degenerate $(m \times n)$ -matrix and $d \in R^m$. Without loss of generality we can assume that

$$\{v \in \text{int } S : Bv = d\} \neq \emptyset.$$

We allow also $B = 0$; in this case $d = 0$.

Let us consider a conic hull of the set S :

$$K = \{(v, \tau) : \tau > 0, \frac{1}{\tau}v \in S\} \cup \{0\}.$$

In view of our assumptions K is a closed convex cone with non-empty interior. The cone dual to K can be represented as follows (see, for example, [3]):

$$K^* = \{\hat{u} = (u, \mu) : \mu \geq \xi_S(-u)\},$$

where $\xi_S(\cdot)$ is a support function of the set S .

Now we can rewrite the problem (4.1) in the following form:

$$\begin{aligned} \text{find } \phi^* &= \max\{\langle \hat{Q}\hat{x}, \hat{x} \rangle : [\hat{x}]^2 \in \hat{\mathcal{F}}\}, \\ \text{find } \phi_* &= \min\{\langle \hat{Q}\hat{x}, \hat{x} \rangle : [\hat{x}]^2 \in \hat{\mathcal{F}}\}, \end{aligned} \quad (4.3)$$

where $\hat{x} \in R^{n+1}$, $\hat{Q} = \begin{pmatrix} Q & 0_n \\ 0_n^T & 0 \end{pmatrix}$, $\hat{\mathcal{F}} = \{z = (v, \tau) \in K : \hat{A}z = b\}$ and

$$\hat{A} = \begin{pmatrix} B & -d \\ 0_n^T & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0_m \\ 1 \end{pmatrix}.$$

Note that the problems in (4.3) satisfy Assumptions 2.1, 2.2. Therefore for their relaxation values ψ^* and ψ_* all statements of Theorems 2.1, 3.1 are valid. Let us find the expressions for ψ^* , ψ_* , τ^* and τ_* in terms of the initial objects of the problem (4.1). It is clear that

$$\begin{aligned} \tau^* &= \max\{\langle d(\hat{Q}), z \rangle : z \in \hat{\mathcal{F}}\} \\ &= \max\{\langle d(Q), v \rangle : Bv = \tau d, \tau = 1, v/\tau \in S\} = \xi(d(Q)), \\ \tau_* &= \min\{\langle d(\hat{Q}), z \rangle : z \in \hat{\mathcal{F}}\} \\ &= \min\{\langle d(Q), v \rangle : Bv = \tau d, \tau = 1, v/\tau \in S\} = -\xi(-d(Q)). \end{aligned}$$

Further, in view of Lemma 2.2 the upper relaxation value ψ^* can be represented as follows:

$$\begin{aligned} \psi^* &= \min_{\hat{y} \in R^{m+1}, \hat{u} \in R^{n+1}} \{\langle b, \hat{y} \rangle : \hat{Q} + D(\hat{u}) \preceq D(\hat{A}^T \hat{y}), \hat{u} \in K^*\} \\ &= \min_{(y, \gamma) \in R^{m+1}, (u, \mu) \in R^{n+1}} \{\gamma : Q + D(u) \preceq D(B^T y), \mu \leq \gamma - \langle d, y \rangle, \mu \geq \xi_S(-u)\} \\ &= \min_{u, y} \{\xi_S(-u) + \langle d, y \rangle : D(B^T y - u) \succeq Q\} \\ &= \min_{u, y} \{\xi_S(u - B^T y) + \langle d, y \rangle : D(u) \succeq Q\}. \end{aligned}$$

Note that in the last expression y does not enter the constraints. Therefore we can replace the objective function of this problem by its minimum in y . That is

$$\begin{aligned} \min_y \{\xi_S(u - B^T y) + \langle d, y \rangle\} &= \min_y \max_{v \in S} \{\langle u - B^T y, v \rangle + \langle d, y \rangle\} \\ &= \max_{v \in S} \min_y \{\langle u, v \rangle + \langle d - Bv, y \rangle\} \\ &= \max_{v \in S} \{\langle u, v \rangle : Bv = d\} = \xi(u). \end{aligned}$$

Thus, we get the representation (4.2) for ψ^* . The representation of ψ_* can be obtained in a similar way:

$$\begin{aligned}
\psi^* &= \max_{\hat{y} \in R^{m+1}, \hat{u} \in R^{n+1}} \{ \langle b, \hat{y} \rangle : \hat{Q} \succeq D(\hat{u}) + D(\hat{A}^T \hat{y}), \hat{u} \in K^* \} \\
&= \max_{(y, \gamma) \in R^{m+1}, (u, \mu) \in R^{n+1}} \{ \gamma : Q \succeq D(u) + D(B^T y), 0 \geq \mu + \gamma - \langle d, y \rangle, \mu \geq \xi_S(-u) \} \\
&= \max_{u, y} \{ -\xi_S(-u) + \langle d, y \rangle : Q \succeq D(u + B^T y) \} \\
&= \max_{u, y} \{ -\xi_S(B^T y + u) + \langle d, y \rangle : Q + D(u) \succeq 0 \} \\
&= \max_u \{ -\min_y \{ \xi_S(B^T y + u) - \langle d, y \rangle \} : Q + D(u) \succeq 0 \} \\
&= \max_u \{ -\xi(u) : Q + D(u) \succeq 0 \}.
\end{aligned}$$

□

Let us present an example of application of Theorems 4.1, 3.1. Consider the following problem:

$$\begin{aligned}
\text{find } \phi^* &= \max \{ \langle Qx_1, x_2 \rangle : [(x_1, x_2)]^2 \in \mathcal{F} \}, \\
\text{find } \phi_* &= \min \{ \langle Qx_1, x_2 \rangle : [(x_1, x_2)]^2 \in \mathcal{F} \},
\end{aligned} \tag{4.4}$$

where Q is a $(k \times n)$ -matrix, $x_1 \in R^k$, $x_2 \in R^n$ and $\mathcal{F}$ is a closed convex set, which satisfies Assumption 2.1. Since the quadratic objective function in this problem is bilinear, we conclude that $\phi_* = -\phi^*$ and $\tau^* = \tau_* = 0$.

The conic relaxation for this problem is defined as follows:

$$\begin{aligned}
\psi^* &= \min_{u=(u_1, u_2)} \left\{ \xi(u) : \begin{pmatrix} D(u_1) & -Q^T \\ -Q & D(u_2) \end{pmatrix} \succeq 0 \right\}, \\
\psi_* &= \max_{u=(u_1, u_2)} \left\{ -\xi(u) : \begin{pmatrix} D(u_1) & Q^T \\ Q & D(u_2) \end{pmatrix} \succeq 0 \right\}.
\end{aligned}$$

It is clear that $\psi_* = -\psi^*$. At the same time, $\beta^* = \beta_* = \frac{1}{2}$. Therefore,

$$\alpha^* = \max \left\{ \frac{2}{\pi} \omega(\beta_*), 1 - \beta^* \right\} = \frac{2}{\pi} \omega\left(\frac{1}{2}\right).$$

Therefore, in view of Theorem 3.1 we have:

$$\psi^* \geq \phi^* \geq \psi(\alpha^*) = (2\alpha^* - 1)\psi^*.$$

Note that $\alpha^* = \frac{2}{\pi} \left(\frac{1}{2} \arcsin \frac{1}{2} + \frac{\sqrt{3}}{2} \right) = \frac{\sqrt{3}}{\pi} + \frac{1}{6}$. Thus, we have proved the following theorem.

Theorem 4.2 *In the problem (4.4) the optimal and relaxation values are related as follows:*

$$\psi^* \geq \phi^* \geq \gamma \psi^*$$

with $\gamma = \frac{2\sqrt{3}}{\pi} - \frac{2}{3} > 0.43$.

5 Why are linear constraints difficult?

In previous sections we have derived constant relative accuracy estimates for a quadratic maximization problem with convex constraints on *squared* variables. Such type of constraints are rather specific. Therefore it is natural to try to extend the results onto the problems with convex constraints on the variables of the quadratic form. However, it appears that this is not trivial. In this section we show that even a single linear constraint can make a quadratic problem completely intractable by the presented technique.

Consider the following optimization problem:

$$\begin{aligned} \phi^* = \max \quad & \langle Qx, x \rangle, \\ \text{s.t.} \quad & x \in \{-1, 1\}^n, \\ & \langle c, x \rangle = \beta, \end{aligned} \tag{5.1}$$

where Q is an $(n \times n)$ -matrix, $c \in R^n$ and $\beta > 0$. Define ϕ_* as a minimal value of the objective function in (5.1). A natural relaxation for this problem is as follows:

$$\psi^* = \max\{\langle Q, X \rangle : \langle Xc, c \rangle = \beta^2, d(X) = 1_n, X \succeq 0\}. \tag{5.2}$$

Let us show that this relaxation can be arbitrary bad in terms of relative accuracy.

Denote by $v_i, i = 1, \dots, 2^n$ the nodes of the boolean unit box $\{-1, 1\}^n$. Let us assume that there exists only one node v_* , which satisfies the linear constraint of the problem (5.1). Moreover, let us assume that there are two other nodes, v_+ and v_- such that

$$0 < \langle c, v_- \rangle < \beta < \langle c, v_+ \rangle. \tag{5.3}$$

Note that in view of our assumption we have $\phi^* = \phi_*$ independently on our choice of the matrix Q .

Let us define a convex polytope $\mathcal{P}_n$ of positive semidefinite $(n \times n)$ -matrices:

$$\mathcal{P}_n = \text{Conv}\{V_i = v_i v_i^T, i = 1, \dots, 2^n\}.$$

Lemma 5.1 *Any V_i is an extreme point of $\mathcal{P}_n$. Any pair of nodes V_i, V_j is connected by an edge.*

Proof:

Since V_i is a rank-one matrix, the first statement is evident. In order to prove the second statement note that the edge $[V_i, V_j]$ is not exposed if and only if there exist some coefficients $\lambda_k > 0, k \in \mathcal{I}, i, j \notin \mathcal{I}$ such that

$$\alpha V_i + (1 - \alpha) V_j = \sum_{k \in \mathcal{I}} \lambda_k V_k, \quad \sum_{k \in \mathcal{I}} \lambda_k = 1,$$

for some $\alpha \in (0, 1)$. Since all nodes of $\mathcal{P}_n$ are positive semidefinite rank-one matrices, we conclude that

$$v_k \in \{v : v = \alpha v_i + \beta v_j, (\alpha, \beta) \in R^2\}, \quad \forall k \in \mathcal{I}.$$

A simple calculation shows that it is possible only for $v_k = \pm v_i$ or $v_k = \pm v_j$. □

Note that in view of our assumption (5.3) there exists a matrix $\tilde{V} \in \mathcal{P}_n$ such that

$$\tilde{V} = \alpha v_- v_-^T + (1 - \alpha) v_+ v_+^T, \quad \alpha \in (0, 1), \quad \langle \tilde{V} c, c \rangle = \beta^2.$$

Let us choose now $Q = \tilde{V} - v_* v_*^T$. Note that the feasible set of the relaxation problem (5.2) contains $\mathcal{P}_n$. Therefore

$$\psi^* \geq \langle Q, \tilde{V} \rangle > \langle Q, v_* v_*^T \rangle = \phi^*.$$

The lower relaxation value ψ_* never exceed $\phi_* = \phi^*$. Therefore, for our example the value $\psi^* - \psi_*$ is strictly positive. This means that the relative accuracy of the value ψ^* is infinitely bad.

Note that the main source of our troubles in the above example is that the linear constraint $\langle Xc, c \rangle = \beta^2$ intersects an edge of the matrix polytope $\mathcal{P}_n$. That can happens with any value of β except $\beta = 0$. Thus, we still can hope that for the problems with homogeneous linear constraints the conic relaxation can work. In the next sections we will see some problems for which it is true.

6 Maximization with a smooth constraint

In previous sections we have established some constant bounds on relative accuracy of the conic relaxations (2.4) for a quadratic maximization problem with convex constraints for the squared variables. At the same time, in Section 5 we have seen that some linear constraints on the initial variables can make the problem intractable in terms of relative accuracy. In this section we present another approach for deriving the conic relaxations. This approach is based on the standard second order optimality conditions and it allows to treat the quadratic maximization problems over l_p -boxes, $p \geq 2$, with homogeneous linear equality constraints (see Section 7). However, the quality of relaxation in this framework becomes dependent on p .

Let $f(y)$, $y \in R^m$, be a homogeneous function of degree p :

$$f(\tau y) = \tau^p f(y), \quad y \in R^m, \quad \tau \geq 0. \quad (6.1)$$

We assume that $f(y)$ is non-negative and twice continuously differentiable at any non-zero point of R^m (notation $f \in H_p$). Recall, that for homogeneous functions we have the following simple relations.

Lemma 6.1 *If $f(y)$ is homogeneous of degree p then for any $y \in R^m$ and $\tau \geq 0$ we have*

$$f'(\tau y) = \tau^{p-1} f'(y), \quad (6.2)$$

$$f''(y)y = (p-1)f'(y), \quad (6.3)$$

$$\langle f'(y), y \rangle = pf(y), \quad (6.4)$$

$$\langle f''(y)y, y \rangle = p(p-1)f(y). \quad (6.5)$$

Proof:

Indeed, if we differentiate (6.1) in y we get (6.2). If we differentiate (6.2) in τ and take $\tau = 1$ we get (6.3). In order to get (6.4) we differentiate (6.1) in τ and take $\tau = 1$. Finally, (6.3) and (6.4) give (6.5). $\square$

Let Q be a symmetric $(m \times m)$ -matrix. Consider the following maximization problem:

$$\text{find } \phi^*(Q) = \max\{\langle Qy, y \rangle : f(y) \leq 1\}. \quad (6.6)$$

If $Q \preceq 0$ then (6.6) is a concave maximization problem and $\phi^*(Q) = 0$. In the other cases we need some necessary conditions to characterize the local solutions of the problem (6.6).

Lemma 6.2 *Let $f \in H_p$ with $p > 0$. Then for any local maximum y_* of the problem (6.6) with $\langle Qy_*, y_* \rangle > 0$ we have $f(y_*) = 1$. Moreover, there exists a value $\lambda = \lambda(y_*) > 0$ such that*

$$\langle Qy_*, y_* \rangle = p\lambda, \quad (6.7)$$

$$Qy_* = \lambda f'(y_*), \quad (6.8)$$

$$Q \preceq \lambda \left[f''(y_*) - \frac{p-2}{p} f'(y_*) f'(y_*)^T \right]. \quad (6.9)$$

Proof:

Since $\langle Qy_*, y_* \rangle > 0$ and $f(y)$ is a homogeneous function of positive degree, we necessarily have $f(y_*) = 1$. Let us write down a Lagrangean for this problem:

$$\mathcal{L}(y, \lambda) = \frac{1}{2} \langle Qy, y \rangle - \lambda [f(y) - 1].$$

Then, the second order necessary conditions for the problem (6.6) can be written as follows:

$$\mathcal{L}'_y(y_*, \lambda) = 0, \quad (6.10)$$

$$\langle \mathcal{L}''_{yy}(y_*, \lambda)h, h \rangle \preceq 0, \quad \forall h : \langle f'(y_*), h \rangle = 0, \quad (6.11)$$

with some $\lambda \in R$. Equation (6.10) is exactly (6.8). Multiplying (6.8) by y_* and using (6.4) we get

$$\langle Qy_*, y_* \rangle = \lambda \langle f'(y_*), y_* \rangle = \lambda p f(y_*) \lambda p > 0,$$

and that is (6.7). Finally, since $\langle f'(y_*), y_* \rangle = p > 0$, any $h \in R^m$ such that $\langle f'(y_*), h \rangle = 0$ can be represented in the form

$$h = \left(I - \frac{1}{p} y_* f'(y_*)^T \right) u, \quad u \in R^n.$$

Therefore the condition (6.11) can be rewritten as

$$\left(I - \frac{1}{p} f'(y_*) y_*^T \right) \mathcal{L}''_{yy}(y_*, \lambda) \left(I - \frac{1}{p} y_* f'(y_*)^T \right) \preceq 0. \quad (6.12)$$

Note that $\mathcal{L}''_{yy}(y_*, \lambda) = Q - \lambda f''(y_*)$ and

$$\begin{aligned} & \left(I - \frac{1}{p} f'(y_*) y_*^T \right) Q \left(I - \frac{1}{p} y_* f'(y_*)^T \right) \\ &= Q - \frac{1}{p} f'(y_*) y_*^T Q - \frac{1}{p} Q y_* f'(y_*)^T + \frac{1}{p^2} \langle Qy_*, y_* \rangle f'(y_*) f'(y_*)^T \\ &= Q - \frac{\lambda}{p} f'(y_*) f'(y_*)^T \end{aligned}$$

in view of (6.8) and (6.7). Similarly, since $f(y_*) = 1$ we have

$$\begin{aligned}
& (I - \frac{1}{p}f'(y_*)y_*^T)f''(y_*)(I - \frac{1}{p}y_*f'(y_*)^T) \\
&= f''(y_*) - \frac{1}{p}f'(y_*)y_*^Tf''(y_*) - \frac{1}{p}f''(y_*)y_*f'(y_*)^T + \frac{1}{p^2}\langle f''(y_*)y_*, y_* \rangle f'(y_*)f'(y_*)^T \\
&= f''(y_*) - 2\frac{p-1}{p}f'(y_*)f'(y_*)^T + \frac{p(p-1)}{p^2}f'(y_*)f'(y_*)^T = f''(y_*) - \frac{p-1}{p}f'(y_*)f'(y_*)^T,
\end{aligned}$$

in view of (6.3) and (6.5). Substituting these expressions in (6.12) we get

$$Q \preceq \lambda f''(y_*) + \frac{\lambda}{p}f'(y_*)f'(y_*)^T - \lambda \frac{p-1}{p}f'(y_*)f'(y_*)^T = \lambda \left[f''(y_*) - \frac{p-2}{p}f'(y_*)f'(y_*)^T \right].$$

□

We will use Lemma 6.2 in order to estimate the quality of relaxations for some non-convex maximization problems. Let $A = (a_1, \dots, a_n) \in R^{m \times n}$ be a non-degenerate $(m \times n)$ -matrix. Consider the following function:

$$f_A(y) = \sum_{i=1}^n |\langle a_i, y \rangle|^p,$$

where $p \geq 2$. The problem we are going to address now is as follows:

$$\text{find } \phi^*(Q, A) = \max\{\langle Qy, y \rangle : f_A(y) \leq 1\}. \quad (6.13)$$

For this problem we can introduce the following relaxation:

$$\psi_p^*(Q, A) = \min_u \{\|u\|_q : AD(u)A^T \succeq Q\}, \quad (6.14)$$

where $q = (\frac{p}{2})^* = \frac{p}{p-2}$ (compare with (4.2)). Now we can prove the main result of this section.

Theorem 6.1 *Let the feasible set of the problem (6.13) be bounded. Then*

$$\frac{1}{p-1}\psi_p^*(Q, A) \leq \phi^*(Q, A) \leq \psi_p^*(Q, A). \quad (6.15)$$

Moreover, any local maximum y_ of the problem (6.13) with positive value of the objective function satisfies inequality $\langle Qy_*, y_* \rangle \geq \frac{1}{p-1}\psi_p^*(Q, A)$.*

Proof:

Indeed, let u be feasible for the problem (6.14). Then for any $y \in R^m$ with $f_A(y) \leq 1$ we have

$$\langle Qy, y \rangle \leq \langle AD(u)A^T y, y \rangle = \langle u, [A^T y]^2 \rangle.$$

At the same time,

$$\| [A^T y]^2 \|_{p/2}^{p/2} = \sum_{i=1}^n |\langle a_i, y \rangle|^p = f_A(y) \leq 1.$$

Therefore, for any feasible y we have

$$\langle Qy, y \rangle \leq \langle u, [A^T y]^2 \rangle \leq \|u\|_q \cdot \| [A^T y]^2 \|_{p/2} \leq \|u\|_q.$$

Hence, $\phi^*(Q, A) \leq \psi_p^*(Q, A)$.

On the other hand, let y_* be a local maximum of (6.13) with $\langle Qy_*, y_* \rangle > 0$. Then, in view of Lemma 6.2 (6.9) for $\lambda = \lambda(y_*)$ we have:

$$Q \preceq \lambda f''(y_*) = p(p-1)\lambda \sum_{i=1}^n |\langle a_i, y_* \rangle|^{p-2} a_i a_i^T$$

(we have used the condition $p \geq 2$). Thus, the vector $u \in R^n$ with the components

$$u^{(i)} = p(p-1)\lambda |\langle a_i, y_* \rangle|^{p-2}, \quad i = 1, \dots, n,$$

is feasible for the problem (6.14). Note that

$$\begin{aligned} \|u\|_q &= p(p-1)\lambda \left[\sum_{i=1}^n |\langle a_i, y_* \rangle|^{(p-2)q} \right]^{1/q} \\ &= p(p-1)\lambda \left[\sum_{i=1}^n |\langle a_i, y_* \rangle|^p \right]^{1/q} \\ &= p(p-1)\lambda [f_A(y_*)]^{1/q} = p(p-1)\lambda. \end{aligned}$$

Hence, in view of (6.7) we have

$$\langle Qy_*, y_* \rangle = p\lambda = \frac{\|u\|_q}{p-1} \geq \frac{1}{p-1} \psi_p^*(Q, A).$$

Note that the above proof shows that under assumptions of the theorem the function $\psi_p^*(Q, A)$ is well defined.

Finally, if there is no local maximum of the problem (6.13) with $\langle Qy_*, y_* \rangle > 0$, then $Q \preceq 0$ and in this case we have $\psi_p^*(Q, A) = \phi^*(Q, A) = 0$. $\square$

Let us estimate the relative accuracy of the relaxation (6.14). First, we need the following trivial result.

Lemma 6.3 *Let for some non-negative values ϕ , ψ and γ we have the following relations:*

$$\gamma\psi \leq \phi \leq \psi.$$

Then, for $\beta = \frac{2\gamma}{1+\gamma}$ we have: $|\beta\psi - \phi| \leq (1-\beta)\phi$.

Define $\phi_*(Q, A) = \min\{\langle Qy, y \rangle : f_A(y) \leq 1\}$.

Theorem 6.2 *Let $\psi^* = \frac{2}{p}\psi_p^*(Q, A)$. Then*

$$|\phi^*(Q, A) - \psi^*| \leq (1 - \frac{2}{p})(\phi^*(Q, A) - \phi_*(Q, A)). \quad (6.16)$$

Proof:

Note that $\phi_*(Q, A) \leq 0$. Therefore it is sufficient to prove

$$|\phi^*(Q, A) - \psi^*| \leq (1 - \frac{2}{p})\phi^*(Q, A).$$

Let us choose $\gamma = \frac{1}{p-1}$ and $\beta = \frac{2\gamma}{1+\gamma} = \frac{2}{p}$. Then the above inequality follows from Theorem 6.1 and Lemma 6.3. $\square$

Let us compare now the relaxation (6.14) with the conic relaxation (4.2). Of course, we have to choose a problem which can be treated by both approaches. Consider the problem

$$\max\{\langle Qx, x \rangle : \|x\|_p \leq 1\}, \quad p \geq 2.$$

This problem can be presented in the form (6.13) with $A = I_n$. On the other hand, it can be written in the form (4.1) with

$$\mathcal{F} = \{v : \|v\|_{p/2} \leq 1\}.$$

In this case $\xi(u) = \|u\|_q$ and we can see that (6.14) coincides with (4.2).

7 Some applications

Let us show that the results of the previous section can be extended onto the problems with linear equality constraints. Consider the following quadratic maximization problem:

$$\begin{aligned} \text{find } \phi_p^* &= \max_{x \in R^n} \langle Cx, x \rangle, \\ \text{s.t. } &\|x\|_p \leq 1, \end{aligned} \tag{7.1}$$

$$Bx = 0,$$

where C is an arbitrary $(n \times n)$ -matrix, $p \geq 2$ and B is a non-degenerate $((n - m) \times n)$ -matrix with $n > m$. Let the rows of some $(m \times n)$ -matrix A span the null space of the matrix B :

$$BA^T y = 0, \quad \forall y \in R^m.$$

Then we can change variables $x = A^T y$ and obtain a problem, which is equivalent to (7.1):

$$\phi_p^* = \max_{y \in R^m} \{\langle ACA^T y, y \rangle : f_A(y) \leq 1\} = \phi^*(ACA^T, A).$$

Thus, in view of Theorem 6.1 and Lemma 6.3 we get the following result.

Theorem 7.1 *For any $p \geq 2$ we have*

$$\frac{1}{p-1} \psi_p^*(ACA^T, A) \leq \phi_p^* \leq \psi_p^*(ACA^T, A).$$

The value $\psi^ = \frac{2}{p} \psi_p^*(ACA^T, A)$ approximates the solution of the problem (7.1) with $(1 - \frac{2}{p})$ relative accuracy.*

Now, let us consider the case when the objective function of the problem (7.1) has a non-zero linear term:

$$\begin{aligned}
\text{find } \hat{\phi}_p^* &= \max_{x \in R^n} \quad \langle Cx, x \rangle + 2\langle c, x \rangle, \\
\text{s.t. } &\|x\|_p \leq 1, \\
&Bx = 0.
\end{aligned} \tag{7.2}$$

This problem can be homogenized in a standard way:

$$\begin{aligned}
&\max_{(x, \tau) \in R^{n+1}} \quad \langle Cx, x \rangle + 2\tau \langle c, x \rangle, \\
\text{s.t. } &\|x\|_p \leq 1, \quad |\tau| \leq 1, \\
&Bx = 0.
\end{aligned} \tag{7.3}$$

Clearly, the optimal value of this problem is $\hat{\phi}_p^*$. However, this problem has two separate constraints for x and τ . Therefore, in order to apply the results of Section 6 we need to replace them by a single functional inequality. Consider the following problem:

$$\begin{aligned}
\text{find } \bar{\phi}_p^* &= \max_{(x, \tau) \in R^{n+1}} \quad \langle Cx, x \rangle + 2\tau \langle c, x \rangle, \\
\text{s.t. } &\|(x, \tau)\|_p \leq 1, \\
&Bx = 0,
\end{aligned} \tag{7.4}$$

Denote by $\bar{\psi}_p^*$ the value of the conic relaxation for the last problem.

Theorem 7.2 *Let $p \geq 2$. For $\psi_a^* = 2^{2/p} \bar{\psi}_p^*$ we have:*

$$\frac{1}{2^{2/p}(p-1)} \psi_a^* \leq \hat{\phi}_p^* \leq \psi_a^*.$$

The value $\psi_r^ = \frac{2}{p+2^{-2/p}-1} \bar{\psi}_p^*$ has at least $(1 - \frac{1}{2p})$ relative accuracy.*

Proof:

Note that the problems (7.3) and (7.4) have the same objective function and the same system of linear equations. Denote by $\mathcal{F}_0$ the feasible set of the problem (7.4) and by $\mathcal{F}_1$ the feasible set of the problem (7.3). Clearly, $\mathcal{F}_0 \subset \mathcal{F}_1 \subset 2^{1/p} \mathcal{F}_0$. Therefore

$$\bar{\phi}_p^* \leq \hat{\phi}_p^* \leq 2^{2/p} \bar{\phi}_p^*.$$

On the other hand, in view of Theorem 7.1, we have:

$$\frac{1}{p-1} \bar{\psi}_p^* \leq \bar{\phi}_p^* \leq \bar{\psi}_p^*.$$

Hence, for $\psi_a^* = 2^{2/p} \bar{\psi}_p^*$ we obtain:

$$\psi_a^* = 2^{2/p} \bar{\psi}_p^* \geq 2^{2/p} \bar{\phi}_p^* \geq \hat{\phi}_p^* \geq \bar{\phi}_p^* \geq \frac{1}{p-1} \bar{\psi}_p^* = \frac{1}{2^{2/p}(p-1)} \psi_a^*.$$

In order to get the statement on the relative accuracy, we take $\psi = \psi_a^*$, $\phi = \hat{\phi}_p^*$, $\gamma = \frac{1}{2^{2/p}(p-1)}$ and apply Lemma 6.3. Then the values β and ψ_r^* can be obtained as follow:

$$\beta = \frac{2\gamma}{1+\gamma} = \frac{2}{1+2^{2/p}(p-1)} \geq \frac{1}{2p},$$

$$\psi_r^* = \beta \psi_a^* = \frac{2}{p+2^{-2/p}-1} \bar{\psi}_p^*.$$

□

We see that the quality of conic relaxation decreases as p increase. Therefore, we cannot directly apply the results of Section 6 to a problem with box constraints. However, at the same time, when p increase the shape of l_p balls becomes very close to the shape of the n -dimensional unit box. Therefore, we can use the values $\psi_p^*(ACA^T, A)$ with p large enough in order to get some bounds for ϕ_∞^* .

Theorem 7.3 *Let $p = 2 \ln n$, $\psi_a^* = e\psi_p^*(ACA^T, A)$ and $\gamma = \frac{1}{e(2 \ln n - 1)}$. Then*

$$\gamma \psi_a^* \leq \psi_\infty^* \leq \psi_a^*.$$

The value $\psi_r^ = \frac{2\gamma}{1+\gamma} \psi_a^*$ has at least $(1 - \frac{1}{e \ln n})$ relative accuracy.*

Proof:

It is well known that for any two values $p \geq 2$ we have:

$$\frac{1}{n^{1/p}} \|x\|_p \leq \|x\|_\infty \leq \|x\|_p, \quad x \in R^n.$$

Therefore

$$\{x \in R^n : \|x\|_p \leq 1\} \subset \{x \in R^n : \|x\|_\infty \leq 1\} \subset \{s \in R^n : \|x\|_p \leq n^{1/p}\}.$$

Since the objective function of the problem (7.1) is homogeneous of degree two, this implies that $\phi_p^* \leq \phi_\infty^* \leq n^{2/p} \phi_p^*$. Thus, using Theorem 7.1 we obtain the following:

$$\begin{aligned} \psi_\infty^* &= e\psi_p^*(ACA^T, A) = n^{2/p} \psi_p^*(ACA^T, A) \geq n^{2/p} \phi_p^* \\ &\geq \phi_\infty^* \geq \phi_p^* \geq \frac{1}{p-1} \psi_p^*(ACA^T, A) = \frac{1}{e(2 \ln n - 1)} \psi_a^*. \end{aligned}$$

In order to get the statement on relative accuracy we apply Lemma 6.3 with

$$\beta = \frac{2\gamma}{1+\gamma} = \frac{2}{1+e(2 \ln n - 1)} > \frac{1}{e \ln n}.$$

□

8 Discussion

In the previous sections we have presented some estimates for the quality of the conic relaxation for different non-convex quadratic maximization problems. The constant bounds of Sections 2, 3 can be applied to a quite large class of non-convex problems and we can expect that they can be used in many practical applications. The bounds we get in Section 6 are not so good. Indeed, they can be applied only to a rather special feasible set, that is an intersection of an l_p -ball, $p \geq 2$, with a linear subspace. Moreover, the quality of these bounds decrease as p increase.

Nevertheless, the results of Section 6 suggest some interesting conclusions. Firstly, the relative accuracy we get from the relaxation (6.14) is $(1 - \frac{2}{p})$. Thus, the accuracy goes to zero as p approaches two. For p small enough the results of Theorem 6.1 become even better than the bounds of Section 2. An important advantage of the estimates (6.15) is that we get the separate bounds for the minimal and the maximal value of the problem. The lower estimate for the maximal value remains positive even if the minimal value of the problems is a large negative value.

Secondly, Theorem 6.1 tells us that the value of the objective function of the problem (6.13) at *any* local solution is not worse than the lower bound we get from the conic relaxation. In fact, this statement is a kind of surprise. Indeed, if we measure a hardness of a problem as a largest ratio of the values of the objective function at the global and a local maximum, it appears that the problem (6.13) is not so difficult, at least for p small enough. Usually the general methods of nonlinear optimization are quite efficient in finding a local solution. Since the computational cost of such schemes is much less than that of the schemes of semidefinite programming, we can conclude that for practical applications the traditional schemes look quite attractive.¹

Finally, in Section 7 we have shown that the results of Theorem 6.1 provides us with some bounds for very difficult problems. Indeed, during last years there were obtained many negative results related to the possibilities to find an approximate solution of an NP -hard problem under hypothesis that $P \neq NP$. The results relevant to the topic of our paper can be found in [1]:

Consider a quadratic optimization problem in the following form:

$$\max\{\langle Cx, x \rangle : Bx \leq b, 0 \leq x \leq 1_n\}. \quad (8.1)$$

Denote by $\tilde{P}$ the class of languages recognizable in quasi-polynomial time.

Theorem 1.2. *Assume $NP \not\subseteq \tilde{P}$. Then there is a constant $\delta > 0$ such that the problem (8.1) has no polynomial time, $(1 - 2^{-\log^\delta n})$ -approximation algorithm.*

Theorem 1.3. *Assume $P \neq NP$. Then there is a constant $\mu \in (0, \frac{1}{3})$ such that a μ -approximation of the problem (8.1) cannot be found in polynomial time.*

In these statements the μ -approximation is understood in a weak sense. We need to compute an estimate for the value of the objective function only.

¹Of course, in non-convex case we cannot prove any global efficiency estimates. Moreover, in general we cannot guarantee a convergence to a point, which satisfies the necessary second order optimality conditions. This negative result is valid even for the second order methods.

Note that using Theorems 7.2 and 7.3, we can approximate in polynomial time the optimal value of the problem

$$\max\{\langle Cx, x \rangle : Bx = \frac{1}{2}1_n, 0 \leq x \leq 1_n\}. \quad (8.2)$$

with $(1 - O(\frac{1}{\ln n}))$ relative accuracy. This result is better than the limiting bound of Theorem 1.2 [1]. At the same time, the optimization problem, which is used in the proof of Theorems 1.2, 1.3 [1], has, in fact, only linear equalities constraint:

$$\max\{\langle Cx, x \rangle : Bx = b, 0 \leq x \leq 1_n\}. \quad (8.3)$$

Thus, the difference in the formulations (8.2) and (8.3) looks very minor. Indeed, any system of linear equations $Bx = b$ can be rewritten in the following form:

$$\bar{B}x = \frac{1}{2}1_n, \quad \langle a, x \rangle = 1,$$

with some matrix $\bar{B}$ and a vector $a \in R^n$. Hence, the feasible set of the problem (8.3) differs from the feasible set of the problem (8.2) just by a single linear equation, which does not pass through the center of the box. However, it appears that this linear equation makes the problem (8.3) completely different.

Let us look at the concrete form of the problem (8.3) ([1], p.438). Denote by X and Y two $(n \times n)$ -matrices. And let $\phi(X, Y)$ be a bilinear form in X and Y with all non-negative coefficients. Then the problem (8.3) is as follows:

$$\begin{aligned} & \max \phi(X, Y), \\ & \text{s.t. } X1_n = 1_n, Y1_n = 1_n, \\ & 0 \leq X, Y \leq 1_{n \times n}. \end{aligned} \quad (8.4)$$

Now we can see the source of our troubles. Indeed, the technique of Section 6 can be applied only to l_p boxes with $p \geq 2$. However, if we will try to approximate the feasible set of the problem (8.4) with the boxes $\mathcal{B}_p = \{x : \|x - \frac{1}{2}1_n\|_p \leq \frac{1}{2}\}$, we need to choose p very large. It is necessary to take $p = O(n \ln n)$ just to have a non-empty intersection of the box $\mathcal{B}_p$ with the system of linear constraints in (8.4).

Thus, we conclude that the feasible set of the problem (8.4) is too far from the center of the box. On the other hand, it is clear that the box structure in (8.4) is quite artificial: the constraint $X, Y \leq 1_{n \times n}$ can be eliminated without changing the feasible set of the problem. Note that we can easily rewrite the problem (8.4) in a more symmetric form:

$$\begin{aligned} & \max \phi(X, Y), \\ & \text{s.t. } \|Xe_i\|_1 \leq 1, i = 1, \dots, n, \\ & \|Ye_i\|_1 \leq 1, i = 1, \dots, n. \end{aligned} \quad (8.5)$$

Since the coefficients of the form $\phi(X, Y)$ are non-negative, the optimal value of the problem (8.5) is the same as that of (8.1). The polyhedral structure of the feasible set in

(8.5) can be seen as a combination of l_∞ -structure with l_1 -structure. However, it appears the latter structure is exactly that one, for which no reasonable bounds for quadratic problems are known.

Thus, the above discussion highlights the following unsolved problem:

Find some bounds for the optimal value of the following quadratic problem:

$$\phi^* = \max\{\langle Qx, x \rangle : \|x\|_p \leq 1, x \in R^n\}, \quad 1 \leq p < 2. \quad (8.6)$$

For an indefinite Q a trivial bound for ϕ^* is given by its maximal eigenvalue $\lambda_{\max}(Q)$:

$$\lambda_{\max}(Q) \geq \phi^* \geq \lambda_{\max}(Q) \cdot n^{1-\frac{2}{p}}, \quad 1 \leq p \leq 2.$$

For $p = 1$ we can suggest for the problem (8.6) a kind of semidefinite relaxation:

$$\begin{aligned} \psi^* &= \max_{X, u} \{\langle Q, X \rangle : D(u) \succeq X, \langle 1_n, u \rangle \leq 1, X \succeq 0\} \\ &= \min_{S, \lambda} \{\lambda : \lambda 1_n = d(S), S \succeq Q, S \succeq 0\}. \end{aligned} \quad (8.7)$$

Note that for any x , $\|x\|_1 \leq 1$, the pair $(X = xx^T, u = \text{abs}[x])$ is feasible for the primal form of the relaxation (8.7). Therefore we can guarantee that $\psi^* \geq \phi^*$. However, the relative accuracy of such a bound is not known.

References

- [1] M.Bellare and P.Rogaway, "The complexity of approximating a nonlinear program," *Mathematical Programming*, 69(1995) 429–441.
- [2] M.X. Goemans and D.P. Williamson, "Improved Approximation Algorithms for maximum cut and satisfiability problem using semi-definite programming," *Journal of ACM*, 42 (1996) 1115–1145.
- [3] J.-B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag. 1993.
- [4] Yu.Nesterov, "Quality of semidefinite relaxation for nonconvex quadratic optimization," CORE Discussion Paper #9719, CORE 1997.
- [5] Yu.Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optimization Methods and Software*, 12 (1997) 1–20.
- [6] Yu. Nesterov and A. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
- [7] Y.Ye, "Approximating quadratic programming with bound constraints," Working paper, Department of Management Sciences, University of Iowa, Iowa City, 1997.