Approximating quadratic programming with bound constraints

Yinyu Ye*
Department of Management Sciences
The University of Iowa
Iowa City, Iowa 52242, U.S.A.

March 31, 1997

Abstract

We consider the problem of approximating the global maximum of a quadratic program (QP) with n variables subject to bound constraints. Based on the results of Goemans and Williamson [4] and Nesterov [6], we show that a 4/7 approximate solution can be obtained in polynomial time.

Key words. Quadratic programming, global maximizer, approximation algorithm

*This author is supported in part by NSF grant DMI-9522507.

1 Introduction

Consider the quadratic programming (QP) problem

$$\text{(QP)} \qquad \begin{array}{rll} \bar{q}(Q) := & \text{Maximize} & q(x) := x^T Qx \\ & \text{Subject to} & -e \le x \le e, \end{array}$$

where $Q \in \Re^{n\times n}$ is given and $e \in \Re^n$ is the vector of all ones. Let $\bar{x} = \bar{x}(Q)$ be a maximizer of the problem. In this paper, without loss of generality, we assume that $\bar{x} \neq 0$.

Normally, there is a linear term in the objective function: $q(x) = x^T Qx + c^T x$. However, the problem can be homogenized as

$$\begin{array}{ll} \text{Maximize} & q(x) := x^T Qx + tc^T x \\ \text{Subject to} & -e \le x \le e, \ -1 \le t \le 1 \end{array}$$

by adding a scalar variable t. There always is an optimal solution $(\bar{x}, \bar{t})$ for this problem in which $\bar{t} = 1$ or $\bar{t} = -1$. If $\bar{t} = 1$, then $\bar{x}$ is also optimal for the non-homogeneous problem; if $\bar{t} = -1$, then $-\bar{x}$ is optimal for the non-homogeneous problem. Thus, without loss of generality, we can let $q(x) = x^T Qx$ throughout this paper.

The function $q(x)$ has a minimizer and a maximizer over the bounded feasible set $-e \le x \le e$. Let $\underline{q} := -\bar{q}(-Q)$ and $\bar{q} := \bar{q}(Q)$ denote their minimal and maximal objective values, respectively. An ϵ-maximal solution or ϵ-maximizer, $\epsilon \in [0, 1]$, for (QP) is defined as an $-e \le x \le e$ such that

$$\frac{\bar{q} - q(x)}{\bar{q} - \underline{q}} \le \epsilon.$$

Note that according to this definition any feasible solution x is a 1-maximizer.

Recently, there were several significant results on approximating specific quadratic problems. Goemans and Williamson [4] proved an approximation result for the Maxcut problem where $\epsilon \le 1 - 0.878$. Nesterov [6] generalized their result to approximating a boolean QP problem

$$\begin{array}{ll} \text{Maximize} & q(x) = x^T Qx \\ \text{Subject to} & |x_j| = 1, \ j = 1, \ldots, n. \end{array}$$

where $\epsilon \le 4/7$. Some negative results were given by Bellare and Rogaway [1].

There are also several approximation algorithms developed for approximating (QP) when the feasible set is a convex polytope. Pardalos and Rosen [8] developed a partitioning and linear programming based algorithm with an approximation bound $\epsilon = \epsilon(Q)$, where $\epsilon(Q)$, a function of the QP data, is less than 1. Vavasis [10] and Ye [11] developed a polynomial-time algorithm, based on solving a ball-constrained quadratic problem, to compute an $(1 - \frac{1}{n^2})$-maximal solution. When the polytope is $\{x : -e \le x \le e\}$, Fu, Luo and Ye [2] further proved a $(1 - \frac{1}{n})$ polynomial-time algorithm.

In this note, we extend Goemans and Williamson and Nesterov's result to approximating (QP). We establish the same 4/7 result for approximating this problem. This result is based on a modification of Goemans and Williamson's algorithm and a generalization of Nesterov's proving technique.

2 Positive Semi-Definite Relaxation

The approximation algorithm for (QP) is to solve a positive semi-definite programming (SDP) relaxation problem

$$\text{(SDP)} \qquad \begin{array}{rll} \bar{s}(Q) := & \text{Maximize} & \langle Q, X \rangle \\ & \text{Subject to} & d(X) \le e, \ X \succeq 0. \end{array} \tag{1}$$

Here, $X \in \Re^{n \times n}$ is a symmetric matrix, $\langle \cdot, \cdot \rangle$ is the matrix inner product $\langle Q, X \rangle = \text{trace}(QX)$, $d(X)$ is a vector containing the diagonal components of X, and $X \succeq Z$ means that $X - Z$ is positive semi-definite.

The dual of the problem is

$$\begin{array}{rll} \bar{s}(Q) = & \text{Minimize} & e^T y \\ & \text{Subject to} & D(y) \succeq Q, \ y \ge 0, \end{array} \tag{2}$$

where $D(y)$ is the diagonal matrix such that $d(D(y)) = y \in \Re^n$. Denote by $\bar{X}(Q)$ and $\bar{y}(Q)$ an optimal solution pair for the primal (1) and dual (2).

The positive semi-definite relaxation was first proposed by Lovász and Shrijver [5], also see recent papers by Fujie and Kojima [3] and Polijak, Rendl and Wolkowicz [9]. This relaxation problem can be solved in polynomial time, e.g., see Nesterov and Nemirovskii [7].

We have the following relations between (QP) and (SDP).

Proposition 1 *Let $\bar{q} = \bar{q}(Q)$, $\underline{q} = -\bar{q}(-Q)$, $\bar{s} = \bar{s}(Q)$, $\underline{s} = -\bar{s}(-Q)$, and $\underline{y} = -\bar{y}(Q)$. Then,*

1. *$\underline{q}$ is the minimal objective value of $x^T Q x$ in the feasible set of (QP);*
2. *$\underline{s} = e^T \underline{y}$ and it is the minimal objective value of $\langle Q, X \rangle$ in the feasible set of (SDP);*
3. $$\underline{s} = -\bar{s}(-Q) \le \underline{q} = -\bar{q}(-Q) \le \bar{q}(Q) = \bar{q} \le \bar{s}(Q) = \bar{s}.$$

Proof. The first and second statements are straightforward to verify. Let $X = \bar{x}(Q)\bar{x}(Q)^T \in \Re^{n\times n}$. Then $X \succeq 0$, $d(X) \leq e$ and $\langle Q, X\rangle = q(\bar{x}(Q)) = \bar{q}(Q)$. Thus, we have $\bar{q}(Q) = \langle Q, X\rangle \leq \bar{s}(Q)$. Similarly, we can prove $\bar{q}(-Q) \leq \bar{s}(-Q)$, or $-\bar{s}(-Q) \leq -\bar{q}(-Q)$. ■

In what follows, we also let $\bar{x} = \bar{x}(Q)$, $\bar{X} = \bar{X}(Q)$. Since $\bar{X}$ is positive semi-definite, there is a factorization matrix $\bar{V} = (\bar{v}_1, \ldots, \bar{v}_n) \in \Re^{n\times n}$, i.e., $\bar{v}_j$ is the jth column of $\bar{V}$, such that $\bar{X} = \bar{V}^T\bar{V}$. The algorithm, similar to Goemans and Williamson [4], generates a random vector u uniformly distributed on an n-dimensional unit ball and then assigns

$$\hat{x} = \bar{D}\sigma(\bar{V}^T u), \tag{3}$$

where

$$\bar{D} = \mathrm{diag}(\|\bar{v}_1\|, \ldots, \|\bar{v}_n\|) = \mathrm{diag}(\sqrt{\bar{x}_{11}}, \ldots, \sqrt{\bar{x}_{nn}}),$$

and for any $x \in \Re^n$, $\sigma(x)$ is the vector whose components are $\mathrm{sign}(x_j)$, $j = 1, \ldots, n$, that is, $\mathrm{sign}(x_j) = 1$ if $x_j \geq 0$ and $\mathrm{sign}(x_j) = -1$ otherwise.

It is easily see that $\hat{x}$ is a feasible point for (QP) and we will show later that the expected objective value, $\mathrm{E}_u q(\hat{x})$, satisfies

$$\frac{\bar{q} - \mathrm{E}_u q(\hat{x})}{\bar{q} - \underline{q}} \leq \frac{\pi}{2} - 1 \leq \frac{4}{7}.$$

3 Approximation Analysis

The following two lemmas are analogues to Lemmas 1 and 2 of Nesterov [6].

Lemma 1

$$\bar{q}(Q) = \quad \textit{Maximize} \quad \sigma(V^T u)^T DQD\sigma(V^T u)$$

$$\textit{Subject to} \quad \|v_j\| \leq 1,\ j = 1, \ldots, n,\ \|u\| = 1,$$

where

$$D = \mathit{diag}(\|v_1\|, \ldots, \|v_n\|).$$

Proof. Since $D\sigma(V^T u)$ is a feasible point for (QP) for any feasible V and u, we have

$$\bar{q}(Q) \geq \sigma(V^T u)^T DQD\sigma(V^T u).$$

On the other hand, for any fixed u with $\|u\| = 1$, we let $v_j = \bar{x}_j u$, $j = 1, \ldots, n$. Then $D\sigma(V^T u) = \bar{x}$. Thus, for a particular feasible V and u we have

$$\bar{q}(Q) = q(\bar{x}) \leq \sigma(V^T u)^T DQD\sigma(V^T u).$$

These two give the desired result. ■

Lemma 2

$$\begin{array}{rll} \bar{q}(Q) = & \textit{Maximize} & \mathrm{E}_u(\sigma(V^T u)^T DQD\sigma(V^T u)) \\ & \textit{Subject to} & \|v_j\| \leq 1, \; j = 1, \ldots, n, \end{array}$$

where

$$D = diag(\|v_1\|, \ldots, \|v_n\|).$$

Proof. Again, since $D\sigma(V^T u)$ is a feasible point for (QP), we have for any feasible V

$$\bar{q}(Q) \geq \mathrm{E}_u(\sigma(V^T u)^T DQD\sigma(V^T u)).$$

On the other hand, for any fixed u with $\|u\| = 1$, we have

$$\mathrm{E}_u(\sigma(V^T u)^T DQD\sigma(V^T u)) = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}\|v_i\|\|v_j\|\mathrm{E}_u(\sigma(v_i^T u)\sigma(v_j^T u)). \tag{4}$$

Let us choose $v_i = \frac{\bar{x}_i}{\|\bar{x}\|}\bar{x}$, $i = 1, \ldots, n$. Then

$$\mathrm{E}_u(\sigma(v_i^T u)\sigma(v_j^T u)) = \begin{cases} 1 & \text{if } \sigma(\bar{x}_i) = \sigma(\bar{x}_j) \\ -1 & \text{otherwise.} \end{cases}$$

Thus,

$$\|v_i\|\|v_j\|\mathrm{E}_u(\sigma(v_i^T u)\sigma(v_j^T u)) = \bar{x}_i\bar{x}_j$$

which implies that for a particular feasible V

$$\bar{q}(Q) = q(\bar{x}) \leq \mathrm{E}_u(\sigma(V^T u)^T DQD\sigma(V^T u)).$$

These two give the desired result. ■

For any function of one variable $f(t)$ and $X \in \Re^{n\times n}$, let $f[X] \in \Re^{n\times n}$ be the matrix with the components $f(x_{ij})$. For example, $[X]^p$ denotes a matrix with the components x_{ij}^p. Nesterov [6] has also proved the next technical lemma.

Lemma 3 *Let* $X \succeq 0$ *and* $d(X) \leq 1$. *Then* $\arcsin[X] \succeq X$. ■

Now we are ready to prove the following theorem.

Theorem 1

$$\begin{array}{rll} \bar{q}(Q) = & \textit{Supremum} & \frac{2}{\pi}\langle Q, D\arcsin[D^{-1}XD^{-1}]D\rangle \\ & \textit{Subject to} & d(X) \leq e, \; X \succ 0, \end{array}$$

where

$$D = diag(\sqrt{x_{11}}, \ldots, \sqrt{x_{nn}}).$$

Proof. For any $X = V^TV \succ 0$, $d(X) \le e$, we have

$$\mathrm{E}_u(\sigma(v_i^Tu)\sigma(v_j^Tu)) = 1 - 2\Pr\{\sigma(v_i^Tu) \ne \sigma(v_j^Tu)\} = 1 - 2\Pr\{\sigma(\frac{v_i^Tu}{\|v_i\|}) \ne \sigma(\frac{v_j^Tu}{\|v_j\|})\}.$$

From Lemma 1.2 of Goemans and Williamson [4], we have

$$\Pr\{\sigma(\frac{v_i^Tu}{\|v_i\|}) \ne \sigma(\frac{v_j^Tu}{\|v_j\|})\} = \frac{1}{\pi}\arccos(\frac{v_i^Tv_j}{\|v_i\|\|v_j\|}).$$

Using the above lemma and equality (4) and noting $\arcsin(t) + \arccos(t) = \frac{\pi}{2}$ give the desired result. ■

Theorem 1 leads us to

Theorem 2 *We have*

1.

$$\bar{q} - \underline{s} \ge \frac{2}{\pi}(\bar{s} - \underline{s}).$$

2.

$$\bar{s} - \underline{q} \ge \frac{2}{\pi}(\bar{s} - \underline{s}).$$

3.

$$\bar{s} - \underline{s} \ge \bar{q} - \underline{q} \ge \frac{4-\pi}{\pi}(\bar{s} - \underline{s}).$$

Proof. Recall $\underline{y} = -\bar{y}(-Q) \le 0$, $\underline{s} = -\bar{s}(-Q) = e^T\underline{y}$, and $Q - D(\underline{y}) \succeq 0$. Thus, for any $X \succ 0$, $d(X) \le e$ and $D = \mathrm{diag}(\sqrt{x_{11}}, \ldots, \sqrt{x_{nn}})$, we have from Theorem 1

$$\begin{aligned}
\bar{q} = \bar{q}(Q) &\ge \frac{2}{\pi}\langle Q, D\arcsin[D^{-1}XD^{-1}]D\rangle \\
&= \frac{2}{\pi}\langle Q - D(\underline{y}) + D(\underline{y}), D\arcsin[D^{-1}XD^{-1}]D\rangle \\
&= \frac{2}{\pi}\Big(\langle Q - D(\underline{y}), D\arcsin[D^{-1}XD^{-1}]D\rangle + \langle D(\underline{y}), D\arcsin[D^{-1}XD^{-1}]D\rangle\Big) \\
&\ge \frac{2}{\pi}\Big(\langle Q - D(\underline{y}), DD^{-1}XD^{-1}D\rangle + \langle D(\underline{y}), D\arcsin[D^{-1}XD^{-1}]D\rangle\Big) \\
&\quad (\text{since } Q - D(\underline{y}) \succeq 0 \text{ and } \arcsin[D^{-1}XD^{-1}] \succeq D^{-1}XD^{-1}) \\
&= \frac{2}{\pi}\Big(\langle Q - D(\underline{y}), X\rangle + \langle D(\underline{y}), D\arcsin[D^{-1}XD^{-1}]D\rangle\Big) \\
&= \frac{2}{\pi}\Big(\langle Q, X\rangle - \langle D(\underline{y}), X\rangle + \langle D(\underline{y}), D\arcsin[D^{-1}XD^{-1}]D\rangle\Big) \\
&= \frac{2}{\pi}\Big(\langle Q, X\rangle - \underline{y}^Td(X) + \underline{y}^Td(D\arcsin[D^{-1}XD^{-1}]D)\Big)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{\pi}\Big(\langle Q, X\rangle - \underline{y}^T d(X) + \underline{y}^T(\frac{\pi}{2} d(X))\Big) \\
&= \frac{2}{\pi}\Big(\langle Q, X\rangle + (\frac{\pi}{2} - 1)\underline{y}^T d(X)\Big) \\
&\geq \frac{2}{\pi}\Big(\langle Q, X\rangle + (\frac{\pi}{2} - 1)\underline{y}^T e\rangle\Big) \\
&\quad \text{(since } 0 \leq d(X) \leq e \text{ and } \underline{y} \leq 0) \\
&= \frac{2}{\pi}\Big(\langle Q, X\rangle + (\frac{\pi}{2} - 1)\underline{s}\Big).
\end{aligned}
$$

Let X converge to $\bar{X}$, then $\langle Q, X\rangle \to \bar{s}$ and we have the desired first inequality.

Replacing Q with $-Q$ proves the second inequality in the theorem.

Adding the first two inequalities gives the third statement in the theorem. ∎

The result indicates that the positive semi-definite relaxation value $\bar{s} - \underline{s}$ is a constant approximation of $\bar{q} - \underline{q}$.

The following corollary can be derived from the proof of the above theorem.

Corollary 1 *Let $X = V^T V \succ 0$, $d(X) \leq e$, $D = diag(\sqrt{x_{11}}, \ldots, \sqrt{x_{nn}})$, and $\hat{x} = D\sigma(V^T u)$ where u with $\|u\| = 1$ is a random vector uniformly distributed on the unit ball. Moreover, let $X \to \bar{X}$. Then,*

$$
\lim_{X \to \bar{X}} \mathrm{E}_u(q(\hat{x})) = \lim_{X \to \bar{X}} \frac{2}{\pi}\langle Q, D \arcsin[D^{-1} X D^{-1}] D\rangle \geq \frac{2}{\pi}\bar{s} + (1 - \frac{2}{\pi})\underline{s}.
$$

Finally, we have

Theorem 3 *Let $\hat{x}$ be generated above from $X = \bar{X}$. Then*

$$
\frac{\bar{q} - \mathrm{E}_u q(\hat{x})}{\bar{q} - \underline{q}} \leq \frac{\pi}{2} - 1.
$$

Proof. Noting that

$$
\bar{s} \geq \bar{q} \geq \frac{2}{\pi}\bar{s} + (1 - \frac{2}{\pi})\underline{s} \geq (1 - \frac{2}{\pi})\bar{s} + \frac{2}{\pi}\underline{s} \geq \underline{q} \geq \underline{s}
$$

we have

$$
\begin{aligned}
\frac{\bar{q} - \mathrm{E}_u q(\hat{x})}{\bar{q} - \underline{q}} &\leq \frac{\bar{q} - \frac{2}{\pi}\bar{s} - (1 - \frac{2}{\pi})\underline{s}}{\bar{q} - \underline{q}} \\
&\leq \frac{\bar{q} - \frac{2}{\pi}\bar{s} - (1 - \frac{2}{\pi})\underline{s}}{\bar{q} - (1 - \frac{2}{\pi})\bar{s} - \frac{2}{\pi}\underline{s}} \\
&\leq \frac{\bar{s} - \frac{2}{\pi}\bar{s} - (1 - \frac{2}{\pi})\underline{s}}{\bar{s} - (1 - \frac{2}{\pi})\bar{s} - \frac{2}{\pi}\underline{s}}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(1-\frac{2}{\pi})(\bar{s}-\underline{s})}{\frac{2}{\pi}(\bar{s}-\underline{s})} \\
&= \frac{(1-\frac{2}{\pi})}{\frac{2}{\pi}} = \frac{\pi}{2} - 1.
\end{aligned}
$$

■

References

[1] M. Bellare and P. Rogaway, "The complexity of approximating a nonlinear program," *Mathematical Programming* 69 (1995) 429-442.

[2] M. Fu, Z.-Q. Luo and Y. Ye "Approximation algorithms for quadratic programming ," manuscript, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, CANADA L8S 4K1, 1996.

[3] T. Fujie and M. Kojima, "Semidefinite programming relaxation for nonconvex quadratic programs," Research Report B-298, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo 152, May 1995. To appear in *Journal of Global Optimization.*

[4] M. X. Goemans and D. P. Williamson, " Improved approximation algorithms for Maximum Cut and Satisfiability problems using semidefinite programming," *Journal of ACM* 42 (1995) 1115–1145.

[5] L. Lovász and A. Shrijver, "Cones of matrices and setfunctions, and $0-1$ optimization," *SIAM Journal on Optimization* 1 (1990) 166-190.

[6] Yu. E. Nesterov, "Quality of semidefinite relaxation for nonconvex quadratic optimization," CORE Discussion Paper, #9719, Belgium, March 1997.

[7] Yu. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms* (SIAM Publications, SIAM, Philadelphia, 1993).

[8] P. M. Pardalos and J. B. Rosen, *Constrained Global Optimization: Algorithms and Applications* (Springer-Verlag, Lecture Notes in Computer Sciences 268, 1987).

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

[10] S. A. Vavasis, *Nonlinear Optimization: Complexity Issues* (Oxford Science, New York, 1991).

[11] Y. Ye, "On affine scaling algorithms for nonconvex quadratic programming," *Mathematical Programming* 56 (1992) 285-300.