Dipartimento di Informatica e Sistemistica "Antonio Ruberti"
Università degli Studi di Roma "La Sapienza"

Efficient and Cheap Bounds for (Standard) Quadratic Optimization

Immanuel M. Bomze
Marco Locatelli
Fabio Tardella

Technical Report n. 10
2005

Aracne

I *Technical Reports* del Dipartimento di Informatica e Sistemistica "Antonio Ruberti" svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche.
Copie della presente pubblicazione possono essere richieste alla Redazione.

Dipartimento di Informatica e Sistemistica "Antonio Ruberti"
Università degli Studi di Roma "La Sapienza"

Via Eudossiana, 18 - 00184 Roma
Via Buonarroti, 12 - 00185 Roma
Via Salaria, 113 - 00198 Roma
www.dis.uniroma1.it

www.aracneeditrice.it
info@aracneeditrice.it

Redazione:
00173 Roma
via Raffaele Garofalo, 133 A/B
06 93781065
telefax 72678427

ISBN 88–548–0195-X

I edizione: luglio 2005

Finito di stampare nel mese di luglio del 2005
dalla tipografia « Grafica Editrice Romana S.r.l. » di Roma
per conto della « Aracne editrice S.r.l. » di Roma
Printed in Italy

EFFICIENT AND CHEAP BOUNDS FOR (STANDARD) QUADRATIC OPTIMIZATION[1]

Immanuel M. Bomze*, Marco Locatelli** and Fabio Tardella***

*Dept. of Statistics and Decision Support Systems
University of Vienna
Universitaetsstrasse 5, A-1010 Wien, Austria

**Dipartimento di Informatica
Università di Torino
Corso Svizzera 185, 10149 Torino, Italy

***Dipartimento di Matematica per le Decisioni
Economiche Finanziarie e Assicurative
Università di Roma "La Sapienza"
Via del Castro Laurenziano 9, 00161 Roma, Italy

e-mail (Bomze): immanuel.bomze@univie.ac.at
e-mail (Locatelli): locatelli@di.unito.it
e-mail (Tardella): fabio.tardella@uniroma1.it

Technical Report 10-05, DIS **July 2005**

[1]The research of M. Locatelli was partially supported by MIUR, FIRB 2001 Research Program *Large Scale Nonlinear Optimization*, Roma, Italy. I. Bomze thanks the Dipartimento di Informatica e Sistemistica at the Università di Roma "La Sapienza" for the hospitality as well as the fruitful and creative atmosphere he enjoyed while spending one month there on research purposes. Support for this stay has been provided under the visiting professor program of the University of Rome "La Sapienza" for the year 2004 (D.R. of 13.4.2004 n. 139).

ABSTRACT

A standard quadratic optimization problem (StQP) consists in minimizing a quadratic form over a simplex. A number of problems can be transformed into a StQP, including the general quadratic problem over a polytope and the maximum clique problem in a graph.

In this paper we present several polynomial-time bounds for StQP ranging from very simple and cheap ones to more complex and tight constructions. The main tools employed in the conception and analysis of most bounds are Semidefinite Programming and decomposition of the objective function into a sum of two quadratic functions, each of which is easy to minimize.

We provide a complete diagram of the dominance, incomparability, or equivalence relations among the bounds proposed in this and in previous works. In particular, we show that one of our new bounds dominates all the others. Furthermore, a specialization of such bound dominates Schrijver's improvement of Lovász's θ function bound for the maximum size of a clique in a graph.

Key Words: standard quadratic optimization, Semidefinite Programming, Quadratic Programming, maximum clique, resource allocation.

1 Introduction

A standard quadratic optimization problem (StQP) consists of finding (global) minimizers of a quadratic form over the standard simplex, i.e., we consider an optimization problem of the form

$$\ell_Q = \min\left\{x^\top Qx : x \in \Delta\right\}\ , \qquad (1)$$

where Q belongs to the class $\mathcal{M}$ of symmetric $n \times n$ matrices; a $^\top$ denotes transposition; and Δ is the standard simplex in the n-dimensional Euclidean space $\mathbb{R}^n$:

$$\Delta = \{x \in \mathbb{R}^n : e^\top x = 1,\ x \geq o\},$$

where $e = [1, \dots, 1]^\top \in \mathbb{R}^n$.

Note that a non homogeneous quadratic function $x^\top Ax + 2c^\top x$ over Δ can be easily homogenized by considering the rank-two update $Q = A + ec^\top + ce^\top$ in (1). Indeed, $x^\top(A + ec^\top + ce^\top)x = x^\top Ax + 2c^\top x$ over Δ.

While problem (1) seems to be a very special Quadratic Program (QP), it actually retains most of the complexity of the general case where Δ is replaced by any polyhedron P. Indeed, it is well known that (1) is NP-hard.

Furthermore, as shown in Sections 1.2 and 1.3 below, every Quadratic Program with a bounded feasible region can be reformulated as a Standard QP in higher dimension, or relaxed to a Standard QP in the same dimension.

Finally, Bomze also showed that *global* optimality of local solutions of general Quadratic Programs can be characterized by a finite number of copositivity conditions (in fact, not more than the number of non-binding constraints plus one) over polyhedral cones. These copositivity conditions in turn can be reformulated into Standard QPs (generally, in higher dimensions). For details see, e.g., [5] and the references therein.

An important tool for many exact or approximate solution methods for optimization problems is the availability of good and/or efficiently computable bounds on the optimum value of the problem. This well-known fact has induced some authors to propose a number of bounds for the Standard Quadratic Problem [1, 6, 9, 10, 12, 15, 29, 32]. However, in most cases no relation has been provided among the proposed bounds.

In this paper we present several new bounds for StQP and establish dominance, incomparability, or equivalence relations among them, as well as with respect to other previously introduced bounds. In particular, we show that one of our new bounds dominates all the others.

Furthermore, a specialization of such bound dominates Schrijver's improvement of Lovàsz's θ function bound for the maximum size of a clique in a graph.

The paper is organized as follows: after introducing some notation, we describe in more detail some relations between the Standard QP and the general QP. We also illustrate a reformulation of the Quadratic Resource Allocation Problem (including the portfolio selection problem) as a Standard QP, and we describe some connections of StQP with the maximum (weight) clique problem in a graph. Section 2 presents some cheap closed-form bounds while Sections 3, 4 and 5 are devoted to Lagrangian bounds, Convex Underestimation bounds, and Nowak's bound, respectively. Section 6 deals with Copositive bounds, and Section 7 introduces Decomposition bounds. In Section 8, we establish the relations between the previously discussed bounds,while Sections 9 and 10 return to the applications sketched in the Introduction.

1.1 Notation and cones of matrices

We now introduce some notation and present the cones of matrices that will be used in the sequel.

Let A, B $n \times n$ symmetric matrices (i.e., $A, B \in \mathcal{M}$). Recall that the *trace* of a matrix is the sum of its diagonal elements, and that for $A, B \in \mathcal{M}$

$$A \bullet B = \text{trace}(AB) = \sum_{ij} a_{ij} b_{ij}$$

is the standard inner product in $\mathcal{M}$.

If v is a vector in $\mathbb{R}^n$, we denote by Diag (v) the diagonal $n \times n$ matrix A with $a_{ii} = v_i$, for $i = 1, \ldots, n$. Conversely, for an $n \times n$ matrix A, diag (A) denotes the n-dimensional vector formed by the diagonal elements of A. Furthermore, Ddiag (A) denotes the matrix obtained from A by replacing all the off-diagonal entries with 0, i.e., Ddiag $(A) =$ Diag (diag (A)). Note that we have $v^\top \text{diag}\ (A) = \text{Diag}\ (v) \bullet A$. We denote by $I_n = \text{Diag}\ (e)$ the $n \times n$ identity matrix and its ith column (the standard basis vector) by e^i. Further, let $E^{ij} = \frac{1}{2}\left[e^i(e^j)^\top + e^j(e^i)^\top\right] \in \mathcal{M}$ be the matrix all entries of which are zero with the exception of two entries $(E^{ij})_{ij} = (E^{ij})_{ji} = \frac{1}{2}$ if $i \neq j$ while $E^{ii} = \text{Diag}\ (e^i)$.

In addition to the cone $\mathcal{M}$ of symmetric matrices we will use the following smaller convex cones:

- the cone $\mathcal{P}$ of all positive semidefinite symmetric matrices;
- the cone $\mathcal{N}$ of all nonnegative symmetric matrices;

- the cone $\mathcal{P} \cap \mathcal{N}$ of *doubly nonnegative* matrices;
- the cone of *copositive* matrices
$\mathcal{C} = \{C \in \mathcal{M} : x^\top C x \geq 0 \text{ for all } x \in \mathbb{R}^n_+\}$;
- the cone of *completely positive* matrices
$\mathcal{C}^* = \{D \in \mathcal{M} : D = YY^\top,\ Y \text{ some } n \times k \text{ matrix with } Y_{ij} \geq 0,\ \text{all } i, j\}$.

On the set $\mathcal{M}$ of symmetric matrices we will use both the standard partial order $\leq$ defined by componentwise inequalities, i.e., $A \geq B$ whenever $a_{ij} \geq b_{ij}$ for all i and j, and the Löwner partial order $\succeq$ induced on $\mathcal{M}$ by the cone $\mathcal{P}$ of positive semidefinite matrices. Thus we write $A \succeq B$ whenever $A - B \in \mathcal{P}$.

Recall that the (convex) dual cone of a cone $\mathcal{K}$ of matrices with respect to the standard inner product of $\mathcal{M}$ is the cone

$$\mathcal{K}^* = \{Y \in \mathcal{M} : X \bullet Y \geq 0,\ \text{for all } X \in \mathcal{K}\}.$$

It is well known that the completely positive cone is the dual of the copositive cone (which justifies the notation $\mathcal{C}^*$), and that the nonnegative and semidefinite cones are self-dual with respect to this inner product.

Recall also that

$$\mathcal{K}_0 = \mathcal{P} + \mathcal{N}$$

is a (zero-order) inner approximation [10] of the copositive cone: $\mathcal{K}_0 \subseteq \mathcal{C}$, with $\mathcal{K}_0 = \mathcal{C}$ if and only if $n \leq 4$ [14]. As $\mathcal{N}^* = \mathcal{N}$ and $\mathcal{P}^* = \mathcal{P}$, we then have

$$\mathcal{C}^* \subseteq \mathcal{K}_0^* = (\mathcal{P} + \mathcal{N})^* = \mathcal{P} \cap \mathcal{N}.$$

1.2 StQP formulation of a general QP over a polytope

If the vertices of a polytope P are known, then the problem of minimizing a function over P can be easily transformed into the problem of minimizing a function on the standard simplex Δ with the change of variables $x = Vy$, where V is the matrix whose column vectors are the vertices $v^1, \ldots, v^N$ of P.

Clearly the transformed problem might have a number of variables that is exponential with respect to the original one. Nevertheless, by means of the above transformation, several theoretical and algorithmic results for StQP can be applied to the problem of minimizing a quadratic function on a more general polytope. In Section 10.1 we use this remark to provide some bounds for the minimum of a quadratic function on the unit ball in the ℓ^1 norm, improving an SDP bound proposed by Nesterov [30] for this problem.

1.3 StQP relaxation of bounded Quadratic Programs

Consider the Quadratic Program

$$\min\{y^\top Cy + 2c^\top y : \; y \in P\} \tag{2}$$

where $P = \{y \in \mathbb{R}^n : Ay = b, \; y \geq o\}$, and A is an $m \times n$ matrix. If P is bounded and $P \neq \{o\}$, there are a vector $p \geq e$ and a number $\pi > 0$ - which can both be obtained by solving a single LP of the size of A - such that P is contained in the intersection of the hyperplane $p^\top y = \pi$ and of the non-negative orthant (see Section 10.2 for details). Thus, by setting

$$D = \pi(\mathrm{Diag}\; p)^{-1} \quad \text{and} \quad Q = DCD + Dce^\top + ec^\top D, \tag{3}$$

we obtain that the Standard QP (1) is a valid relaxation of (2). This obviously implies that any valid lower bound for (1) is also a valid lower bound for (2). This fact will be used in Section 10.2 to provide lower bounds for a general QP with a bounded feasible region.

1.4 Quadratic Resource Allocation and Portfolio Optimization

Given n activities using a common resource R with intensities $a_1, \ldots, a_n$, a *resource allocation problem* consists in finding the levels $x_1, \ldots, x_n$ of these activities that maximize a utility function $f(x_1, \ldots, x_n)$ (see, e.g., [23]). This problem can be formulated as follows

$$\max\left\{f(x) : a^\top x \leq R, \; x \geq o\right\} . \tag{4}$$

Note that a simple scaling of the variables, like in Section 1.3, and the introduction of an additional slack variable allows to transform the constraints of this problem into the form $x \in \Delta$, so that a resource allocation problem with a quadratic utility function is essentially a StQP.

In many applications the utility function f is assumed to be a separable convex function, which considerably simplifies the problem. However, some non-separable and non-convex models are also needed in some cases. An important example is the familiar (Markowitz) mean/variance portfolio selection problem (see, e.g. [26, 27]), which can be formalized as follows: suppose there are n securities to invest in, at an amount expressed in relative shares $x_i \geq 0$ of an investor's budget. Thus, the budget (resource allocation) constraint reads $e^\top x = 1$, and the set of all feasible portfolios is given by Δ. Now, given the expected return m_i of security i during the forthcoming period, and an $n \times n$ covariance matrix

C across all securities, the investor faces the multiobjective problem to maximize expected return $m^\top x$ and simultaneously minimize the risk $x^\top Cx$ associated to her decision x.

One of the most popular approaches to such type of problems is that the user prespecifies a parameter β which in her eyes balances the benefit of high return and low risk. This leads to the parametric QP

$$\max \left\{ f_\beta(x) = m^\top x - \beta x^\top Cx : x \in \Delta \right\} . \tag{5}$$

Note that, for fixed β, this is again a Standard QP.

In theory the matrix C is, as an exact covariance matrix, positive semidefinite (and it could be singular in many applications, see [27]), so that (5) is a convex problem. On the other hand, securities usually are highly correlated, and in time-series analysis one frequently encounters the situation that some of the most reliable estimators $\tilde{C}$ of the unknown covariance matrix C lack semidefiniteness properties [31], [35, pp.134 ff]. Hence, the portfolio optimization problem can be transformed into a (possibly non-convex) parametric Standard QP. Furthermore, Best and Ding [2] show how to reduce the parametric problem to a *single* Standard QP with indefinite Q under some assumptions.

1.5 StQP formulation of the maximum weight clique problem

Consider an undirected graph $\mathcal{G} = (N, \mathcal{E})$ with n nodes. A *clique* S is a subset of the *node set* N which induces a complete subgraph of $\mathcal{G}$ (i.e., any pair of nodes in S is joined by an edge in $\mathcal{E}$, the *edge set*). A clique S is *maximal* if there is no larger clique containing S. A (maximal) clique is a *maximum* clique if it has the largest number of elements among all cliques. In the weighted case we associate weights $w_i > 0$ to the nodes, and define a (separable) *weight function* $W(S) = \sum_{i \in S} w_i$ on the subsets S of nodes. A *maximum weight clique* is then a clique that maximizes the weight function $W(S)$ among all cliques of the graph.

Motzkin and Straus [28] were the first to observe that the maximum clique problem can be formulated as a very special (non-convex) Standard QP. This obviously implies NP-hardness of the general StQP. The StQP formulation of the maximum clique problem has been extended to the weighted case in [18] by exploiting an idea of Lovász. Tardella [40] showed that the same result can also be derived as a consequence of an extension of the Fundamental Theorem of Linear Programming, and proved, in addition, a somewhat converse relation. Indeed, it is shown in [40] that the StQP (1) can be solved by finding a maximal clique in an associated graph that maximizes a suitable non-separable weight function $\widetilde{W}(S)$. Recently, Pavan and Pelillo described a similar reduction of a non-standard maximum weight clique problem used for clustering and image segmentation to StQP [33]. These relations can be obviously exploited to transform results (and, in particular bounds) for Standard QPs into results for the maximum (weight) clique problem, and conversely (see Section 9).

In order to describe the relation between the maximum weight clique problem and the Standard QP we need to introduce a subset of the class $\mathcal{M}$ of symmetric matrices called the Motzkin-Straus class of matrices $\mathcal{M}(w, \mathcal{G})$ associated to a graph $\mathcal{G} = (N, \mathcal{E})$ and a vector w in $\mathbb{R}^n$ with $w_i > 0$ for all i. We define

$$\mathcal{M}(w, \mathcal{G}) = \{B \in \mathcal{M} : b_{ij} \geq \tfrac{b_{ii}+b_{jj}}{2}, \text{ if } \{i,j\} \notin \mathcal{E}\,; \, b_{ij} = 0, \text{ if } \{i,j\} \in \mathcal{E}\,; \text{ and } b_{ii} = \tfrac{1}{w_i}, \text{ for } i \in N\}\,. \tag{6}$$

In [18] it is shown that the StQP (1) attains the same optimal value for every Q in $\mathcal{M}(w, \mathcal{G})$ and that the inverse of this value coincides with the value of a maximum weight clique on the graph $\mathcal{G}$ with weights w on the nodes, i.e.,

$$\ell_Q = \min\left\{x^\top Q x : x \in \Delta\right\} = 1/W(S^*) \quad \text{for all } Q \in \mathcal{M}(w, \mathcal{G})\,, \tag{7}$$

where S^* is a maximum weight clique of $\mathcal{G}$. Regularization techniques, i.e., modifying the class $\mathcal{M}(w, \mathcal{G})$ may be necessary to extract such an S^* from the solution of (7), see [4].

In the special case where $E = ee^\top$ is the matrix with all ones and $A_\mathcal{G}$ is the node-node adjacency matrix of the graph $\mathcal{G}$, we have $E - A_\mathcal{G} \in \mathcal{M}(e, \mathcal{G})$. Thus, the size $\omega(\mathcal{G})$ of a maximum clique (the *clique number*) of $\mathcal{G}$ is given by

$$\omega(\mathcal{G}) = 1/\ell_{E-A_\mathcal{G}}\,, \tag{8}$$

which is a reformulation of the classical result of Motzkin and Straus [28].