1

Ellipsoidal Approach to Box-Constrained Quadratic Problems

PASQUALE L. DE ANGELIS[1], IMMANUEL M. BOMZE[2] and GERARDO TORALDO[3]
[1]*Institute of Statistic and Mathematics, University of Naples 'PARTHENOPE', Naples, Italy*
[2]*Department of Statistics and Decision Support Systems, University of Vienna, Viennna, Austria*
[3]*CPS-CNR & Department of Agricultural Engineering and Agronomy, University of Naples 'Federico II', Naples, Italy*

Abstract. We present a new heuristic for the global solution of box constrained quadratic problems, based on the classical results which hold for the minimization of quadratic problems with ellipsoidal constraints. The approach is tested on several problems randomly generated and on graph instances from the DIMACS challenge, medium size instances of the Maximum Clique Problem. The numerical results seem to suggest some effectiveness of the proposed approach.

Key words: Nonconvex quadratic programming, Ellipsoidal constraints, Global optimization

1. Introduction

We consider global optimization problems of the form

$$\min\{f(x)=\frac{1}{2}x^{\top}Ax+b^{\top}x:x\in[-1,1]^{n}\} \tag{1.1}$$

where $A=[a_{ij}]\in\mathbb{R}^{n\times n}$ is a symmetric $n\times n$ matrix and $x,b\in\mathbb{R}^{n}$ with $x^{\top}$ denoting the transposed vector. Problems of such form appear in a number of applications. Moreover (1.1) arises as subproblem in the solution of more complex optimization problems. For the convex case, which is known to be polynomially solvable [12], many different approaches have been proposed. Also for the concave case, which is equivalent to the 0–1 quadratic problem, several algorithms exist based on the combinatorial nature of that problem. Here public domain packages like Q01SUBS [18] are available. However, there are very few algorithms for the general indefinite case, where the matrix A has eigenvalues of mixed sign (see, for example, a recent paper [20] or [7] for a survey on bound constrained quadratic programming). In general only problems with few variables can be handled with exact algorithms, and therefore, for medium-size and large problems, heuristics have to be devised, which iteratively produce a sequence of local solutions. Obviously, the hard task calling for heuristic approaches is to escape from an inefficient local solution towards an improving one.

The paper is organized as follows. In Section 2 we first review some results taken from the trust region literature, about global minimization of a quadratic

function over a ball. The key result for the purposes addressed in this paper is apparently new: the Lagrange multiplier for the global solution (which is known to be unique) depends continuously, and in an decreasing manner, on the radius of the ball. In fact, this dependence is even smooth, with the — possible, but from a generic viewpoint highly improbable — exception of one kink. We also show that a global solution can be chosen such that it depends continuously (again, generically, even smoothly) on the radius. Note that in degenerate cases, the global solution need not be unique. In Section 3 we present a heuristic algorithm for the problem (1.1), in which a local solver is used, using a 'promising' starting point whose computation is motivated by the results in Section 2. Such an algorithm can naturally be embedded in a branch and bound strategy. Finally, Section 4 presents some computational results on a wide variety of test problems. In particular the algorithm is tested on small and medium sized graph instances from the DIMACS Challenge.

2. Local and global minimizers for quadratic functions over a ball

In this section we study the auxiliary problem

$$\min\{f(x) = \frac{1}{2}x^\top A x + b^\top x : x \in \mathbb{R}^n \text{ with } x^\top x \leqslant \delta\} \tag{2.2}$$

It is well-known in the trust-region literature, see, e.g., [14, 13, 19], that there is at most one local solution to (2.2) which is not global, and generically, the global solution to (2.2) is unique. Moreover, we exhibit a way how to obtain a global solution which depends continuously on the ball radius, or, equivalently, on δ (as usual, we denote by $\|x\| = \sqrt{x^\top x}$). It will turn out that this dependence is in fact smooth, with the possible exception of a single kink at some critical δ_{crit} in degenerate cases. Similarly, the corresponding Lagrange multiplier depends (again, with the possible exception of one kink at δ_{crit}) smoothly and in a monotonically decreasing way on the radius. While most of the used technology is folklore by now, the monotonicity and smoothness results seem to be new, and in order to establish them we anyhow need the prerequisites which were used previously. Thus we decided to present also short proofs of known results, to make the presentation self-contained, and in particular to shed some light on the genericity conditions which exclude degeneration of the problem.

We start with some second-order optimality conditions for Karush/Kuhn/Tucker (KKT) points $\bar{x}$ of (2.2).

Recall that such a point satisfies

$$\text{(i)} \quad (A + \mu I_n)\bar{x} + b = 0; \quad \text{(ii)} \quad \mu(\bar{x}^\top \bar{x} - \delta) = 0;$$
$$\text{with (iii)} \quad \bar{x}^\top \bar{x} \leqslant \delta \text{ and } \mu \geqslant 0. \tag{2.3}$$

The last conditions (iii) are the primal (for $\bar{x}$), and dual (for the *Lagrange multiplier* μ) feasibility conditions, respectively, while (i) and (ii) represent first-order

optimality and complementarity. Note that as Slater's constraint qualifications are satisfied for the problem (2.2), all local solutions must satisfy (2.3).

First observe that if there is no local solution at the boundary of the ball, then from compactness of the feasible region there must be a local solution $\bar{x}$ in the interior (namely at least the global one), and therefore the unrestricted second-order optimality conditions familiar from elementary calculus imply that A must be positive semidefinite. Clearly, $A\bar{x} = -b$ but there there is no alternative solution to $Ax = -b$ (otherwise we would obtain a global solution also at the boundary). Hence A cannot be singular, and therefore A must be positive-definite. Consequently, the optimal solution reads $x^\star(\delta) = -A^{-1}b$ for all $\delta \geqslant b^\top A^{-2}b$. For smaller δ, we again must have (local) solutions on the boundary of the ball. So without loss of generality we may and do investigate in the sequel only solutions that satisfy $\bar{x}^\top \bar{x} = \delta$, regardless of the definiteness of A.

Returning to the general case, the second-order condition necessary for local optimality of a KKT point $\bar{x}$ can be expressed as a lower bound for μ in terms of the spectrum of A. Denote by

$$\sigma_1 = \sigma_2 = \cdots = \sigma_k < \sigma_{k+1} \leqslant \cdots \leqslant \sigma_n$$

the eigenvalues of A, where the smallest eigenvalue of A has multiplicity k (possibly $k = 1$). Observe that the case where all eigenvalues coincide, i.e., where $A = \sigma_1 I_n$, is trivial: unless $b = 0$ (where $x^\star = o$ is the global solution for any $\delta > 0$ if $\sigma_1 \geqslant 0$ while $x^\star$ is some x with $x^\top x = \delta$ if $\sigma_1 < 0$), we get $x^\star = -b/(\sigma_1 + \mu)$ where $\mu = -\sigma_1 + \sqrt{(b^\top b)/\delta}$ if this quantity is positive, or else $x^\star = -b/\sigma_1$ in the interior of the ball. So again without loss of generality let us assume in the sequel that A has a nonzero spectral spread.

The next three results were established by Martínez [14]. For the readers' convenience, we provide a self-contained (and slightly simplified) presentation. In the sequel, we sometimes use coordinates relative to the orthogonal eigensystem $u_1, \ldots u_n$ of A, where u_i satisfy $u_i^\top u_i = 1$ and $Au_i = \sigma_i u_i$. Then A is replaced with $U^\top AU = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$, and b with $c = U^\top b$, while the constraint $x^\top x = vU^\top Uv = v^\top v \leqslant \delta$ remains unchanged.

THEOREM 1 (**Lemma 2.4 in [14]**) *If $\bar{x}$ is a local solution to (2.2), then $\bar{x}$ satisfies (2.3), i.e., $\bar{x}$ is a KKT point of (2.2), and the Lagrange multiplier μ must necessarily satisfy*

$$\mu \geqslant -\sigma_2. \tag{2.4}$$

Proof. We adapt a similar proof for *global* optimality conditions in [11]. Put $\varepsilon > 0$ (to be chosen later) and let d be some direction with $-\varepsilon < d^\top \bar{x} < 0$. Then for $t = -2d^\top \bar{x}/d^\top d$, we get

$$(\bar{x} + td)^\top (\bar{x} + td) = \bar{x}^\top \bar{x} - 4\frac{d^\top \bar{x}}{d^\top d}d^\top \bar{x} + 4\frac{(d^\top \bar{x})^2}{(d^\top d)^2}d^\top d = \delta,$$

so that $\bar{x}+td$ is feasible to (2.2). Now if $\rho > 0$ is so small that $f(x) \geqslant f(\bar{x})$ for every feasible x with $\|x-\bar{x}\| < \rho$, choose $\varepsilon = \rho\frac{\|d\|}{3}$. Then, by construction of t, we get for $x = \bar{x}+td$ that $\|x-\bar{x}\| < \rho$ and thus $f(x) \geqslant f(\bar{x})$. But

$$
\begin{aligned}
f(\bar{x}+td) - f(\bar{x}) &= td^\top(A\bar{x}+b) + \tfrac{t^2}{2}d^\top Ad \\
&= td^\top(-\mu\bar{x}) + \tfrac{t^2}{2}d^\top Ad \\
&= \tfrac{t^2}{2}d^\top(A+\mu I_n)d + \tfrac{\mu}{2}\left[-2td^\top\bar{x} - t^2 d^\top d\right] \\
&= \tfrac{t^2}{2}d^\top(A+\mu I_n)d + \tfrac{\mu}{2}\left[\|\bar{x}\|^2 - \|\bar{x}+td\|^2\right] \\
&= \tfrac{t^2}{2}d^\top(A+\mu I_n)d.
\end{aligned}
\tag{2.5}
$$

Thus we arrive at the condition $d^\top(A+\mu I_n)d \geqslant 0$ for all directions d with $-\varepsilon < d^\top\bar{x} < 0$. Now let v be a direction perpendicular to $\bar{x}$. Then v can be arbitrarily well approximated by such a direction d, and continuity of the quadratic form finally yields

$$
v^\top(A+\mu I_n)v \geqslant 0 \quad \text{for all } v \perp \bar{x}.
\tag{2.6}
$$

Evidently, this cannot happen if $A+\mu I_n$ has more than one negative eigenvalue (including multiplicities). Hence the result. □

Theorem 1 can be also derived from much more general second-order optimality conditions for nonlinear optimization problems, see, e.g., [6] (and observe that Abadie's constraint qualification for $M_\lambda = \{x \in \mathbb{R}^n : \|x\| = \delta\}$ if $\lambda = \mu > 0$ are satisfied in the present context). However, for the readers' convenience we here provided a concise direct proof.

The global case is also treated in Theorem 2.1 of [11] (earlier references are Lemmas 2.2 and 2.3 in [14] or, e.g., [17]):

THEOREM 2 *Suppose that $\bar{x}$ satisfies (2.3), i.e., $\bar{x}$ is a KKT point of (2.2). Then $\bar{x}$ is a global solution to (2.2) if and only if the Lagrange multiplier μ satisfies*

$$
\mu \geqslant -\sigma_1.
\tag{2.7}
$$

Proof. The assertion is equivalent to the requirement that $A+\mu I_n$ be positive semi-definite, as formulated in [11]. □

The next step deals with a non-generic case which complicates analysis, and shows that then no non-global solution may exist:

LEMMA 3 (**Lemma 3.2 in [14]**). *If $b \in \{u_1, \dots, u_k\}^\perp$, then no non-global local solution to (2.2) can exist.*

Proof. Let μ be the Lagrange multiplier of a non-global local solution $\bar{x}$ to (2.2). Then, by Theorems 1 and 2, $-\sigma_2 \leqslant \mu < -\sigma_1$. Thus necessarily $\sigma_1 < \sigma_2$ so that $k = 1$. Now by hypothesis,

$$0 = -b^\top u_1 = u_1^\top (A + \mu I_n)\bar{x} = (\sigma_1 + \mu)u_1^\top \bar{x}$$

and $\mu \neq -\sigma_1$ imply $\bar{x} \perp u_1$. On the other hand, $u_1^\top (A + \mu I_n)u_1 = \sigma_1 + \mu < 0$, contradicting (2.6). Hence the result. □

As a consequence of the previous proof, existence of a local, non-global solution implies simplicity of the smallest eigenvalue σ_1 of A (i.e., $k = 1$).

Next we attack the question how many local solutions to (2.2) may coexist: to this end the so-called *secular function* [19]

$$\varphi(t) = \sum_{i=1}^{n} \frac{c_i^2}{(\sigma_i + t)^2}$$

plays an important role. This function is well-defined on $D_\varphi = \mathbb{R}_+ \setminus \{-\sigma_1, \ldots, -\sigma_n\}$, and φ is strictly convex on D_φ. Observe that if the multiplier $\mu \in D_\varphi$, then μ determines the KKT point uniquely since $A + \mu I_n$ is then non-singular and

$$\bar{x} = -(A + \mu I_n)^{-1}b, \tag{2.8}$$

as can be seen from the conditions (2.3)(i). But the remaining (ii) and (iii) also imply that for any $\mu \in D_\varphi \setminus \{0\}$, we get $\bar{x} = U\bar{v}$ with

$$\bar{v}_i = -\frac{c_i}{\sigma_i + \mu},$$

where μ satisfies $\varphi(\mu) = \sum_{i=1}^{n}(\bar{v}_i)^2 = \delta$. This establishes the significance of the secular function φ.

We next show that under conditions which are generically (w.r.t. A and b) satisfied, there are at most three local solutions to (2.2).

THEOREM 4   *Suppose that $b \notin [\{u_1, \ldots, u_k\}^\perp \cup \{u_{k+1}, \ldots, u_m\}^\perp]$ where $\{u_{k+1}, \ldots, u_m\}$ is the eigensystem to the eigenvalue σ_{k+1} with multiplicity $m - k + 1 \geqslant 1$ (if the smallest two eigenvalues of A are simple, this condition reduces to $(b^\top u_1)(b^\top u_2) \neq 0$). Then there are at most three local solutions to (2.2) at the boundary of the ball. Every such local solution $\bar{x}$ satisfies $\bar{x} = U\bar{v}$ with $\bar{v}_i = -c_i/(\sigma_i + \mu)$, where $\mu > -\sigma_{k+1}$ satisfies $\mu \neq -\sigma_1$ and $\varphi(\mu) = \delta$.*

Proof. In coordinates w.r.t. the eigensystem of A, the assumption $b \notin \{u_1, \ldots, u_k\}^\perp$ reads $c_i = b^\top u_i \neq 0$ for some $i \in \{1, \ldots, k\}$. Without loss of generality we assume $c_1 \neq 0$. But as

$$\text{diag}\,(\sigma_1 + \mu, \ldots, \sigma_n + \mu)\bar{v} = U^\top(A + \mu I_n)UU^\top \bar{x} = U^\top(-b) = -c$$

we get $(\sigma_1 + \mu)v_1 = -c_1 \neq 0$, so that $\mu \neq -\sigma_1$. Similarly we obtain from $b \notin \{u_{k+1}, \ldots, u_m\}^\perp$ (again, after rearranging indices in case of $m - k + 1 > 1$ if necessary) from $c_{k+1} \neq 0$ also $\mu \neq -\sigma_{k+1}$. Hence we obtain via Theorem 1 that $\mu \in D_\varphi \cap \,] -\sigma_{k+1}, +\infty[$. On this region, the secular function φ has only one pole located at $-\sigma_1$. For $t \to \sigma_1$, we get $\varphi(t) \to +\infty$, and also $\varphi(t) \to +\infty$ as $t \searrow -\sigma_{k+1}$ while $\varphi(t) \searrow 0$ as $t \nearrow +\infty$. From strict convexity of φ on this region, we infer that there are at most three distinct solutions $\mu_{1,2,3}$ to the equation $\varphi(\mu) = \delta$ (at least one for every δ, two for a particular value of δ, and three for all larger δ). To every such μ_i there corresponds exactly one KKT point, and so there are at most three local solutions. $\qquad\square$

Note that similar but refined arguments are used by Lucidi *et al.* in [13] to bound the number of KKT points of (2.2). Also, the last result can be sharpened considerably for global solutions:

THEOREM 5 *Suppose that $b \notin \{u_1, \ldots, u_k\}^\perp$ (if the smallest eigenvalue of A is simple, this reduces to $b^\top u_1 \neq 0$). Then there is exactly one global solution to (2.2) at the boundary of the ball. This solution $\bar{x}$ satisfies $\bar{x} = U\bar{v}$ with $\bar{v}_i = -c_i/(\sigma_i + \mu)$, where $\mu > -\sigma_1$ satisfies $\varphi(\mu) = \delta$.*

Proof. The arguments are virtually the same as in the previous theorem. Indeed, using Theorem 2, we see that now the relevant part of D_φ shrinks to the region $\mathbb{R}_+ \cap \,] -\sigma_1, +\infty[$, upon which the secular function is even smooth. $\qquad\square$

Now we prove the sharpening of Theorem 4 due to Martínez:

THEOREM 6 (**Lemma 3.2 in [14]**) *There is at most one local, non-global solution $\bar{x}$ to (2.2), and this solution has a Lagrange multiplier μ which satisfies*

$$\varphi'(\mu) \geqslant 0$$

where φ' is the derivative of the secular function.

Proof. We only have to establish the asserted inequality. The remainder follows from the arguments in Theorem 4 (use Lemma 3). Starting point of our reasoning is the relation (2.6). We first construct a basis of $\bar{x}^\perp$. Consider the vectors

$$w_j = \frac{c_j}{\sigma_j + \mu} u_1 - \frac{c_1}{\sigma_1 + \mu} u_j, \quad j = 2, \ldots, n.$$

Since $c_1 \neq 0$ as shown in the proof of Theorem 4, these vectors $\{w_2, \ldots, w_n\}$ are linearly independent and they all satisfy $w_j \perp \bar{x}$, because of $\bar{x} = U\bar{v} = -\sum_{i=1}^n c_i u_i/(\sigma_i + \mu)$, so that

$$w_j^\top \bar{x} = \sum_{i=1}^n \frac{c_i}{\sigma_i + \mu} \left[\frac{c_j}{\sigma_j + \mu} u_1^\top u_i - \frac{c_1}{\sigma_1 + \mu} u_j^\top u_i \right] = 0.$$

Thus, forming the $n \times (n-1)$ matrix $B = [w_2, \ldots, w_n]$, we see that condition (2.6) can be rephrased as a positive-semidefiniteness condition for the symmetric $(n-1) \times (n-1)$ matrix $B^\top(A + \mu I_n)B$. Now observe that

$$(A + \mu I_n)w_j = \frac{c_j(\sigma_1 + \mu)}{\sigma_j + \mu}u_1 - \frac{c_1(\sigma_j + \mu)}{\sigma_1 + \mu}u_j,$$

whence we deduce that the entry in the ith column and jth row of $B^\top(A + \mu I_n)B$ equals

$$w_i^\top(A + \mu I_n)w_j = \frac{c_i c_j(\sigma_1 + \mu)}{(\sigma_i + \mu)(\sigma_j + \mu)} + \frac{c_1^2(\sigma_i + \mu)^2}{(\sigma_1 + \mu)^2}\delta_{ij},$$

where δ_{ij} is the Kronecker symbol. Written in compact form, we thus obtain

$$B^\top(A + \mu I_n)B = D - ss^\top,$$

where D is a diagonal, positive-definite matrix, and s has coordinates

$$s_j = \sqrt{|\sigma_1 + \mu|}\frac{c_j}{\sigma_j + \mu},$$

for $j = 2, \ldots, n$. Hence $B^\top(A + \mu I_n)B$ is positive-semidefinite if and only if $I_{n-1} - (D^{-1/2}s)(D^{-1/2}s)^\top$ is so, with

$$D^{-1/2} = \mathrm{diag}\left(\frac{|\sigma_1 + \mu|}{|c_1|\sqrt{\sigma_j + \mu}}\right)_{j=2,\ldots,n}$$

designating the inverse of the square-root factorization of D. But a symmetric rank-one update of I_{n-1} in this form, $I_{n-1} - vv^\top$ is positive-semidefinite if and only if $v^\top v \leqslant 1$, as a direct argument via $v^\perp$ and the Cauchy-Schwarz-Bunyakovsky inequality shows. Finally, we calculate

$$1 \geqslant (D^{-1/2}s)^\top(D^{-1/2}s) = s^\top D^{-1}s = -\frac{(\sigma_1 + \mu)^3}{c_1^2}\sum_{j=2}^{n}\frac{c_j^2}{(\sigma_j + \mu)^3},$$

which obviously is equivalent to $\varphi'(\mu) \geqslant 0$. □

Finally, we obtain a simple genericity condition guaranteeing that there are at most two local solutions:

COROLLARY 7 *If there is a local, non-global solution $\bar{x}$ to (2.2), then there is no other local-nonglobal solution. If, furthermore, $b^\top u_1 \neq 0$, then there are in total two local solutions, namely $\bar{x}$ and the (unique) global solution x^*.*

Proof. Existence of $\bar{x}$ implies, as observed in the proof of Lemma 3, $k = 1$. Hence the result follows by Theorems 6 and 5.   □

At the end of this section, we investigate continuous dependence of the global solution from the parameter δ (the square of the ball radius). For completeness, we also deal with the non-generic case where $b \in \{u_1, \ldots, u_k\}^\perp$, and provide an explicit construction both for the multiplier $\mu^\star(\delta)$ and the optimal solution $x^\star(\delta)$.

THEOREM 8 *For varying δ, we can choose a global solution $x^\star(\delta)$ to (2.2), which depends continuously upon δ. This dependence is smooth if $b \notin \{u_1, \ldots, u_k\}^\perp$, whereas it has exactly one point of non-differentiability in case of $b \in \{u_1, \ldots, u_k\}^\perp$. Also the corresponding multiplier $\mu^\star(\delta)$ exhibits the same smoothness and continuity behavior.*

Proof. The first assertion can be derived from Theorem 5, invoking the Implicit Function Theorem, applied to the equation $F(t, \delta) = \varphi(t) - \delta$ on the region $]-\sigma_1, +\infty[\times]0, +\infty[$ where F is well-defined and smooth. However, if $b \in \{u_1, \ldots, u_k\}^\perp$, we have to argue more carefully since there may be more than one global solution, and since the case $\mu = -\sigma_1$ (the largest potential pole of φ) is not ruled out a priori. However, in this case we get

$$\varphi(t) = \sum_{i > k} \frac{c_i^2}{(\sigma_i + t)^2},$$

which is well defined for all $t > -\sigma_{k+1}$ (in particular, for $t = -\sigma_1$). Now put $\delta_{\text{crit}} = \varphi(-\sigma_1)$. If $\delta < \delta_{\text{crit}}$, then there is a unique $\mu^\star(\delta) = \mu > -\sigma_1$ such that $\varphi(\mu) = \delta$, since φ decreases strictly on $[-\sigma_1, +\infty[$. In this case, put

$$v_i^\star(\delta) = 0 \text{ if } i \leqslant k \quad \text{whereas} \quad v_i^\star(\delta) = -\frac{c_i}{\sigma_i + \mu} \text{ if } i > k.$$

Then $x^\star(\delta) = Uv^\star(\delta)$ is the only KKT point with multiplier $\mu^\star(\delta) \geqslant -\sigma_1$, and thus the only global solution. Again invoking the Implicit Function Theorem, we obtain smooth dependence of the above quantities upon δ.
If, however, $\delta \geqslant \delta_{\text{crit}}$, then define $\mu^\star(\delta) = -\sigma_1$, a constant in δ, and put

$$v_i^\star(\delta) = \sqrt{[\delta - \varphi(-\sigma_1)]/k} \text{ if } i \leqslant k \quad \text{whereas} \quad v_i^\star(\delta) = -\frac{c_i}{\sigma_i - \sigma_1} \text{ if } i > k.$$

Then the continuity assertion is a consequence of $\mu^\star(\delta) \searrow -\sigma_1$ as $\delta \nearrow \delta_{\text{crit}}$, which in turn follows from strict monotonicity and continuity of φ on $[-\sigma_1, +\infty[$. Finally, we obtain via $\varphi'(-\sigma_1) < 0$ and the Implicit Function Theorem that the left-hand side derivative of $\mu^\star(\delta)$ at δ_{crit} satisfies

$$\frac{d\mu^\star}{d\delta}(\delta_{\text{crit}}) = -\frac{\frac{\partial F}{\partial \delta}(-\sigma_1, \delta_{\text{crit}})}{\frac{\partial F}{\partial t}(-\sigma_1, \delta_{\text{crit}})} = -\frac{-1}{\varphi'(-\sigma_1)} < 0,$$

so that there must be a kink at δ_{crit} for $\mu^\star(\delta)$. $\qquad\square$

Here we should stress that it is not suggested to employ the above explicit construction in numerical application. It serves only for establishing the theoretical result. Only in the highly unlikely case of $b \in \{u_1, \ldots, u_k\}^\perp$ one might be forced to return to the above (or similar) constructions. Observe that for testing the latter conditions, we need only information on the eigenspace for the smallest eigenvalue rather than complete spectral decomposition of A.

Finally, we obtain an important monotonicity result for the multiplier:

COROLLARY 9 *The Lagrange multiplier $\mu^\star(\delta)$ is (can be chosen as) decreasing in δ.*
If $b \in \{u_1, \ldots, u_k\}^\perp$, then $\mu^\star(\delta)$ strictly decreases on the interval $[0, \delta_{\mathrm{crit}}]$ (and is constant for larger δ), while if $b \notin \{u_1, \ldots, u_k\}^\perp$, then $\mu^\star(\delta)$ decreases strictly on the whole $I\!\!R_+$.

3. The algorithm

Once we have established continuous dependence of the global solution $x^\star(\delta)$ upon the radius, we can proceed as follows:

3.1. algorithm ebh

1. Start with $\delta_0 = n$ (then the ball is containing the box B), determine $\mu_0 = \mu^\star(\delta_0)$ and $x_0 = x^\star(\delta_0)$ by solving the corresponding trust region problem.
2. If $x_0 \in B$, stop and return x_0 that is the exact solution of the BCQP; else determine the multiplier $\mu_\infty = \mu^\star(1)$ for the ball completely contained in B and put $k = 1$.
3. Calculate

$$\mu_k = (\mu_0 + \mu_\infty)/2,$$
$$x_k = -(A + 2\mu_k I_n)^{-1} b,$$

   (then for $\delta_k = x_k^\top x_k$ we have $x_k = x^\star(\delta_k)$ and $\mu_k = \mu^\star(\delta_k)$, according to (2.8) and Theorem 5).
4. If $x_k \in B$ and near the boundary ∂B of B, stop and return x_k as the promising point (to be used as starting point for any local solver); else, if $x_k \in B$ then put $\mu_\infty = \mu_k$ else put $\mu_0 = \mu_k$.
5. Replace k with $k + 1$ and repeat from 3.

From Corollary 9 and Theorem 8 we can be sure that the above procedure is finite. The next step is a branch-and-bound approach with the help of the obtained promising point x_k. We cut the box B in two halves by bisecting along that coordinate direction i which is most promising in the sense that the minimal deviation of the i-th coordinate of x_k from ± 1 is largest. Repeating this recursively, we generate

2^m subboxes by m such cuts. Of course, subboxes for which the minimum of $f(x)$ on the including ball is larger than the best current solution can be discarded from further analysis. This cannot be used as an exact algorithm, because of the prohibitive computational cost that it would require, so a heuristic stopping criterion must be used. In our computational experiments we used as a stopping criterion a prefixed number `maxit` of cuts.

4. Computational results

A prototype `Matlab` version of algorithm EBH was tested on a set of test problems. For the minimization of $f(x)$ over a ball, we used the code `trustpen` [13]. As a local solver for the box-constrained quadratic problem solved we used the `Matlab` routine `quadprog` [15]. As already pointed out, for a prefixed maximum number `maxit` of cuts, at most 2^{maxit} problems of the form (2.2) had to be solved.

4.1. TEST PROBLEMS WITH KNOWN GLOBAL SOLUTION

We first try test problems with known global, generated as in [10], moving from the sufficient global optimality condition originally proposed by Neumaier [16]. However, we found such problems very easy, in the sense that the `Matlab` local solver was almost always able to find the global solution.

Therefore for constructing test problems, we used the sufficient global optimality conditions recently described by Beck and Teboulle [2] for quadratic optimization problems with binary constraints, that can be easily rephrased for concave box-constrained quadratic problems.

THEOREM 10 *Consider the problem (1.1), with A negative definite. Then x is a global solution for (1.1) if x is feasible and*

$$\lambda_{max}(A)e \geqslant XAXe + Xb \tag{4.9}$$

where $e = [1, \ldots, 1]^\top \in I\!R^n$, $X = diag(x_i)$, *and* $\ll_{max}(A)$ *is the largest eigenvalue of the matrix A.*

Based on Theorem 10, concave test problems were generated with $A = HDH$, with D a diagonal matrix whose elements are randomly generated from the interval $]-2, -1[$; H a Householder matrix of a random vector; the prefixed solution x randomly generated with components in $\{-1, +1\}$, and the vector b is computed so that the condition (4.9) holds. Test problems were generated with $n = 50, 100, 200$; for each dimension 10 different problems were generated. The value of `maxit` was set equal to 15 in our experiments.

From the computational results (Table I), we see that for such problems, the ellipsoidal approach show to be very effective, since only for one instance of such problems the global solution was not found; to be fair, we note that problems

Table I. Failure
rates for EBH

n	Failure
50	0
100	1
200	0

with known solution are usually not very difficult, because of the sufficient global optimality conditions they are based on. So are the problems we generated, for which the linear term is very strong. Nevertheless, to check that the problems are not trivial, for each of them we run the linear solver $10n$ times, with randomly generated starting points; we found that for none of them, the stationary point is unique. For example, for the problems of size 100, we found for each of them 25 stationary points on the average. Also, for the local solvers, the problems shown to be harder than the ones we generated as in [10].

4.2. TEST PROBLEMS WITH UNKNOWN GLOBAL SOLUTION

Next, we generated a first set of random test problems with unknown solution. The matrix A has the form $A = HDH$, where H is a Householder matrix of a random vector and $D = \text{diag}(d)$, with d a random vector such that $|d_i| \in]1, 2[$. For each matrix we generate five different problems, by taking $b = d/10^p$, with $p = 0, 1, 2, 3, 4, 5$. For $p = 5$ we have problems in which the quadratic part of the objective function dominates, and this is the opposite which happens for $p = 0$. Several problems were generated varying the size n and the percentage *Neg* of negative eigenvalues of the Hessian matrix. To check the quality of the solution that was found with the ellipsoidal approach, for each problem we run n times the two local solvers with randomly generated starting points, and then we compared the best solution (random solution), to the one we found using the ellipsoidal approach. As expected, the number of stationary points for the test problems with unknown global solution is much higher than for the test problems with known global solution. For the problems with 100 variables, in finding the random solution, 25 different stationary points were found on the average.

From Table II, it is observed that the ellipsoid algorithm was able to find the best (known) solution in 24 out of the 30 test problems; actually, for 13 problems the solution that the algorithm was able to detect, was better than the one which was found by the random approach. We also point out that for the six problems in which the algorithm failed, just doubling the maximum number of iterations `maxit`, the EBH algorithm was eventually able to get the best solution in four cases. The number of failures (i.e., the problems in which the random solution was better than the one found by the algorithm) is very low; although there is of course no guarantee that the solution the algorithm was able to find is the global one, even the suboptimal solutions found were of considerable high quality (measured by

Table II. Performance of EBH on random test problems (See text)

n	*Neg*	Failures
50	25	1
50	50	1
50	75	1
100	25	0
100	50	1
100	75	2

the fraction of EBH-result over the maximum over all random solutions found). It seems that the effectiveness of the algorithm does not deteriorate when the number of negative eigenvalues increases (since the number of failures increases just slightly). On the other hand, for $Neg = 75$ and $Neg = 50$, the ellipsoidal approach beats the random approach in 11 problems out of 20; in the problems with $Neg = 25$, the two approaches almost always get the same solution.

4.3. MAXIMUM CLIQUE PROBLEMS

The maximum clique problem is a classical problem in graph theory which can be formulated as follows. Let $G = G(V, E)$, with vertex set $V = \{1, 2, \ldots, n\}$ and edge set $E \subseteq V \times V$, be an undirected graph. The maximum clique problem is the problem of finding a complete subgraph of G of maximal cardinality. This problem has many equivalent formulations as an integer problem, or a continuous nonconvex optimization problem, for a recent survey see [4]. Here we follow a recent continuous reformulation approach from [1]: starting from the following quadratic zero-one reformulation

$$\min\{f(x) = x^\top A x - e^\top x: \quad \text{s.t.} \quad x \in \{0, 1\}^n\}, \tag{4.10}$$

where A is the adjacency matrix of the complement graph of G and e is the vector of all ones, consider the continuous relaxation problem

$$\min\{f(x) = x^\top A x - e^\top x: \quad \text{s.t.} \quad x \in [0, 1]^n\}. \tag{4.11}$$

Since the matrix A has zeroes on the diagonal, this problem has at least one 0–1 solution, then the global solution of (4.10) is equal to (one of the) global solution(s) of (4.11), and therefore a maximum clique can be computed by solving the problem (4.11).

The EBH algorithm was then tested on problem (4.11) generated by a set of some benchmark graph instances from the DIMACS challenge [9]. We selected only graphs with less than 400 vertices. The results are shown in Table 3. On purpose, we did not include the results of the similar reformulation approach in [1], because not all instances in Table III were considered there. In short, the authors

Table III. Maximum clique problems from the DIMACS challenge (see text)

Problem	n	Dens. (%)	Clique size	EBH (cuts)	CBH	SR	AR
c-fat200-1	200	07.7	12	12 (1)	12	12	12
c-fat200-2	200	16.3	24	24 (1)	24	24	24
c-fat200-5	200	42.6	58	58 (1)	58	58	58
hamming6-2	64	90.5	32	32 (1)	32	32	32
hamming6-4	64	34.9	4	4 (1)	4	4	4
hamming8-2	256	96.9	128	128 (1)	128	128	128
hamming8-4	256	63.9	16	16 (1)	16	16	16
johnson8-2-4	28	55.6	4	4 (1)	4	4	4
johnson8-4-4	70	76.8	14	14(1)	14	14	14
johnson16-2-4	120	76.5	8	8 (1)	8	8	8
p_hat300-1	300	24.4	8	8 (2)	8	6	8
p_hat300-2	300	48.9	25	24 (3)	25	22	25
p_hat300-3	300	74.5	36	32 (1)	36	32	35
MANN_a9	45	92.7	16	16 (1)	16	12	16
MANN_a27	378	99.0	126	125 (4)	121	117	117
keller4	171	64.9	11	9 (8)	10	7	8
brock200_1	200	74.5	21	20 (1)	20	17	19
brock200_2	200	49.6	12	10 (1)	12	8	10
brock200_3	200	60.5	15	11 (1)	14	9	13
brock200_4	200	65.8	17	15 (5)	16	12	14
san200_0.7_1	200	70.0	30	15 (3)	15	15	15
san200_0.7_2	200	70.0	18	12 (1)	12	12	12
san200_0.9_1	200	90.0	70	46 (2)	46	45	45
san200_0.9_2	200	90.0	60	40 (10)	36	36	39
san200_0.9_3	200	90.0	44	34 (9)	30	32	31
sanr200_0.7	200	69.7	18	16 (7)	18	14	16
sanr200_0.9	200	89.8	≥ 42	35 (3)	41	37	41

report in [1] the same figures as in Table III for the first 10 lines. Their algorithm has equal performance as EBH on the Mannino graphs MANN_ax ($x \in \{9, 27\}$) and the Sanchis graphs san200_0.9_x ($x \in \{2, 3\}$), beats EBH by one node for san200_0.9_1, and hits the maximum clique in the Keller graph keller4. All other instances were unfortunately not covered in [1]. The columns labelled by CBH, SR, and AR, refer to different but related continuous-based heuristics for the maximum clique problem. Essentially, all these are based upon the Motzkin–Strauß program which reformulates the Maximum Clique Problem as a Standard Quadratic Optimization Problem, i.e. to find the maximum of a quadratic form over the standard simplex (rather than over a box). Here, the matrix generating the quadratic form is the adjacency matrix of the graph.

The first method considered is the Continuous Based Heuristic (CBH) [8], which employs a parameterized version of the original Motzkin–Strauß program. The problem is divided into a series of subproblems with the simplex constraints relaxed into spherical ones (this is a related feature to our approach). A combinatorial post-processing phase is needed to round the solutions produced by the procedure that solves the subproblems back to the standard simplex.

The second procedure is the Simple Replicator Dynamics approach (SR) [5] which basically finds local solution to a regularized version of the Motzkin-Strauß program (adjacency matrix plus 1/2 times the identity matrix) with the help of replicator dynamics, a discrete time (non-Euclidean gradient) dynamical system operating on the standard simplex (thus feasibility is guaranteed throughout the algorithm).

The third algorithm is Annealed Replication (AR) [3]. It uses a parameterized quadratic form emerging from perturbing similar as in SR the adjacency matrix (here plus α times the identity matrix), and iteratively searches for local solutions via replicator dynamics. In order to escape inefficient local solutions, thereafter the parameter α is varied as in simulated annealing, but motivated by more principled arguments.

5. Conclusions

Moving from some new theoretical results concerning the global minimization of a quadratic function over a ball, and more specifically the continuous dependence of the Lagrange multiplier for the global solution on the radius of the ball, we proposed an heuristic algorithm for the global solution of box constrained quadratic programming problems. The algorithm, embedded in a branching strategy, showed to be successful in many of the test problems we used in our computational experiences. In particular, our approach gave reasonable good results on a wide set of some benchmark graph instances from the DIMACS challenge, although no special care was taken to fit the algorithm to the maximum clique problem. We therefore believe that the proposed algorithm can be fruitfully used, at least, as an excellent procedure to give a very good starting point to any local solver.

Acknowledgement

References

1. J. Abello, S. Butenko, P.M. Pardalos and M.G.C. Resende (2001), Finding independent sets in a graph using continuous multivariable polynomial formulations. *J. Global Optimization*. 21, 111–137.
2. A. Beck and M. Teboulle (2000), Global optimality conditions for quadratic optimization problems with binary constraints. *SIAM J. Optim.* 11, 179–188.
3. I.M. Bomze, M. Budinich, M. Pelillo and C. Rossi (2001), Annealed replication: A new heuristic for the maximum clique problem, to appear in: *Discrete Appl. Math.*
4. I.M. Bomze, M. Budinich, P.M. Pardalos and M. Pelillo (1999), The maximum clique problem. In: D.Z. Du and P.M. Pardalos (Eds), Handbook of Combinatorial Optimization, suppl. Vol. A, 1–74. Kluwer Academic Publishers, Dordrecht.

5. I.M. Bomze, M. Pelillo and R. Giacomini (1997), Evolutionary approach to the maximum clique problem: Empirical evidence on a larger scale. In: I. M. Bomze, T. Csendes, R. Horst and P. M. Pardalos (Eds.), *Developments in Global Optimization*, pp. 95–108. Kluwer Academic Publishers, Dordrecht.

6. G. Danninger and I.M. Bomze (1994), Generalized convexity for second-order optimality conditions. In: S. Komlósi, T. Rapcsák, S. Schaible (Eds.), *Generalized Convexity,* pp. 137–144. Springer, Berlin.

7. P.L. De Angelis, P.M. Pardalos and G. Toraldo (1997), Quadratic Programming with box constraints. In: I.M.Bomze, T. Csendes, R. Horst and P.M. Pardalos (Eds.), *Developments in Global Optimization*, pp. 73–93. Kluwer Academic Publishers, Dordrecht/Boston/London.

8. L. E. Gibbons, D. W. Hearn and P. M. Pardalos (1996), A continuous based heuristic for the maximum clique problem. In: D. S. Johnson and M. Trick (Eds.), *Cliques, Coloring, and Satisfiability—Second DIMACS Implementation Challenge*, pp. 103–124. American Mathematical Society, Providence, RI.

9. J. Hasselberg, P.M. Pardalos and G. Vairaktarakis (1993), Test case generators and computational results for the maximum clique problem, *Journal of Global Optimization* 3, 463–482.

10. C.G. Han, P.M. Pardalos and Y. Ye (1992), On the solution of indefinite quadratic problems using an interior point algorithm, *Informatica* 3, 474–496.

11. J.B. Hiriart-Urruty (1995), Conditions for global optimality. In: R. Horst and P. M. Pardalos (Eds.), *Handbook of Global Optimization*, pp. 1–26. Kluwer Academic Publishers, Dordrecht.

12. M.K. Kozlov, S.P. Tarasov and L.G. Khacian (1979), Polynomial solvability of convex quadratic programming, *Soviet. Math. Dokl.* 20, 1108–1111.

13. S. Lucidi, L. Palagi and M. Roma (1998), On some properties of quadratic programs with a convex quadratic constraint, *SIAM J. Optim.* 8, 105–122.

14. J.M. Martínez (1994), Local minimizers of quadratic functions on Euclidean balls and spheres, *SIAM J. Optim.* 4, 159–176.

15. Matlab - Optimization Tolbox 2.1 (2001), The MathWorks, Inc.

16. A. Neumaier (1992), An optimality criterion for global quadratic optimization, *J. Global Optimization* 2, 201–208.

17. J. Moré and D.C. Sorensen (1983), Computing a trust region step, *SIAM J. Sci. Stat. Comput.* 4, 553–572.

18. P.M. Pardalos and G.P. Rodgers (1990), Computational aspects of branch and bound algorithms for quadratic zero-one programming, *Computing* 45, 131-144.

19. R.J. Stern and H. Wolkowicz (1994), Trust region problems and nonsymmetric eigenvalue perturbations, *SIAM J. Matrix Anal. Appl.* 15, 755–778.

20. Y. Yajima, M.V. Ramana and P.M. Pardalos (2001), Cuts and Semidefinite Relaxations for Nonconvex Quadratic Problems, In: N. Hadjisavvas, J.E. Martinez-Legaz and J.P. Penot (Eds.), *'Generalized Convexity and Generalized Monotonicity,' Lecture Notes in Economics and Mathematical Systems*, Vol. 502, pp. 48–70. Springer, Berlin.