Journal of Global Optimization **22:** 17–37, 2002.

Branch-and-bound approaches to standard quadratic optimization problems

IMMANUEL M. BOMZE
Department of Statistics and Decision Support Systems, University of Vienna, Vienna, Austria
(E-mail: immanuel.bomze@univie.ac.at)

Abstract. This paper explores several possibilities for applying branch-and-bound techniques to a central problem class in quadratic programming, the so-called Standard Quadratic Problems (StQPs), which consist of finding a (global) minimizer of a quadratic form over the standard simplex. Since a crucial part of the procedures is based on efficient local optimization, different procedures to obtain local solutions are discussed, and a new class of ascent directions is proposed, for which a convergence result is established. Main emphasis is laid upon a d.c.-based branch-and-bound algorithm, and various strategies for obtaining an efficient d.c. decomposition are discussed.

Key words: D.C. decomposition, Semidefinite relaxation, Replicator dynamics

1. Introduction

This paper explores several possibilities for applying branch-and-bound techniques to a central problem class in quadratic programming, the so-called Standard Quadratic Problems (StQPs). A StQP consists of finding a (global) maximizer of a quadratic form over the standard simplex. Problems of such type occur frequently as subproblems in escape procedures for general quadratic optimization, but also have manifold direct applications, e.g., in portfolio selection and in the maximum weight clique problem for undirected graphs. As such, StQPs are an important instance for successful application of continuous-based techniques in combinatorial optimization.

Formally speaking, we consider (global) optimization problems of the form

$$\max\{f(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} : \mathbf{x} \in \Delta\}\,, \tag{1}$$

where A is an arbitrary symmetric $n \times n$ matrix; the sign $^\top$ denotes transposition; and Δ is the standard simplex in n-dimensional Euclidean space $\mathbb{R}^n$,

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n : x_i \geqslant 0 \text{ for all } i \in \{1, \ldots, n\}\,,\ \mathbf{e}^\top\mathbf{x} = 1\}\,,$$

where $\mathbf{e} = \sum_{i=1}^n \mathbf{e}_i = [1, \ldots, 1]^\top$ while $\mathbf{e}_i$ denotes the i-th standard basis vector in $\mathbb{R}^n$. More generally, the question of finding maximizers of a inhomogeneous quadratic function $\mathbf{x}^\top Q\mathbf{x} + 2\mathbf{c}^\top\mathbf{x}$ over Δ can be homogenized by considering the

Dedicated to Reiner Horst on occasion of his 60th birthday.

rank-two update $A = Q + \mathbf{e}\mathbf{c}^\top + \mathbf{c}\mathbf{e}^\top$ in (1) which has the same objective values. Finally, if a general, possibly inhomogeneous quadratic function as above is considered over a polytope $P = \mathrm{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_q)$, then every such quadratic problem can be rewritten as a StQP (of course, this is only recommended if q is small and full vertex enumeration is easy). See Bomze (1998) for more practical arguments in favour of StQPs.

Of course, quadratic optimization problems like (1) — even regarding the detection of local solutions — are NP-hard (Horst et al., 1995). Nevertheless, there are several exact procedures which try to exploit favourable data constellations in a systematic way, and to avoid the worst-case behaviour whenever possible.

The article is organized as follows: we start with a short review on traditional local solution strategies in Section 2, and present in Section 3 a new class of local ascent methods for StQPs. We want to stress that the popular gradient-projection method is by far not the dominating one, and propose to use different, gradient-like methods having much nicer features. Efficiency of these methods is important in our context as we also need a locally optimizing procedure for obtaining upper bounds in the d.c. approach proposed below. Note that for the problem class considered, feasibility never is a problem, so that lower bounds are readily available, although they can be considerably improved by employing local search.

The following Section 4 reports shortly on algorithms which use strategies differing from the branch-and-bound approach. Escape strategies like annealed replication, the G.E.N.F. algorithm, copositive programming, and the $\mathcal{R}$-strategy are mentioned for the sake of completeness, the latter in particular because it also relies on a d.c. decomposition of the quadratic objective, which plays a central role in the following Section 5. Here, a convergence result for general d.c.-decomposition strategies is (re-)stated, and several variants of these are discussed. On the basis of a recently developed local optimality criterion (Dür, 1999), and also for practical reasons, we here propose to favour the spectral decomposition and prove that in some sense, this decomposition is minimal. However, we also specify an example which shows that the hope for a globally minimal d.c. decomposition is in vain. Finally, we suggest SDP-based criteria for selecting an appropriate d.c. decomposition.

2. Traditional approaches to solve StQPs: local optimization

2.1. GRADIENT PROJECTION FOR STQPS

There are several local optimization strategies for solving StQPs. A more or less straight-forward projected gradient method is used, e.g., in (Nowak, 1999) as a subroutine for obtaining an improved upper bound based on a better concave overestimator than the usual LP-bounds. Note that we need concave overestimators rather than convex underestimators because we maximize the objective function.

While the gradient of the objective $\frac{1}{2}\mathbf{x}^\top A\mathbf{x}$ is cheaply computed to $A\mathbf{x}$, the difficulty with this method is to obtain a feasible direction via projection onto the

feasible set Δ. An exact but involved algorithm for the (non-linear) least-distance projection can be found, e.g., in (Cegielski, 1993), and references therein. A classical alternative is Rosen's gradient projection method, which (ortho-)projects linearly on supporting hyperplanes (or a subspace thereof of one dimension less, if necessary) of the feasible polyhedron.

To be more precise, denote by I_k the $k \times k$ identity matrix and by E_k the $k \times k$ matrix with unit entries only. Now suppose that we start at a feasible point $\mathbf{x} \in \Delta$ with $x_i > 0$ if and only if $1 \leqslant i \leqslant k$ for some $k \in \{1, \ldots, n\}$. Then the above-mentioned orthoprojection is given by the $n \times n$ matrix $P_{\mathbf{x}}$ which contains $I_k - \frac{1}{k} E_k$ as its leading (upper left) principal $k \times k$ submatrix, and has zero entries elsewhere. Thus, the search direction $\mathbf{w}(\mathbf{x}) = P_{\mathbf{x}} A\mathbf{x}$ satisfies

$$\begin{array}{ll} w_i(\mathbf{x}) = & [A\mathbf{x}]_i - \frac{1}{k} \sum_{j=1}^{k} [A\mathbf{x}]_j \,, \quad \text{if } 1 \leqslant i \leqslant k \,, \\ w_i(\mathbf{x}) = & 0 \qquad\qquad\qquad\qquad\quad \text{else.} \end{array} \tag{2}$$

It is readily seen that $\mathbf{w}(\mathbf{x}) = \mathbf{o}$ if and only if $\mathbf{x}$ is a *generalized* Karush/Kuhn/Tucker (KKT) point, i.e., a critical point for the Lagrange function

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_{i=1}^{n} \lambda_i x_i + \lambda_{n+1}(\mathbf{e}^\top \mathbf{x} - 1)$$

where no sign restrictions on the multiplicators λ_i are imposed. Observe that generalized KKT points are exactly the stationary points under the replicator dynamics (10) treated later in this section. There is a resolution strategy for the case $P_{\mathbf{x}} A\mathbf{x} = \mathbf{o}$: if $[A\mathbf{x}]_i \leqslant \frac{1}{k} \sum_{j=1}^{k} [A\mathbf{x}]_j$ for all $i \in \{k+1, \ldots, n\}$ (this includes the case $k = n$), then $\mathbf{x}$ is a usual KKT point (i.e. with the classical dual constraints $\lambda_i \geqslant 0$ for all $i \in \{1, \ldots, n\}$). Otherwise, ignore one of the binding constraints (say, $x_{k+1} = 0$) such that $[A\mathbf{x}]_{k+1} > \frac{1}{k} \sum_{j=1}^{k} [A\mathbf{x}]_j$, and form the resulting projection matrix $\widehat{P}_{\mathbf{x}}$ which now contains $I_{k+1} - \frac{1}{k+1} E_{k+1}$ as its leading (upper left) principal $(k+1) \times (k+1)$ submatrix, and zeroes elsewhere. Then $\widehat{\mathbf{w}}(\mathbf{x}) = \widehat{P}_{\mathbf{x}} A\mathbf{x}$ is given by (always provided that $P_{\mathbf{x}} A\mathbf{x} = \mathbf{o}$):

$$\begin{array}{ll} \widehat{w}_i(\mathbf{x}) & = [A\mathbf{x}]_i - \frac{1}{k+1} \sum_{j=1}^{k+1} [A\mathbf{x}]_j \\ & = \frac{1}{k+1} \Big(\frac{1}{k} \sum_{j=1}^{k} [A\mathbf{x}]_j - [A\mathbf{x}]_{k+1} \Big) \,, \qquad \text{if } 1 \leqslant i \leqslant k \,, \\ \widehat{w}_i(\mathbf{x}) & = \frac{k}{k+1} \Big([A\mathbf{x}]_{k+1} - \frac{1}{k} \sum_{j=1}^{k} [A\mathbf{x}]_j \Big) \,, \qquad \text{if } i = k+1, \\ \widehat{w}_i(\mathbf{x}) & = 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{else.} \end{array} \tag{3}$$

The complete procedure thus reads as follows: for any $\mathbf{x} \in \Delta$, let

$$\begin{array}{l} \mathbf{v}(\mathbf{x}) = \mathbf{w}(\mathbf{x}) \,, \text{ with } \mathbf{w}(\mathbf{x}) \text{ as in (2), if } \mathbf{w}(\mathbf{x}) \neq \mathbf{o} \,, \\ \mathbf{v}(\mathbf{x}) = \widehat{\mathbf{w}}(\mathbf{x}) \,, \text{ with } \widehat{\mathbf{w}}(\mathbf{x}) \text{ as in (3), else.} \end{array} \tag{4}$$

Then, according to Theorem 10.3.4 of (Bazaraa and Shetty, 1979), $\mathbf{v}(\mathbf{x})$ represents an ascent direction (in particular $\mathbf{v}(\mathbf{x}) \neq \mathbf{o}$) if and only if $\mathbf{x}$ is no KKT point (in the traditional sense). In this paper we use 'ascent direction' synonymous for 'improving feasible direction', which means (a) that $\mathbf{x} + t\mathbf{v}(\mathbf{x}) \in \Delta$ for all $t > 0$ small enough; and (b) that

$$\psi(t) = f\Big(\mathbf{x} + t\mathbf{v}(\mathbf{x})\Big) = \mathbf{x}^\top A\mathbf{x} + 2t\,\mathbf{x}^\top A\mathbf{v}(\mathbf{x}) + t^2\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) \tag{5}$$

satisfies $\psi(t) > \psi(0) = \mathbf{x}^\top A\mathbf{x}$ for all $t > 0$ small enough. Feasibility of $\mathbf{x} + t\mathbf{v}(\mathbf{x})$ for those t means that

$$\overline{t}(\mathbf{x}) = \min_{1 \leqslant i \leqslant n} \left\{ \frac{x_i}{-v_i(\mathbf{x})} : v_i(\mathbf{x}) < 0 \right\} . \tag{6}$$

is always strictly positive, provided that $\mathbf{x} \in \Delta$ is no KKT point. Finally, denote by $t(\mathbf{x})$ the maximizer of $\psi(t)$ over the feasible interval $[0, \overline{t}(\mathbf{x})]$. Then

$$\begin{array}{ll} t(\mathbf{x}) & = \min\left\{\overline{t}(\mathbf{x}), \frac{\mathbf{x}^\top A\mathbf{v}(\mathbf{x})}{-\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x})}\right\} , \quad \text{if } \mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) < 0 , \\ t(\mathbf{x}) & = \overline{t}(\mathbf{x}) , \qquad\qquad\qquad\qquad\quad \text{else.} \end{array} \tag{7}$$

Observe that the above case distinction is made according to whether ψ is strictly concave or convex. Now generate a (possibly finite) sequence $(\mathbf{x}_\nu)$ of feasible points as follows: take an arbitrary starting point $\mathbf{x}_0 \in \Delta$ and put

$$\mathbf{x}_{\nu+1} = \mathbf{x}_\nu + t(\mathbf{x}_\nu)\mathbf{v}(\mathbf{x}_\nu) , \quad \text{all } \nu \in \mathbb{N}_0 . \tag{8}$$

Then by definition, $\mathbf{x}_\nu \in \Delta$ and $f(\mathbf{x}_{\nu+1}) \geqslant f(\mathbf{x}_\nu)$ for all $\nu \in \mathbb{N}$.

Remark 1. By compactness of Δ and continuity of f, we have convergence of the objective value sequence $f(\mathbf{x}_\nu)$ regardless whether or not the sequence of points $(\mathbf{x}_\nu)$ converges.

A notorious obstacle for gradient projection methods is that the map $\mathbf{x} \mapsto \mathbf{v}(\mathbf{x})$ fails to be continuous. As a consequence, no convergence results seem to be available; see (Bazaraa and Shetty, 1979), pp. 398f.

2.2. REDUCED GRADIENTS

A familiar alternative is Wolfe's reduced gradient method. Applied to our StQP, for any non-stationary point $\mathbf{x} \in \Delta$ with largest coordinate $\mathbf{x}_m > 0$, say, this yields

$$\begin{array}{ll} v_i(\mathbf{x}) & = [A\mathbf{x}]_i - [A\mathbf{x}]_m , \qquad \text{if } [A\mathbf{x}]_m \leqslant [A\mathbf{x}]_i \text{ and } i \neq m , \\ v_i(\mathbf{x}) & = x_i\Big([A\mathbf{x}]_i - [A\mathbf{x}]_m\Big) , \quad \text{if } [A\mathbf{x}]_m > [A\mathbf{x}]_i , \end{array} \tag{9}$$

and $v_m(\mathbf{x}) = -\sum_{i \neq m} v_i(\mathbf{x})$. Then $\mathbf{v}(\mathbf{x}) \in \mathbf{e}^\perp$ and $\mathbf{x} + t\mathbf{v}(\mathbf{x}) \in \Delta$ if $t > 0$ is small enough, i.e., if $0 \leqslant t \leqslant \overline{t}(\mathbf{x})$ with $\overline{t}(\mathbf{x})$ as in (6) with $\mathbf{v}(\mathbf{x})$ as in (9). Note that,

again, we have $\bar{t}(\mathbf{x}) > 0$. Further, $\mathbf{v}(\mathbf{x}) \neq \mathbf{o}$ if and only if $\mathbf{x}$ is no KKT point, and $\mathbf{v}(\mathbf{x})$ is an ascent direction. Finally, if the sequence $(\mathbf{x}_\nu)$ is generated as in (8) with $\mathbf{v}(\mathbf{x})$ as in (9), then, contrasting to the situation in Subsection 2.1, one can prove that every accumulation point of the sequence $(\mathbf{x}_\nu)$ is a KKT point, according to Theorem 10.4.3 in (Bazaraa and Shetty, 1979) (note the assumptions there are satisfied for our StQP).

However, convergence of the whole sequence $(\mathbf{x}_\nu)$ is, in general, not guaranteed (which in practice admittedly poses little problems). In the following section we present a new class of ascent methods with the same properties, which has the additional advantage of arbitrary high degree of smoothness. In a nutshell, we propose to compare the coordinates $[A\mathbf{x}]_i$ of the gradient with the average $\mathbf{x}^\top A\mathbf{x} = \sum_i x_i [A\mathbf{x}]_i$ rather than with one fixed coordinate $[A\mathbf{x}]_m$.

2.3. REPLICATOR DYNAMICS AND COMPLEMENTARY PIVOTING FOR LOCAL SOLUTIONS TO STQPS

A different method for local solution of StQPs, which recently became increasingly popular, employs a dynamical system which plays the role of a gradient system (and, in fact, is one, albeit not w.r.t. Euclidean geometry but the Shahshahani metric rather): the so-called replicator dynamics. These dynamics have the advantage that they converge, contrasting to all other local optimization methods, without any further qualifications on the instance. Even in the generically rare case of halting at a non-optimal point, a single perturbation immediately yields an improving local solution. By concept (only), this is similar to the resolution strategy in (3). Again, the key idea here is comparison of $[A\mathbf{x}]_i$ with $\mathbf{x}^\top A\mathbf{x}$, and a corresponding update. Observe that the maximizers of (1) remain the same if A is replaced with $A + \gamma E_n$ (note that the curvature of $f(\mathbf{x})$ *over* Δ remains unaffected under this transition), where γ is an arbitrary constant, say $\gamma = \max_{i,j} a_{ij} + 1$, so without loss of generality we can assume that all entries of A are positive, ensuring $\mathbf{x}^\top A\mathbf{x} > 0$ for all $\mathbf{x} \in \Delta$. Then the following dynamics is well-defined and leaves the simplex Δ invariant:

$$x_i(\tau + 1) = x_i(\tau) \frac{[A\mathbf{x}(\tau)]_i}{\mathbf{x}(\tau)^\top A\mathbf{x}(\tau)}, \quad i \in \{1, \ldots, n\}, \text{ all } \tau \in \mathbb{N}_0. \tag{10}$$

The stationary points $\mathbf{x}$ under (10) coincide with those $\mathbf{x} \in \Delta$ satisfying $P_{\mathbf{x}} A\mathbf{x} = \mathbf{o}$ from (2), and all local solutions to the StQP (1) are among these. Of course, there are quite many stationary points, e.g., all vertices $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of Δ. However, only those $\mathbf{x}$ are serious candidates for strict local solutions which are asymptotically stable, which means that every solution to (10) which starts close enough to $\mathbf{x}$, will converge to $\mathbf{x}$ as $\tau \nearrow \infty$. For proofs, more details and background, see Bomze (1998) and the references therein. Here let us stress that, irrespective of the curvature of the objective function, replicator dynamics yield, with probability one regarding the choice of a starting point, in the long run *local solutions* rather than

merely KKT points[1] (like all other local optimization procedures covered here), provided a simple condition on A (namely, that no principal minor vanishes) is satisfied (Bomze and Stix, 1999), which is generically true. Interestingly enough, a similar genericity result is used in Section 3 below, there to ensure convergence of the whole sequence $(\mathbf{x}_\nu)$ rather than of a subsequence.

Another attack rephrases the first-order conditions for local optimality into a linear complementarity problem and solves this with pivoting methods à la Lemke. Surprisingly enough, the points obtained were always (good) local solutions, at least for the special case of the maximum weighted clique problem. Decisive for this success seems to be an intricate degeneracy resolution strategy, combined with a one-step look ahead technique, which applies also to the general StQP (1), without exploiting the discrete structure information of the (combinatorial) graph theoretic background. Note that theory would predict KKT points (only) as the outcome of the algorithm rather than local solutions. As a detailed description of this procedure would be beyond the scope of this paper, let us refer to (Massaro et al., 2001) for more detail.

3. A new class of ascent methods for StQPs, and a general convergence result

The class of ascent directions we now propose can be viewed as a compromise between the (non-smooth) Wolfe-direction in (9) and the smooth replicator dynamics (10). It also involves the gradient $A\mathbf{x}$, but in a way inspired to some extent by the notion of payoff-monotone dynamics in evolutionary game theory (Weibull, 1995). The formula we employ is much older than these recent developments, and can be traced back to the method of existence proof for Nash equilibria via Brouwer's fixed point theorem.

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an arbitrary function satisfying $\varphi(t) = 0$ if $t \leqslant 0$ but $\varphi(t) > 0$ if $t > 0$. Examples are abundant: from the truncated sign function $\varphi(t) = [\operatorname{sign} t]_+$ (which is elementary but discontinuous) through $\varphi(t) = [t_+]^\beta$ (with a degree of smoothness which increases with $\beta > 0$) to the C^∞-smooth function $\varphi(t) = \exp(-\frac{1}{t})$ if $t > 0$, and $\varphi(t) = 0$ if $t \leqslant 0$. Here and in the sequel, we denote, for a real number t, by $t_+ = \max\{t, 0\}$ and by $t_- = \max\{-t, 0\}$ so that $t = t_+ - t_-$ and $t_+ t_- = 0$.

Now denote by $r_i(\mathbf{x}) = [A\mathbf{x}]_i - \mathbf{x}^\top A\mathbf{x}$, all $i \in \{1, \ldots, n\}$ and by

$$v_i(\mathbf{x}) = \varphi\Big(r_i(\mathbf{x})\Big) - \sigma(\mathbf{x}) x_i \,, \quad i \in \{1, \ldots, n\} \,, \tag{11}$$

with the convention

$$\sigma(\mathbf{x}) = \sum_{j=1}^{n} \varphi\left(r_j(\mathbf{x})\right) \,.$$

[1] Of course, in case of concave maximization this distinction vanishes as all KKT points are global solutions then.

The latter enforces $\mathbf{v}(\mathbf{x}) \in \mathbf{e}^{\perp}$ for all $x \in \Delta$. Further, defining $\overline{t}(\mathbf{x})$ as in (6) with the new definition of $\mathbf{v}(\mathbf{x})$ from (11), we see that $\mathbf{v}(\mathbf{x})$ is an ascent direction unless $\mathbf{x}$ is a KKT point:

THEOREM 1. *Let $\mathbf{x} \in \Delta$. For $\mathbf{v}(\mathbf{x})$ as in (11), we have $\mathbf{v}(\mathbf{x}) \neq \mathbf{o}$ if and only if $\mathbf{x}$ is not a KKT point. In this case,*
(a) $\sigma(\mathbf{x}) > 0$*;*
(b) $\overline{t}(\mathbf{x}) \geqslant \frac{1}{\sigma(\mathbf{x})}$*;*
(c) $\mathbf{v}(\mathbf{x})$ *is an ascent direction.*

Proof. First observe that the KKT conditions on $\mathbf{x}$ for the StQP are equivalent to the Nash equilibrium requirement $[A\mathbf{x}]_i \leqslant \mathbf{x}^{\top}A\mathbf{x}$ for all i, with equality for those i such that $x_i > 0$, see, e.g., Bomze (1998). Hence $r_i(\mathbf{x}) \leqslant 0$ for all i at KKT points $\mathbf{x}$, so that $\varphi\Big(r_i(\mathbf{x})\Big) = 0$ and also $\sigma(\mathbf{x}) = 0$, whence $\mathbf{v}(\mathbf{x}) = \mathbf{o}$ follows readily. To show the converse, note that $v_i(\mathbf{x}) = 0$ means $\varphi\Big(r_i(\mathbf{x})\Big) = \sigma(\mathbf{x})x_i$. Now if $\sigma(\mathbf{x})$ were strictly positive, then $r_i(\mathbf{x}) > 0$ also would hold for all i with $x_i > 0$, which would yield $[A\mathbf{x}]_i > \mathbf{x}^{\top}A\mathbf{x}$ for all such i, contradicting $\mathbf{x}^{\top}A\mathbf{x} = \sum_i x_i[A\mathbf{x}]_i$: thus we obtain $\sigma(\mathbf{x}) = 0$, which entails $r_i(\mathbf{x}) \leqslant 0$ or $[A\mathbf{x}]_i \leqslant \mathbf{x}^{\top}A\mathbf{x}$ for all i, with equality if $x_i > 0$ by the same averaging argument as above. Thus $\mathbf{x}$ must be a KKT point if $\mathbf{v}(\mathbf{x}) = \mathbf{o}$, which in turn is, as established just now, equivalent to $\sigma(\mathbf{x}) = 0$. Therefore also assertion (a) is proved. Now if $\mathbf{x} \in \Delta$ is an arbitrary non-KKT point, then $v_i(\mathbf{x}) < 0$ implies $x_i > 0$ because of $\varphi\Big(r_i(\mathbf{x})\Big) \geqslant 0$. Hence

$$\frac{x_i}{-v_i(\mathbf{x})} = \frac{x_i}{\sigma(\mathbf{x})x_i - \varphi\left(r_i(\mathbf{x})\right)} \geqslant \frac{1}{\sigma(\mathbf{x})}$$

for all i such that $v_i(\mathbf{x}) < 0$, which establishes (b). Finally, to show (c), note that by virtue of (5), the derivative of ψ at $t = 0$ equals $\dot{\psi}(0) = 2\mathbf{x}^{\top}A\mathbf{v}(\mathbf{x})$. By symmetry of A, we get

$$\begin{aligned}
\mathbf{x}^{\top}A\mathbf{v}(\mathbf{x}) &= \sum_{i,j=1}^{n} x_i a_{ij}\left[\varphi\Big(r_j(\mathbf{x})\Big) - \sigma(\mathbf{x})x_j\right] \\
&= \sum_{j=1}^{n} \varphi\Big(r_j(\mathbf{x})\Big)\left[\sum_{i=1}^{n} a_{ji}x_i - \mathbf{x}^{\top}A\mathbf{x}\right] \qquad (12)\\
&= \sum_{j=1}^{n} \varphi\Big(r_j(\mathbf{x})\Big)\, r_j(\mathbf{x}) \geqslant 0
\end{aligned}$$

with equality if and only if all $\varphi\left(r_j(\mathbf{x})\right) = 0$, which is equivalent to the KKT condition. Hence for every non-KKT point we get $\dot{\psi}(0) > 0$. □

Note that the presented class of ascent directions applies to general quadratic objective functions irrespective of curvature. However, the convex case (concave

maximization in our case) deserves special attention: first, we then know the KKT points are global solutions; second, these instances will appear as subproblems in the d.c. approach below. So before we proceed to prove a general convergence result for any curvature, we shortly digress to establish a helpful result which provides a shortcut to avoid calculating $\overline{t}(\mathbf{x})$ via (6), i.e. we provide conditions sufficient for $t(\mathbf{x}) = -\mathbf{x}^\top A\mathbf{v}(\mathbf{x})/\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x})$:

PROPOSITION 2. *Suppose that A is negative-semidefinite, and that $\mathbf{v}^\top A\mathbf{v} < 0$ holds for all $\mathbf{v} \in \mathbf{e}^\perp$. Then if $\sum_j [A\mathbf{x}]_j \varphi\left(r_j(\mathbf{x})\right) \geqslant 0$ we have*

$$t(\mathbf{x}) = \frac{\mathbf{x}^\top A\mathbf{v}(\mathbf{x})}{-\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x})}\,.$$

Proof. All we have to show is that the latter quantity does not exceed $\overline{t}(\mathbf{x})$. Now put $z_j = \varphi\left(r_j(\mathbf{x})\right)$ and $\mathbf{z} = [z_j] \in \mathbb{R}^n$, to arrive via (12) at

$$\begin{aligned}
\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) &= \sigma^2(\mathbf{x})\mathbf{x}^\top A\mathbf{x} - 2\sigma(\mathbf{x}) \sum_{i,j=1}^{n} x_i a_{ij} \varphi\left(r_j(\mathbf{x})\right) + \mathbf{z}^\top A\mathbf{z} \\
&= \sigma^2(\mathbf{x})\mathbf{x}^\top A\mathbf{x} - 2\sigma(\mathbf{x}) \sum_{j=1}^{n} [A\mathbf{x}]_j \varphi\left(r_j(\mathbf{x})\right) + \mathbf{z}^\top A\mathbf{z} \\
&= -\sigma(\mathbf{x})\mathbf{x}^\top A\mathbf{v}(\mathbf{x}) - \sigma(\mathbf{x}) \sum_{j=1}^{n} [A\mathbf{x}]_j \varphi\left(r_j(\mathbf{x})\right) + \mathbf{z}^\top A\mathbf{z} \\
&\leqslant -\sigma(\mathbf{x})\mathbf{x}^\top A\mathbf{v}(\mathbf{x})
\end{aligned}$$

by assumption, so that finally we get via Theorem 1(b)

$$\frac{\mathbf{x}^\top A\mathbf{v}(\mathbf{x})}{-\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x})} \leqslant \frac{1}{\sigma(\mathbf{x})} \leqslant \overline{t}(\mathbf{x})\,.$$

Hence the assertion. □

The next step is standard: given the definition of $\mathbf{v}(\mathbf{x})$ as in (11), generate the sequence $(\mathbf{x}_\nu)$ as in (8). Now we are ready to establish the convergence result:

THEOREM 3. *If the function φ is chosen continuous, then any accumulation point $\mathbf{x}$ of the sequence $(\mathbf{x}_\nu)$ satisfies the KKT condition.*

Proof. Suppose that $\mathbf{x}_{\nu_k} \to \mathbf{x}$ as $k \to \infty$ along some subsequence (ν_k). Put $\mathbf{v}_k = \mathbf{v}(\mathbf{x}_{\nu_k})$ and $t_k = t(\mathbf{x}_{\nu_k})$ as well as $\sigma_k = \sigma(\mathbf{x}_{\nu_k})$. By continuity, $\mathbf{v}_k \to \mathbf{v}(\mathbf{x})$ and $\sigma_k \to \sigma(\mathbf{x})$ as $k \to \infty$. Now, if $\mathbf{x}$ were no KKT point, then $\sigma(\mathbf{x}) > 0$ and $\mathbf{v}(\mathbf{x}) \neq \mathbf{o}$ by Theorem 1. Further, during the course of iterations, no $\mathbf{x}_\nu$ has been a KKT point,

since otherwise, the sequence would have stopped there. We now show that there is a constant $\rho > 0$ such that

$$t_k \geqslant \rho \quad \text{for infinitely many } k\,. \tag{13}$$

Obviously, we are done if $t_k = \overline{t}(\mathbf{x}_{\nu_k})$ for infinitely many k, since in this case Theorem 1(b) even yields $t_k \geqslant \frac{1}{\sigma_k} \geqslant \frac{1}{2\sigma(\mathbf{x})}$ for infinitely many k. So let us assume the contrary, namely that $t_k \neq \overline{t}(\mathbf{x}_{\nu_k})$ for all $k \geqslant \overline{k}$, which, by (7), yields,

$$t_k = t(\mathbf{x}_{\nu_k}) = \frac{\mathbf{x}_{\nu_k}^\top A\mathbf{v}_k}{-\mathbf{v}_k^\top A\mathbf{v}_k} \quad \text{for all } k \geqslant \overline{k}\,,$$

for some $\overline{k}$ large enough. Further, $t(\mathbf{x}_{\nu_k}) \neq \overline{t}(\mathbf{x}_{\nu_k})$ implies via (7) $\mathbf{v}_k^\top A\mathbf{v}_k < 0$ for all $k \geqslant \overline{k}$, yielding $\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) \leqslant 0$ by continuity. If $\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) < 0$, then a similar limit argument as above applies to arrive via $t_k \to \frac{\mathbf{x}^\top A\mathbf{v}(\mathbf{x})}{-\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x})} > 0$ at condition (13).[2] If, however, $\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) = 0$, then choose an arbitrary small number $\delta > 0$ and $\eta > 0$ so small, that $\frac{\mathbf{x}^\top A\mathbf{v}(\mathbf{x})-\delta}{\eta} > \overline{t}(\mathbf{x}) + \delta$. Next observe that $\limsup_k \overline{t}(\mathbf{x}_{\nu_k}) \leqslant \overline{t}(\mathbf{x})$, because for k large enough, we have $v_i(\mathbf{x}_{\nu_k}) < 0$ whenever $v_i(\mathbf{x}) < 0$ by continuity. Thus we may also pick k so large that both $\overline{t}(\mathbf{x}_{\nu_k}) < \overline{t}(\mathbf{x}) + \delta$ and $\mathbf{x}_{\nu_k}^\top A\mathbf{v}_k \geqslant \mathbf{x}^\top A\mathbf{v}(\mathbf{x}) - \delta$, and also $-\mathbf{v}_k^\top A\mathbf{v}_k \leqslant \eta$ hold (the latter is possible since $\lim_k \mathbf{v}_k^\top A\mathbf{v}_k = \mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x}) = 0$). Now

$$t_k = \frac{\mathbf{x}_{\nu_k}^\top A\mathbf{v}_k}{-\mathbf{v}_k^\top A\mathbf{v}_k} \geqslant \frac{\mathbf{x}_{\nu_k}^\top A\mathbf{v}_k}{\eta} \geqslant \frac{\mathbf{x}^\top A\mathbf{v}(\mathbf{x}) - \delta}{\eta} > \overline{t}(\mathbf{x}) + \delta \geqslant \overline{t}(\mathbf{x}_{\nu_k})$$

which is absurd. Hence (13) is established. We now are in the situation of Lemma 10.2.6 in (Bazaraa and Shetty, 1979), but for the readers' convenience we specify a slightly more direct argument here. Denote by

$$\psi_k(t) = f(\mathbf{x}_{\nu_k} + t\mathbf{v}_k) = \mathbf{x}_{\nu_k}^\top A\mathbf{x}_{\nu_k} + 2t\mathbf{x}_{\nu_k}^\top A\mathbf{v}_k + t^2\mathbf{v}_k^\top A\mathbf{v}_k\,, \quad t \geqslant 0\,.$$

Then by definition and from (13),

$$f(\mathbf{x}_{\nu_k+1}) = \psi_k(t_k) \geqslant \psi_k(\rho) = f(\mathbf{x}_{\nu_k}) + \rho[2\mathbf{x}_{\nu_k}{}^\top A\mathbf{v}_k + \rho\,\mathbf{v}_k^\top A\mathbf{v}_k]$$

for all k large enough. Decreasing $\rho > 0$ if necessary, we may and do ensure that $-\rho\,\mathbf{v}_k^\top A\mathbf{v}_k \leqslant \mathbf{x}_{\nu_k}^\top A\mathbf{v}_k$ for all such k (pass to the limit as $k \to \infty$, and define ρ in terms of $\mathbf{x}^\top A\mathbf{v}(\mathbf{x}) > 0$ and $\mathbf{v}(\mathbf{x})^\top A\mathbf{v}(\mathbf{x})$). Obviously, still $\psi_k(t_k) \geqslant \psi_k(\rho)$ holds. Thus we arrive at

$$f(\mathbf{x}_{\nu_k+1}) - f(\mathbf{x}_{\nu_k}) \geqslant \rho\,\mathbf{x}_{\nu_k}^\top A\mathbf{v}_k \geqslant \frac{\rho}{2}\,\mathbf{x}^\top A\mathbf{v}(\mathbf{x}) > 0$$

for infinitely many k, which contradicts the convergence of the objective values $f(\mathbf{x}_\nu)$ as $\nu \to \infty$ addressed in Remark 1 at the end of Subsection 2.1. Therefore $\mathbf{x}$ must be a KKT point and the convergence result is proved. □

[2] Recall that by (12), $\mathbf{x}^\top A\mathbf{v}(\mathbf{x}) > 0$ if $\mathbf{x}$ is no KKT point.

Again, it should be stressed that, as in Wolfe's reduced gradient algorithm, convergence of the whole sequence $(\mathbf{x}_\nu)$ is not ensured, whereas trajectories under the replicator dynamics (10) always converge. However, there is a simple condition which is generically satisfied, ensuring convergence for a considerably larger class of methods, applied to general problems to maximize a continuous function f over a compact feasible set M: let us call an iterative algorithm 'an iteratively improving KKT solver', if it generates, for any starting point $\mathbf{x}_0 \in M$, a sequence $(\mathbf{x}_\nu)$ of feasible points in M satisfying

$$f(\mathbf{x}_{\nu+1}) \geqslant f(\mathbf{x}_\nu) \quad \text{for all } \nu \in \mathbb{N}\,, \tag{14}$$

and such that any accumulation point of this sequence is a KKT point. Both the methods in Subsection 2.2 and the ones in this section (with continuous φ) are iteratively improving KKT solvers for the StQP (1) in this sense.

PROPOSITION 4. *If any two different KKT points* $\mathbf{x}$ *and* $\mathbf{y}$ *have different objective values,* $f(\mathbf{x}) \neq f(\mathbf{y})$*, then every iteratively improving KKT solver converges, as* $\nu \to \infty$*, to one of the KKT points.*

Proof. Suppose $\mathbf{x}$ and $\mathbf{y}$ were two different accumulation points of $(\mathbf{x}_\nu)$. Then (14) guarantees by monotonicity and continuity of the objective $f(\mathbf{x}_\nu) \to f(\mathbf{x}) = f(\mathbf{y})$ as $\nu \to \infty$, cf. Remark 1 at the end of Subsection 2.1. This contradicts the assumption. Hence there is at most one accumulation point of $(\mathbf{x}_\nu)$. On the other hand, compactness of M ensures there is at least one, so that the sequence $(\mathbf{x}_\nu)$ has a limit, which by the property of KKT solvers must be a KKT point. □

In some sense, this result is similar to the well-known observation that branch-and-bound procedures converge to the global optimizer, given this is unique (Horst and Tuy, 1993).

As a consequence, we obtain convergence of Wolfe's reduced gradient procedure and the new methods proposed in this section, if $\mathbf{x}^\top A\mathbf{x} \neq \mathbf{y}^\top A\mathbf{y}$ is ensured for all KKT points $\mathbf{x} \neq \mathbf{y}$ of (1). Now it can be shown that the set of symmetric $n \times n$ matrices satisfying this condition is open and dense in the space of all symmetric $n \times n$ matrices.

As similarity of numerical complications with either of the preceding different algorithms is not easily foreseen, it is not clear which of these local optimization methods are preferable from a practical point of view. A ranking between them is best established by a large empirical study which lies beyond the scope of this paper.

4. Traditional approaches to solve StQPs: escape strategies and other global optimization methods

The procedures described in the sequel all have in common that they are based upon some (or a special) local optimizer, and then they try to improve the quality of the

result obtained (or, in the ideal case, provide a certificate for global optimality of the current solution).

Following a perturbation approach similar to those below (see Subsection 5.2), in (Bomze et al., 2001) an escape is attempted by varying the objective function 'towards convexity', but with the aim to keep the most efficient local solutions while the others should be avoided. In a nutshell, this so-called 'annealed replication' approach consists of restarting the locally optimizing replicator dynamics with varying perturbation parameter, which is bounded a priori as to guarantee the above-mentioned positive features.

A totally different strategy is followed by 'Copositive Programming' (Bomze et al., 2000), where an StQP is reformulated into a linear optimization problem over a convex matrix cone, and interior primal-dual steps borrowed from the semidefinite programming (SDP) framework are used as escape step.

In a similar vein, the SDP relaxation procedures of (Quist et al., 1998) try to improve inefficient local solutions.

Based on global optimality conditions for general quadratic problems, and using block pivoting strategies, the G.E.N.F. algorithm introduced in (Bomze and Stix, 1999) produces either a certificate of global optimality, or delivers an improving feasible point, by a recursive procedure which decomposes the master problem into a series of considerably smaller subproblems of the same type, so as to employ the same locally optimizing routines.

Yet another approach is reported in Kuznetsova and Strekalovsky (2001) who use a particular d.c. decomposition of the objective function in conjunction with sublevel global optimality conditions. The resulting, so-called improved '$\mathcal{R}$-strategy' has nothing in common with branch-and-bound procedures, but yields results of considerable quality for medium-sized instances of the maximum-clique problem from the DIMACS testbed (Johnson and Trick, 1996).

Using SDP technology for obtaining good upper bounds, branch-and-bound ideas for StQPs are explicitly introduced in (Nowak, 1999). Contrasting with the d.c. approach proposed below, this paper concentrates of improving concave overestimation procedures.

A different, and apparently new, branch-and-bound approach, is presented in (Stix, 2000) and more generally in (Stix, 2001), where local search for a maximum clique is combined with a so-called 'target-branch-and-bound' strategy. In essence, this means that the usual problem tree is re-organized such that in a breadth-first search, the most promising regions are processed first. The author reports successful application, although he employs very rough upper bounds obtained from colouring heuristics. Here, the combinatorial aspect of the problem is the key for the branch-and-bound attack and also here, no d.c. decomposition is used.

Of course, this account is not exhaustive, because a bibliographic survey is rather beyond the scope of this article. A (most probably incomplete) list of other references to branch-and-bound methods for quadratic problems, which in one sense or another come close to the approach treated here, is Horst and Thoai,

Raber, Hansen et al., An and Tao, An and Tao, Phong et al. (1996, 1998, 1993, 1998, 1997, 1996). The first two articles in this list combine LP bounds with simplicial decomposition, while Hansen et al. and An and Tao (1993, 1998) treat quadratic problems over a hypercube, the former with derivatives, the latter with an ellipsoidal approach, which is presented in a more general setup for general polytopes in An and Tao (1997). Concave overestimation techniques are employed in Phong et al. (1996).

5. D.C. decomposition for StQPs

5.1. BASIC IDEAS AND RESULTS

The central idea is very simple, and motivated by the observation that the LP bounds widely used in branch-and-bound algorithms are often quite inefficient (Nowak, 1999): decompose the indefinite $n \times n$ matrix $A = B - C$ such that B and C are positive-semidefinite (psd) matrices, so that $\mathbf{x}^\top B\mathbf{x}$ and $\mathbf{x}^\top C\mathbf{x}$ are convex functions. Then for any simplex $X = \mathrm{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$

$$\beta(X) = \max\{\mathbf{x}^\top B\mathbf{x} : \mathbf{x} \in X\} = \max\left\{\mathbf{v}_i^\top B\mathbf{v}_i : i \in \{1, \ldots, n\}\right\} \tag{15}$$

while every *local* solution $\bar{\mathbf{x}} \in X$ to the problem

$$\gamma(X) = \min\{\mathbf{x}^\top C\mathbf{x} : \mathbf{x} \in X\} \tag{16}$$

is automatically a global one. Therefore, any of the locally optimizing procedures described in Sections 2 and 3, applied to

$$\min\{\mathbf{x}^\top C\mathbf{x} : \mathbf{x} \in X\} = -\max\{\mathbf{y}^\top(-V^\top CV)\mathbf{y} : \mathbf{y} \in \Delta\}$$

with $V = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$, yields $\gamma(X)$. Thus

$$\alpha(X) = \beta(X) - \gamma(X) \tag{17}$$

is an easily obtainable upper bound for $\max\{\mathbf{x}^\top A\mathbf{x} : \mathbf{x} \in X\}$.

We employ a *local* lower bound $\delta(X)$ with the property that

$$\delta(X_k) \to (\mathbf{x}^*)^\top A(\mathbf{x}^*) \quad \text{whenever} \bigcap_{k=1}^{\infty} X_k = \{\mathbf{x}^*\}\,. \tag{18}$$

Any feasible point $\mathbf{x} \in X$ gives such a lower bound $\delta(X) = \mathbf{x}^\top A\mathbf{x}$, for instance. An alternative would be to put $\delta(X)$ equal to the objective result of one of the locally optimizing procedures described in Sections 2 and 3. As always, the final choice of method depends on the balance between quality and effort. However, the better the quality of the locally optimizing procedure, the more cutting power is to be expected and the less subproblems will be generated. Compromises are available with every 'any-time' local optimizer (i.e., those maintaining feasibility during all iterations). Fortunately, all local optimization methods covered here enjoy this property,

with the exception of Lemke's complementary pivoting algorithm, since the latter does not generate (primally) feasible paths. The *global* lower bound needed in the algorithm then results by taking the largest of the $\delta(X)$ values obtained so far.

As the general roster of a branch-and-bound scheme applies, see, e.g., Horst and Tuy (1993) or Horst et al. (1995), we refrain from repeating it here. To select a subsimplex X for next bisection, usually an X with largest $\alpha(X)$ is chosen, in the hope to improve upper bounds as quickly as possible. Alternatively, one may pick a subsimplex X with largest local lower bound $\delta(X)$, to obtain at an early stage good intermediate results in case of premature abortion of the procedure. Regardless of the choice of these variants, we have the following convergence result which is more or less stated and proved in Proposition 3.15 of Horst et al. (1995). We repeat the main arguments here for the readers' convenience.

THEOREM 5. *Define $\alpha(X)$ as in (17) via (15) and (16). Then the branch-and-bound algorithm using simplicial bisection along the longest edge with upper bounds $\alpha(X)$ and any choice of lower bounds $\delta(X)$ satisfying (18) converges.*

Proof. In light of the general results in Chapter IV of Horst and Tuy (1993), see also Theorem 3.8 of Horst et al. (1995), all we have to show is the following: if X_k is an exhaustive sequence of simplices, i.e., shrinking to a singleton $\{\mathbf{x}^*\}$, then $\alpha(X_k) \to (\mathbf{x}^*)^\top A(\mathbf{x}^*)$ as $k \to \infty$. But this is evident as $\beta(X_k)$ obviously converges to $(\mathbf{x}^*)^\top B(\mathbf{x}^*)$ while $\gamma(X_k) \to (\mathbf{x}^*)^\top C(\mathbf{x}^*)$. Further, also $\delta(X_k) \to (\mathbf{x}^*)^\top A(\mathbf{x}^*)$ by assumption (18), which yields $\alpha(X_k) - \delta(X_k) \to 0$. Now exhaustiveness of the simplicial bisection, which means halving the longest edge, is shown in Proposition 3.14 of Horst et al. (1995), and the claimed assertion follows. □

Instead of halving the longest edge, one may bisect it at some ratio ranging in a fixed closed interval excluding 0 and 1, without losing exhaustivity (Horst, 1997). More sophisticated exhaustive subdivision procedures may be found in (Horst et al., 1992; Horst and Tuy, 1993). For variants thereof which also may involve the objective see Horst and Thoai (1989).

5.2. VARIANTS OF D.C. DECOMPOSITIONS

As usual in d.c. programming, the decomposition $A = B - C$ is by no means unique. One possibility is the *perturbation* decomposition $B = \mu I_n$ and $C = \mu I_n - A$ where $\mu \geqslant 0$ is chosen large enough to make C psd. The smallest possible value is of course the spectral bound $\mu = [\lambda_{\max}(A)]_+$ with

$$\lambda_{\max}(A) = \max\{\lambda : \lambda \text{ is an eigenvalue of } A\}\,.$$

Since upper bounds of the spectrum are cheap, this method seems to be of advantage for very large instances. Further, $\beta(\mathrm{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n)) = \mu \max\{\|\mathbf{v}_i\|^2 : i \in \{1, \ldots, n\}\}$ is very easy to update while

$$\gamma(\mathrm{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n)) = \mu \min\{\mathbf{x}^\top V^\top V \mathbf{x} : \mathbf{x} \in \Delta\}\,.$$

Also other positive-definite perturbation matrices D could replace I_n: $B = \mu D$ and $C = \mu D - A$, where μ now must be an upper bound for the Rayleigh quotient $\max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^\top A\mathbf{x}}{\mathbf{x}^\top D\mathbf{x}}$. An alternative perturbation decomposition may employ $B = A+\overline{\mu}\overline{D}$ and $C = \overline{\mu}\overline{D}$. It is merely a question of notation which of these variants are chosen: putting $\mu D = A+\overline{\mu}\overline{D}$ or $\overline{\mu}\overline{D} = \mu D - A$ does the trick. Differences between these two variants mainly occur in numerics, e.g., if D (or $\overline{D}$) are sparse.

A second strategy could employ the *spectral* d.c. decomposition: if $A = ULU^\top$ with U being a orthonormal $n \times n$ matrix consisting of eigenvectors of A while $L = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix made from the eigenvalues of A, then $A = B - C$ is a d.c. decomposition with $B = UL_+U^\top$ while $C = UL_-U^\top$ with

$$L_\pm = \mathrm{diag}([\lambda_1]_\pm, \ldots, [\lambda_n]_\pm)\,.$$

Of course, then, $L_\pm$ are psd, and so are B and C, while $L = L_+ - L_-$ implies $A = B - C$. A drawback of this approach is the necessity to determine the full spectral information about A. On the other hand, this d.c. decomposition is, within a 'natural' class of d.c. decompositions, minimal in the sense of (Dür, 1999). There, a local optimality condition for d.c. functions is presented in terms of ε-subdifferentials, which gives a third aspect of quality, besides size of basin of attraction, and functional quality.

We now establish the addressed minimality result for the spectral decomposition:

THEOREM 6. *Consider two different d.c. decompositions of*

$$f(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} = g_{\mathrm{spec}}(\mathbf{x}) - h_{\mathrm{spec}}(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x})\,,$$

where

$$g_{\mathrm{spec}}(\mathbf{x}) = \mathbf{x}^\top A_+\mathbf{x}\,, \quad h_{\mathrm{spec}}(\mathbf{x}) = \mathbf{x}^\top A_-\mathbf{x}$$

with $A_\pm = UL_\pm U^\top$ *and*

$$g(\mathbf{x}) = \mathbf{x}^\top B\mathbf{x}\,, \quad h(\mathbf{x}) = \mathbf{x}^\top(B - A)\mathbf{x}\,,$$

respectively, where $BA = AB$ *and both* B *and* $B - A$ *are psd matrices, so that* $g, g_{\mathrm{spec}}, h, h_{\mathrm{spec}}$ *are convex functions.*
Then $g_{\mathrm{spec}}(\mathbf{x}) \leqslant g(\mathbf{x})$ *and* $h_{\mathrm{spec}}(\mathbf{x}) \leqslant h(\mathbf{x})$*, and also both* $g - g_{\mathrm{spec}}$ *and* $h - h_{\mathrm{spec}}$ *are convex functions.*

Proof. If B commutes with A, then $B = UMU^\top$ with U containing the (common) orthonormal eigenvectors of A and B while $M = \mathrm{diag}(\mu_1, \ldots, \mu_n)$ contains the eigenvalues of B. Now if B is psd, then necessarily $\mu_i \geqslant 0$ for all $i \in \{1, \ldots, n\}$. Similarly, since $B - A$ is also psd, this means $\mu_i - \lambda_i \geqslant 0$, i.e., $\mu_i \geqslant [\lambda_i]_+$ for all $i \in \{1, \ldots, n\}$, which in turn entails psdness of the matrix $B - A_+$. Hence $g - g_{\mathrm{spec}}$ must be convex, and also non-negative. Now $h - h_{\mathrm{spec}} = (g - f) - (g_{\mathrm{spec}} - f) = g - g_{\mathrm{spec}}$, so that the result for the concave parts follows readily. □

Taking $\mu \geqslant [\lambda_{\max}(A)]_+$ and $B = \mu I_n = U\,\mathrm{diag}(\mu, \ldots, \mu)U^\top$, we readily see that the perturbation approach addressed above falls into this class of commuting decompositions, so that we have

$$h_{\mathrm{pert}}(\mathbf{x}) - h_{\mathrm{spec}}(\mathbf{x}) = g_{\mathrm{pert}}(\mathbf{x}) - g_{\mathrm{spec}}(\mathbf{x}) \geqslant 0$$

if

$$g_{\mathrm{pert}}(\mathbf{x}) = \mu\|\mathbf{x}\|^2 \quad \text{and} \quad h_{\mathrm{pert}}(\mathbf{x}) = \mu\|\mathbf{x}\|^2 - \mathbf{x}^\top A\mathbf{x}\,.$$

Similar to the above result, we obtain the following

LEMMA 7. *Suppose that $A^0_\pm$ are psd with $A = A^0_+ - A^0_-$, and such that $A_\pm - A^0_\pm$ are also psd. Then necessarily $A^0_\pm = A_\pm$.*

Proof. If $\mathbf{u}_i$ is an eigenvector of A corresponding to a nonnegative eigenvalue $\lambda_i \geqslant 0$, then by assumption,

$$0 \leqslant \|\sqrt{A^0_-}\mathbf{u}_i\|^2 = \mathbf{u}_i^\top A^0_-\mathbf{u}_i \leqslant \mathbf{u}_i^\top A_-\mathbf{u}_i = 0\,,$$

(with $\sqrt{A^0_-}$ denoting the symmetric square-root factorization of the psd matrix A^0_-), so that $A^0_-\mathbf{u}_i = \mathbf{o}$ follows. But then

$$\lambda_i\mathbf{u}_i = A\mathbf{u}_i = A^0_+\mathbf{u}_i - \mathbf{o} = A^0_+\mathbf{u}_i\,.$$

Similarly, $A^0_+\mathbf{u}_i = \mathbf{o}$ if $\lambda_i < 0$. Hence both $A^0_\pm$ have the same eigenvectors as A(and hence as $A_\pm$), and the same eigenvalues as $A_\pm$, respectively. Thus $A^0_\pm = A_\pm$. □

One may now hope that there is a "global" answer to the question of minimality. However, the following example shows that this cannot happen in more than one dimension:

Example 1. Let

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ with } A_+ = \tfrac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} \frac{1}{3} & 1 \\ 1 & \frac{7}{2} \end{bmatrix}.$$

Then both B and $B - A$ are psd, but

$$B - A_+ = \begin{bmatrix} -\frac{1}{6} & \frac{1}{2} \\ \frac{1}{2} & 3 \end{bmatrix}$$

is obviously indefinite.

PROPOSITION 8. *There is no 'globally minimal' d.c. decomposition of quadratic functions defined on $\mathbb{R}^n$ if $n \geqslant 2$, in the sense of Theorem 6 above.*

Proof. Without loss of generality we may and do assume that $n = 2$. Now suppose that $A = A^0_+ - A^0_-$ yields a globally minimal d.c. decomposition of $\mathbf{x}^\top A\mathbf{x}$ where A is as in Example 1. Then, by definition of minimality, $A_+ - A^0_+$ and also $A_- - A^0_-$ are psd. But then $A^0_\pm = A_\pm$, by Lemma 7. On the other hand, with B as in Example 1, also $A = B + (A - B)$ yields a d.c. decomposition, we would arrive at psdness of $B - A^0_+$ (and also $(A - B) - A^0_-$) by minimality of $A^0_\pm$ and thus at psdness of $B - A_+$, which as shown is absurd. □

The following subsection provides a possible way out of this dilemma, to select a d.c. decomposition which yields a minimal upper bound for $\alpha(\Delta)$ from Subsection 5.1.

5.3. SELECTING A D.C. DECOMPOSITION WITH MINIMAL UPPER BOUNDS VIA SDP

In the sequel, upper bounds for $\alpha(\Delta)$ are proposed which can be minimized to select a d.c. decomposition. This can be accomplished by SDP procedures which become increasingly popular nowadays. Note that it is *not* suggested to replace $\alpha(X)$ in the branch-and-bound procedure with this new upper bound, but rather use it as a guideline to select one out of many possible d.c. decompositions. This is important as the obtained upper bound is worse (i.e., larger than) $\alpha(\Delta)$, and by no means cheaper to obtain.

Now denote by

$$\lambda_{\min}(C) = \min\{\lambda : \lambda \text{ is an eigenvalue of } C\}\,.$$

PROPOSITION 9. *Given a d.c. decomposition $A = B - C$ with B and C psd, we have*

$$\alpha(\Delta) \leqslant \lambda_{\max}(B) - \tfrac{1}{n}\lambda_{\min}(C).$$

Proof. First observe that

$$\mathbf{x}^\top C\mathbf{x} \geqslant \lambda_{\min}(C)\|\mathbf{x}\|^2 \geqslant \lambda_{\min}(C)\tfrac{1}{n} \quad \text{for all } \mathbf{x} \in \Delta\,, \tag{19}$$

because $\|\mathbf{x}\|\|\mathbf{e}\| \geqslant \mathbf{x}^\top\mathbf{e} = 1$ holds for all $\mathbf{x} \in \Delta$ due to the Cauchy/Schwarz inequality, and because $\|\mathbf{e}\| = \sqrt{n}$. Further, note that

$$\mathbf{x}^\top B\mathbf{x} \leqslant \lambda_{\max}(B)\|\mathbf{x}\|^2 \leqslant \lambda_{\max}(B) \quad \text{for all } \mathbf{x} \in \Delta\,. \tag{20}$$

Combining both estimates yields the result. □

It now remains to reformulate the quest of tighter (i.e., smaller) upper bounds as a semidefinite program: to this end, observe that

$$\begin{aligned} \lambda_{\min}(C) &= \max\{t \geqslant 0 : (C - tI_n) \text{ is psd}\}\,, \\ \lambda_{\max}(B) &= \min\{s \geqslant 0 : (sI_n - B) \text{ is psd}\}\,. \end{aligned} \tag{21}$$

Now put $Y = sI_n - B$ and $Z = C - tI_n$. Then a minimal upper bound in the sense of Proposition 9 is attained if and only if (s, t, B, Y, Z) solve the following optimization problem:

$$\begin{aligned} s - \tfrac{1}{n}t &\to \min\ ! \\ B + Y - sI_n &= O \\ B - Z - tI_n &= A \\ s, t \geqslant 0\,, \quad B, Y, Z &\quad \text{are psd.} \end{aligned} \tag{22}$$

Problem (22) involves decision variables which are organized in symmetric $n \times n$ matrices (B, Y, Z), and constraints which are formulated as definiteness conditions on these matrices, plus linear constraints, and a linear objective. Hence (22) is a semidefinite program (SDP). Note that all the constraints in the last line of (22) can be compactly reformulated as psdness of the block-diagonal $(3n+2) \times (3n+2)$ matrix

$$\begin{bmatrix} s & & & & \\ & t & & & \\ & & B & & \\ & & & Y & \\ & & & & Z \end{bmatrix}$$

where all entries not specified are zero. This reformulation may be useful for primal–dual solution approaches. There is a fully established duality theory for optimization problems of this kind, and interior-point methods have proven to be a powerful tool for solving them. For an excellent recent reference see Renegar (2001).

From a solution $(s^*, t^*, B^*, Y^*, Z^*)$ of (22), psd B^* and $C^* = Z^* + t^* I_n$ can be used for the d.c. decomposition upon which the calculations of $\alpha(X) = \beta(X) - \gamma(X)$ in Subsection 5.1 are based. The rationale behind this proposal is the hope that not only $\alpha(\Delta) \leqslant s^* - \frac{1}{n}t^*$ is as tight as possible but also the upper bounds $\alpha(X)$ for many of the subsimplices X generated by the branching part of the algorithm.

Finally, we present a lower bound $\rho(C)$ on $\gamma(\Delta)$ which is tighter than $\frac{1}{n}\lambda_{\min}(C)$ from (19). To this end we replace I_n by $\frac{1}{n}E_n$, a psd matrix of rank one:

PROPOSITION 10. *If C is psd, define*

$$\rho(C) = \max\{r \geqslant 0 : C - r\,E_n \text{ is psd}\}\,. \tag{23}$$

Then

$$\mathbf{x}^\top C \mathbf{x} \geqslant \rho(C) \geqslant \tfrac{1}{n}\lambda_{\min}(C) \quad \text{for all } \mathbf{x} \in \Delta\,.$$

Proof. We have to establish two inequalities. To prove the leftmost one, observe that psdness of $C - rE_n$ implies $\mathbf{x}^\top C\mathbf{x} - r = \mathbf{x}^\top(C - rE_n)\mathbf{x} \geqslant 0$ for all $\mathbf{x} \in \Delta$, as $\mathbf{x}^\top E_n\mathbf{x} = (\mathbf{e}^\top\mathbf{x})^2 = 1$ for all such $\mathbf{x}$. The rightmost inequality requires a bit more: first, note that the matrix $P = I_n - \frac{1}{n}E_n$ (the orthoprojection matrix onto $\mathbf{e}^\perp$) is psd. Hence, if $C - tI_n$ is psd for some $t \geqslant 0$ and $r = \frac{t}{n}$, then

$$C - rE_n = C - tI_n + tI_n - \tfrac{t}{n}E_n = C - tI_n + tP \quad \text{is psd.}$$

Thus $\{t \geqslant 0 : C - tI_n \text{ is psd }\} \subseteq \{t \geqslant 0 : C - \frac{t}{n}E_n \text{ is psd }\}$, hence

$$\max\{t \geqslant 0 : C - tI_n \text{ is psd }\} \leqslant n \max\{\tfrac{t}{n} \geqslant 0 : C - \tfrac{t}{n}E_n \text{ is psd }\},$$

and the latter expression equals $n\rho(C)$. So we obtain $\frac{1}{n}\lambda_{\min}(C) \leqslant \rho(C)$ with the help of (21). □

Note that the latter inequality is strict, e.g., for $C = E_n$ where $\frac{1}{n}\lambda_{\min}(E_n) = 0 < 1 = \rho(E_n)$. Now Proposition 10 implies

$$\alpha(\Delta) \leqslant \lambda_{\max}(B) - \rho(C) \leqslant \lambda_{\max}(B) - \tfrac{1}{n}\lambda_{\min}(C)\,,$$

and we can try to minimize the tighter bound (center term above): put again $Y = sI_n - B$ but now $Z = C - rE_n$ to arrive at the following SDP which refines (22):

$$\begin{aligned} s - r &\to \min\,! \\ B + Y - sI_n &= O \\ B - Z - rE_n &= A \\ r, s \geqslant 0\,, \quad B, Y, Z &\quad \text{are psd.} \end{aligned} \tag{24}$$

As with (22), from an optimal solution $(r^*, s^*, B^*, Y^*, Z^*)$ to (24) extract positive semidefinite matrices B^* and $C^* = Z^* + r^*E_n$ which give a (hopefully) effective d.c. decomposition of $\mathbf{x}^\top A\mathbf{x}$ in the sense that $\alpha(\Delta) \leqslant s^* - r^*$.

5.4. FINAL REMARKS ON ALTERNATIVE BRANCH-AND-BOUND STRATEGIES VIA SDP

At the end of this paper, let us address the question of attacking problem (1) by a branch-and-bound strategy aiming directly at A rather than using a d.c. decomposition, but also using SDP technology for obtaining upper bounds.

Remark 2. For an alternative upper bound $\tilde{\alpha}(X)$ of the objective on a subsimplex X, we could employ (20) for $V^\top AV$ instead of B (which holds irrespective of definiteness):

$$\tilde{\alpha}(X) = \lambda_{\max}(V^\top AV) \quad \text{with } V = [\mathbf{v}_1, \ldots, \mathbf{v}_n] \text{ if } X = \operatorname{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

Then $\max_{\mathbf{x}\in X} \mathbf{x}^\top A\mathbf{x} = \max_{\mathbf{y}\in\Delta} \mathbf{y}^\top(V^\top AV)\mathbf{y} \leqslant \tilde{\alpha}(X)$ as in (20), where $\tilde{\alpha}(X)$ can be determined also via an SDP, namely

$$\begin{aligned} s &\to \min\,! \\ S - sI_n &= -V^\top AV \\ s \in \mathbb{R}\,, \quad S &\quad \text{is psd.} \end{aligned} \tag{25}$$

Now transition from X_k to a subsimplex X_{k+1} generated by bisection of X_k means a rank-two update of the corresponding matrix $V = V_k$, giving $V_{k+1} = V_k + \mathbf{f}\mathbf{g}^\top$ (if the i-th vertex $V_k\mathbf{e}_i$ is replaced with the midpoint of the edge to the vertex $V_k\mathbf{e}_j$, then $\mathbf{f} = \frac{1}{2}V_k(\mathbf{e}_j - \mathbf{e}_i)$ and $\mathbf{g} = \mathbf{e}_i$). Thus the right-hand side $A_k = -V_k^\top A V_k$ of (25) undergoes a symmetric rank-two update

$$A_{k+1} = A_k - [\mathbf{g}\mathbf{f}^\top A V_k + V_k^\top A \mathbf{f}\mathbf{g}^\top] - (\mathbf{f}^\top A \mathbf{f})\mathbf{g}\mathbf{g}^\top . \tag{26}$$

Still, tight quantitative (right-hand side) sensitivity results for SDPs under updates of this form are by no means immediate. Hence for practical purposes, a d.c. decomposition (possibly obtained by using (24) only once) yielding bounds $\alpha(X)$ as in Subsection 5.1 seems to be preferable than recalculating the alternative upper bounds $\tilde{\alpha}(X_k)$ by solving (25) anew for every subsimplex X_k generated by branching. Returning to the special global case $X = \Delta$, note that

$$\tilde{\alpha}(\Delta) = \lambda_{\max}(A) \leqslant \lambda_{\max}(B) - \lambda_{\min}(C) \leqslant \lambda_{\max}(B) - \tfrac{1}{n}\lambda_{\min}(C) ,$$

refining the bound specified in Proposition 9. However, $\tilde{\alpha}(\Delta)$ may exceed $\alpha(\Delta)$ and even the sharper bound $\lambda_{\max}(B) - \rho(C)$ from Proposition 10 (for instance, if $B = O$ and $C = E_n$: note that then $-V^\top A V = V^\top E_n V = E_n = -A$ for all subsimplices and similarly for B and C, because of $E_n V = \mathbf{e}\mathbf{e}^\top[\mathbf{v}_1, \ldots, \mathbf{v}_n] = \mathbf{e}\mathbf{e}^\top = E_n$, so that the relation $\lambda_{\max}(A) = 0 > -1 = \lambda_{\max}(B) - \rho(C)$ holds throughout all subsimplices).

Remark 3. Now consider the bound $\rho(C)$ defined in (23): by the same argument that led to the proof of Proposition 10, we have

$$\max_{\mathbf{x}\in\Delta} \mathbf{x}^\top A \mathbf{x} = -\min_{\mathbf{x}\in\Delta} \mathbf{x}^\top(-A)\mathbf{x} \leqslant -\rho(-A) ,$$

so that an alternative upper bound is

$$\widehat{\alpha}(X) = -\rho(-V^\top A V) .$$

As with $\tilde{\alpha}(X)$, also this bound $\widehat{\alpha}(X)$ can be expressed as the optimal objective value of a suitable SDP. The above caveat regarding efficient sensitivity results under rank-two updates (26) of course applies also to this SDP, so that tedious repeated calculations of $\widehat{\alpha}(X_k)$ seemingly cannot be avoided.

As an ultimate remark, note that the common observations in Remarks 2 and 3 above also reflect one reason why we did, in this paper, not restrict analysis to the curvature of the objective on the hyperplane $\mathbf{e}^\perp$ rather than on the full $\mathbb{R}^n$: indeed, e.g. $\mathbf{x}^\top B \mathbf{x}$ is convex on Δ if and only if PBP is psd where $P = I_n - \frac{1}{n}E_n$ is the orthoprojector onto $\mathbf{e}^\perp$. But as

$$P(V^\top B V)P \neq V^\top P B P V = V^\top B V - \tfrac{1}{n}[E_n B V + V^\top B E_n] + \tfrac{\mathbf{e}^\top B \mathbf{e}}{n^2} E_n$$

in general (the latter equality is due to $PV = V - \frac{1}{n}E_n$, while there is no similar result for VP), this relative convexity argument seems to be of minor relevance in the present branch-and-bound concept.

6. Conclusion

In this paper, we have discussed some old and new local optimization procedures for a central class in quadratic optimization, the standard quadratic problems (StQP). The reason why we paid so much attention on these in a paper on global optimization is straightforward: we need some of them for obtaining high-quality bounds for branch-and-bound algorithms for StQPs based on d.c. decompositions, for which we considered different variants. Some of them are comparable in terms of minimal curvature, but it turns out that there is no globally optimal decomposition in this sense. Finally, we hinted at the question how to obtain d.c. decompositions with tighter upper bounds via semidefinite programming. Which of these variants should be chosen remains to be decided, possibly via large empirical studies that, hopefully, will be initiated by this article.

7. Acknowledgements

The author is indebted to Mirjam Dür and Grete Einsiedler-Mucha for stimulating discussions and valuable suggestions, as well as to an anonymous referee for helpful remarks.

References

An, L. T. H. and Tao, P. D. Solving a class of linearly constrained indefinite quadratic problems by DC algorithms. *J. Global Optimiz.* 11: 253–285, 1997.

An, L. T. H. and Tao, P. D. A branch and bound method via d. c. optimization algorithms and ellipsoidal technique for box constrained nonconvex quadratic problems. *J. Global Optimiz.* 13: 171–206, 1998.

Bazaraa, M. S. and Shetty, C. M. *Nonlinear programming — theory and algorithms.* Wiley, New York, 1979.

Bomze, I. M. On standard quadratic optimization problems. *J. Global Optimiz.* 13: 369–387, 1998.

Bomze, I. M., Budinich, M., Pardalos, P. M. and Pelillo, M. The maximum clique problem. In D.-Z. Du and P. M. Pardalos, editors, *Handbook of Combinatorial Optimization* suppl. Vol. A:1–74. Kluwer, Dordrecht, 1999.

Bomze, I. M., Budinich, M., Pelillo, M. and Rossi, C. Annealed replication: a new heuristic for the maximum clique problem. To appear in: Discrete Applied Math., 2001.

Bomze, I. M., Dür, M., de Klerk, E., Quist, A. J., Roos, C. and Terlaky, T. On copositive programming and standard quadratic optimization problems. *J. Global Optimiz.* 18: 301–320, 2000.

Bomze, I. M. and Stix, V. Genetical engineering via negative fitness: evolutionary dynamics for global optimization. *Annals of O.R.* 89: 279–318, 1999.

Cegielski, A. The Polyak subgradient projection method in matrix games. *Discuss. Math.* 13: 155–166, 1993.

Dür, M. A Note on Local and Global Optimality Conditions in D.C.-Programming. Research Report No. 56, Dept. of Statistics, Vienna Univ. Econ., 1999.

Hansen, P., Jaumard, B., Ruiz, M. and Xiong, J. Global minimization of indefinite quadratic functions subject to box constraints. *Nav. Res. Logist.* 40: 373–392, 1993.

Horst, R. On generalized bisection of n-simplices. *Math. of Comput.* 66: 691–698, 1997.

Horst, R., Pardalos, P. M. and Thoai, V. N. *Introduction to Global Optimization*. Kluwer, Dordrecht, 1995.

Horst, R. and Thoai, V. N. Modification, implementation and comparison of three algorithms for globally solving linearly constrained concave minimization problems. *Computing* 42: 271–289, 1989.

Horst, R. and Thoai, V. N. A new algorithm for solving the general quadratic programming problem. *Comput. Optim. Appl.* 5: 39–48, 1996.

Horst, R., Thoai, V. N. and de Vries, J. On geometry and convergence of a class of simplicial covers. *Optimization* 25: 53–64, 1992.

Horst, R. and Tuy, H. *Global Optimization*. Springer, Heidelberg, 1993.

Johnson, D. S. and Trick, M. A. (editors). *Cliques, Coloring, and Satisfiability: Second DIMACS Implementation Challenge*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science 26. American Mathematical Society, Providence, RI, 1996.

Kuznetsova, A. and Strekalovsky, A. On solving the maximum clique problem. *J. Global Optimiz.* 21: 265–288, 2001.

Massaro, A., Pelillo, M. and Bomze, I. M. A complementary pivoting approach to the maximum weight clique problem. To appear in: *SIAM J. Optimiz.*, 2001.

Murty, K. G. and Kabadi, S. N. Some NP-complete problems in quadratic and linear programming. *Math. Programming* 39: 117–129, 1987.

Nowak, I. A new semidefinite programming bound for indefinite quadratic forms over a simplex. *J. Global Optimiz.* 14: 357–364, 1999.

Phong, T. Q., An, L. T. H. and Tao, P. D. On globally solving linearly constrained indefinite quadratic minimization problems by decomposition branch and bound method. *RAIRO, Rech. Oper.* 30: 31–49, 1996.

Quist, A. J., de Klerk, E., Roos, C. and Terlaky, T. Copositive relaxation for general quadratic programming. *Optimization Methods and Software* 9: 185–209, 1998.

Raber, U. A simplicial branch-and-bound method for solving nonconvex all-quadratic programs. *J. Global Optimiz.* 13: 417–432, 1998.

Renegar, J. *A mathematical view of interior-point methods in convex optimization*. Forthcoming, SIAM, Philadelphia, PA, 2001.

Stix, V. Global optimization of standard quadratic problems including parallel approaches. Ph.D. thesis, Univ. Vienna, 2000.

Stix, V. Target-oriented branch-and-bound method for global optimization. Preprint, Univ. Vienna, 2001.

Weibull, J. W. *Evolutionary Game Theory*. MIT Press, Cambridge, MA, 1995.