Journal of Global Optimization **18:** 337–356, 2000.

Dual Bounds and Optimality Cuts for All-Quadratic Programs with Convex Constraints

IVO NOWAK
Humboldt-Universität zu Berlin, Rudower Chaussee 25, D-10099 Berlin, Germany (e-mail: ivo@mathematik.hu-berlin.de)

Abstract. A central problem of branch-and-bound methods for global optimization is that often a lower bound do not match with the optimal value of the corresponding subproblem even if the diameter of the partition set shrinks to zero. This can lead to a large number of subdivisions preventing the method from terminating in reasonable time. For the all-quadratic optimization problem with convex constraints we present optimality cuts which cut off a given local minimizer from the feasible set. We propose a branch-and-bound algorithm using optimality cuts which is finite if all global minimizers fulfill a certain second order optimality condition. The optimality cuts are based on the formulation of a dual problem where additional redundant constraints are added. This technique is also used for constructing tight lower bounds. Moreover we present for the box-constrained and the standard quadratic programming problem dual bounds which have under certain conditions a zero duality gap.

AMS subject classification: 90C20, 90C30, 90C06, 65K05

Key words: Global optimization, Nonconvex quadratic programming, Lagrangian relaxation, Optimality cuts, Duality gap

1. Introduction

In this paper we consider the following all-quadratic optimization problem:

$$
\text{(Q)} \qquad \begin{array}{rll}
\text{global minimize} & q_0(x) & \\
\text{subject to} & q_i(x) \le 0, & i \in I_{in} \\
 & q_i(x) = 0, & i \in I_{eq}
\end{array}
$$

where $q_i(x) := \frac{1}{2}x^T A_i x + b_i^T x + c_i$, $A_i \in \mathbb{R}^{(n,n)}$, $b_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, $i \in I_{in} \cup I_{eq} \cup \{0\}$. It is assumed that the inequality constraints of problem (Q) are convex and the equality constraints are linear.

Problem (Q) plays an interesting role in global optimization. Important special cases of (Q) are for example the trust region problem with one or several ellipsoid constraints, the box-constrained quadratic program and the standard quadratic program. It is known that problem (Q) is NP-hard [12]. For applications and solution methods we refer to [8, 2, 6, 13, 26, 27, 19, 20, 22, 3, 4, 21]. Many solution methods for problem (Q) are based on the branch-and-bound (B&B) principle.

A well-known difficulty of B&B algorithms is that often regions containing a global minimizer have to be subdivided very often in order to get sharp lower bounds. In large dimensions this can prevent the method from terminating in reasonable time. Almost all existing bounding methods for problem (Q) produce lower bounds which usually do not match with the optimal value of the corresponding subproblem. This can lead to infinitely many iterations of a B&B algorithm. Finite termination of B&B algorithms can be proved often only for so-called ϵ-optimal solutions.

We propose to use **optimality cuts** for avoiding this difficulty. An optimality cut is a cutting plane which cuts off a part of the feasible set containing a local minimizer. In Section 2 we propose a B&B algorithm using optimality cuts which solves problem (Q) in finite time under certain assumptions. The basis for the construction of optimality cuts is Lagrangian relaxation of problem (Q) which we introduce in Section 3. Based on these results we construct in Section 4 tight lower bounds and derive in Section 5 optimality cuts. Moreover we present in Section 6 for the box-constrained and the standard quadratic programming problem dual bounds which have under certain conditions a zero duality gap. We finish in Section 7 with a short discussion on the implementation of the algorithm and with some conclusions.

2. A Finite Branch-and-Bound Algorithm

In this section we describe a B&B algorithm for solving problem (Q) in finitely many iterations. We begin with the description of basic operations of a B&B algorithm.

2.1. BASIC B&B OPERATIONS

Partition Sets. Denote by $\Omega \subset \mathbb{R}^n$ the feasible set of problem (Q). Let $S_1, .., S_r$ be subsets of $\mathbb{R}^n$ such that

$$\bigcup_{i=1}^{r} S_i \supset \Omega \quad \text{and} \quad \text{int } S_i \cap \text{ int } S_j = \emptyset \text{ for } i \neq j.$$

Each subset S_i is called partition set and the collection of partition sets denoted by $\mathcal{P} := \{S_1, .., S_r\}$ is called a partition of Ω.

Subdivision Methods. A subdivision method defines from a given partition a new partition by subdividing one or several partition sets. A nested subsequence of partition sets $\{S_i\}$, (i.e. $S_{i+1} \subset S_i \quad \forall i$), is called exhaustive if S_i shrinks to a unique point, i.e. $\bigcap_{i=1}^{\infty} S_i = \{x\}$. A partition method is called exhaustive if every nested subsequence of partition sets generated by the subdivision method is exhaustive. Examples for exhaustive partition methods are given in [12].

Lower Bounds. Let S be a closed subset of $\mathbb{R}^n$ The optimal value of $q_0(x)$ over $\Omega \cap S$ is denoted by

$$q^*(S) := \begin{cases} \min\limits_{x \in \Omega \cap S} q_0(x) & : \quad \text{if } \Omega \cap S \neq \emptyset \\ \infty & : \quad \text{else} \end{cases}. \tag{1}$$

A lower bound of $q^*(S)$ is denoted by $\mu(S)$. A lower bounding method is called **tight** if

$$\lim_{i \to \infty} \mu(S_i) = \begin{cases} q_0(\hat{x}) & : \quad \text{if } \hat{x} \in \Omega \\ \infty & : \quad \text{else} \end{cases}$$

where $\{S_i\}$ is an exhaustive nested subsequence with $\bigcap_{i=1}^{\infty} S_i = \{\hat{x}\}$.

Heuristics. A heuristic $F(S) \in \mathbb{R}^n$ for problem (Q) is defined by the following two steps:

(i) In the first step a point $\tilde{x} \in S$ is computed which should be close to a global minimizer of q_0 over $S \cap \Omega$ if $S \cap \Omega \neq \emptyset$. This can be done by simply choosing an arbitrary point in S or by applying a more involved procedure as for example a Lagrange Heuristic (see Remark 1).

(ii) In the second step the initial point $\tilde{x}$ is refined by applying a local search method $LS(\tilde{x})$ starting from the point $\tilde{x}$ for minimizing an exact penalty function as for example $\min\limits_{x \in \mathbb{R}^n} P(x) := q_0(x) + \sum\limits_{i \in I_{in}} \rho_i \max\{0, q_i(x)\} + \sum\limits_{i \in I_{eq}} \rho_i |q_i(x)|$ where ρ_i, $i \in I_{in} \cup I_{eq}$, are sufficiently large penalty parameters.

Note that the point computed by $LS(\tilde{x})$ must not be in $\Omega \cap S$ and is not necessarily a strict local minimizer. We denote by $\Theta(x^*)$ the region of attraction of a local minimizer of $P(x)$ with respect to the local search method, i.e. $\Theta(x^*) := \{x \in \mathbb{R}^n : LS(x) = x^*\}$. We require the following property:

ASSUMPTION 1. *For all global minimizers x^* of problem (Q) there exist a closed ball $B_\delta(x^*)$ with center x^* and diameter $\delta > 0$ such that $B_\delta(x^*) \subset \Theta(x^*)$.*

Assumption 1 can be satisfied for example if the global minimizers of (Q) fulfill a certain constraint qualification (see [25]; Satz 3.6.5).

Upper Bounds. An upper bound of $q^*(S)$ is defined by

$$\gamma(S) := \begin{cases} q_0(F(S)) & : \quad \text{if } F(S) \in \Omega \\ \infty & : \quad \text{else} \end{cases}$$

Optimality Cuts. Given a local minimizer $x^* \in \Omega$ of problem (Q) we call a halfspace $H := \{x \in \mathbb{R}^n : \eta^T x \leq \gamma\}$ where $\eta \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ an optimality cut if $x^* \in \text{ int } H$ and $q^*(H) = q_0(x^*)$.

2.2. BRANCH-AND-BOUND USING OPTIMALITY CUTS

The following B&B algorithm solves problem (Q) in finite time using optimality cuts and tight lower bounds. We denote by q_{opt} the actual estimate of the optimal value and by $\mu(S)$ a tight lower bounding method.

ALGORITHM 1.

1 *determine* $S \subset \mathbb{R}^n$ *such that* $S \supset \Omega$, *compute* $\mu(S)$, *set* $\mathcal{P} = \{S\}$
   and $q_{opt} = +\infty$;
2 **repeat**
3   *choose* $\hat{S} \in \mathcal{P}$ *such that* $\mu(\hat{S}) = \min_{S \in \mathcal{P}} \mu(S)$;
4   *compute* $F(\hat{S})$;
5   **if** $\gamma(\hat{S}) \leq q_{opt}$ *and* $\gamma(\hat{S}) \neq \infty$
6   *set* $q_{opt} = \gamma(\hat{S})$;
7   *try to compute an optimality cut* H *with respect to* $F(\hat{S})$;
8   **if** *this is possible : compute for all* $S \in \mathcal{P}$: $\mu(\hat{S} \setminus int\ H)$,
9   *set* $\mathcal{P} = \{S \setminus int\ H : S \in \mathcal{P}\}$ *and* **goto** *12*
10   **endif**
11   *subdivide* $\hat{S}$ *into* $S_1, .., S_p$, *compute* $\mu(S_1), .., \mu(S_p)$
  and set $\mathcal{P} = \mathcal{P} \cup \{S_1, .., S_p\} \setminus \{\hat{S}\}$;
12   *delete elements* $S \in \mathcal{P}$ *with* $\mu(S) \geq q_{opt}$.
13 **until** $\mathcal{P} = \emptyset$.

PROPOSITION 1. *Assume that* $\Omega \neq \emptyset$, *the lower bounding method* $\mu(S)$ *is tight and the subdivision method of Algorithm 1 is exhaustive. Assume further that Assumption 1 is fulfilled and it is possible to make an optimality cut with respect to* $\hat{x}$ *if* $\hat{x}$ *is a global minimizer of problem (Q). Then Algorithm 1 terminates after finitely many iterations.*

Proof. Assume that Algorithm 1 does not terminate in finite time. Then there exists a nested subsequence of partition elements $\{S_i\}$ generated by Algorithm 1 such that $\mu(S_i)$ is the global lower bound of the corresponding partition, i.e. $\mu(S_i) = \min_{S \in \mathcal{P}_i} \mu(S)$, implying $\mu(S_i) \leq q^*$. Since the partition method is exhaustive we have $\bigcap_{i=1}^{\infty} S_i = \{\hat{x}\}$. We show now that the sequence $\{S_i\}$ is finite which proves the assertion. If $\hat{x}$ is a global minimizer of (Q) there exist $k \in \mathbb{N}$ such that $S_k \subset \Theta(\hat{x})$ due to Assumption 1 implying $F(S_k) = \hat{x}$. In this case the algorithm makes an optimality cut with respect to $\hat{x}$ proving the finiteness of $\{S_i\}$. If $\hat{x}$ is not a global minimizer then either $\hat{x} \notin \Omega$, implying $\mu(S_i) \to \infty$, or $\hat{x} \in \Omega$ and $q_0(\hat{x}) > q^*$, implying $\mu(S_i) \to q_0(\hat{x})$ since $\mu(S_i)$ is tight. In both cases it follows $\mu(S_i) > q^*$ if i is sufficiently large. This contradicts $\mu(S_i) \leq q^*$.

3. Lagrangian Relaxation

We describe now the method for obtaining lower bounds for problem (Q) based on Lagrangian relaxation.

3.1. NOTATION

Let $I \subset I\!N$ be an index set. We define by $\mathbb{R}^I$ the $|I|$-dimensional Euclidean space with the vector indexing defined by I, i.e. $\mathbb{R}^I = \{\{x_i\}_{i\in I} : x_i \in \mathbb{R}, i \in I\}$. We use the notation $A \succeq 0$ for a matrix A to be positive semidefinite. The linear space spanned by vectors $v_1, .., v_p$ is denoted by $\operatorname{span}\{v_1, .., v_p\}$, a linear space which is orthogonal to a given linear space V is denoted by $V^\perp$ and the null-space of a linear map defined by a matrix A is denoted by $\operatorname{kern} A$. Furthermore $|x|$ denotes the Euclidean norm of a vector, $\|A\|_2$ denotes the two-norm of a matrix and $\lambda_1(A)$ denotes the smallest eigenvalue of a matrix.

3.2. BASIC PRINCIPLE

The Lagrange function of problem (Q) is the quadratic form

$$L(x, \alpha) := q_0(x) + \sum_{i\in I} \alpha_i q_i(x),$$

where $I := I_{in} \cup I_{eq}$. Consider the Lagrange problem

$$\Psi(\alpha) := \min_{x\in\mathbb{R}^n} L(x, \alpha). \tag{2}$$

The dual bound is defined by

$$\Psi^* := \max_{\alpha\in\mathbb{R}^I} \quad \Psi(\alpha) \text{subject to} \alpha_i \geq 0, \quad i \in I_{in}. \tag{3}$$

By weak duality we have

$$q^* - \Psi^* \geq 0.$$

The quantity $q^* - \Psi^*$ is called duality gap. It is well-known that due to the nonconvexity of problem (Q) it is possible that a nonzero duality gap can occur. However, if (Q) is convex and satisfies a Slater condition then $q^* - \Psi^* = 0$. Note that most of the existing lower bounding methods do not have this useful property.

3.3. EQUIVALENT FORMULATIONS OF THE DUAL BOUND

Interestingly (3) can be formulated as a semidefinite program (SDP).

LEMMA 1. *Let $\hat{x} \in \mathbb{R}^n$ be an arbitrary point, $Q(\alpha, y) \in \mathbb{R}^{(n+1,n+1)}$ the matrix defined by*

$$Q(\alpha, y) := \begin{pmatrix} \frac{1}{2}\nabla_x^2 L(\hat{x}, \alpha) & \frac{1}{2}\nabla_x L(\hat{x}, \alpha) \\ \frac{1}{2}\nabla_x L(\hat{x}, \alpha)^T & L(\hat{x}, \alpha) - y \end{pmatrix}$$

and $\mathcal{S} \subset \mathbb{R}^I \times \mathbb{R}$ the set defined by

$$\mathcal{S} := \{(\alpha, y) \in \mathbb{R}^I \times \mathbb{R} : Q(\alpha, y) \succeq 0, \alpha_i \geq 0, i \in I_{in}\}. \tag{4}$$

Then

$$\Psi^* = \max_{(\alpha,y)\in\mathcal{S}} y. \tag{5}$$

Proof. Let $(\hat{\alpha}, \hat{y})$ be a solution of (5). Since $Q(\hat{\alpha}, \hat{y}) \succeq 0$ we have

$$\begin{aligned} &\begin{pmatrix} x - \hat{x} \\ 1 \end{pmatrix}^T Q(\hat{\alpha}, \hat{y}) \begin{pmatrix} x - \hat{x} \\ 1 \end{pmatrix} \\ &\quad = \frac{1}{2}(x - \hat{x})^T \nabla_x^2 L(\hat{x}, \hat{\alpha})(x - \hat{x}) + \nabla_x L(\hat{x}, \hat{\alpha})^T (x - \hat{x}) + L(\hat{x}, \hat{\alpha}) - \hat{y} \\ &\quad = L(x, \hat{\alpha}) - \hat{y} \geq 0 \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

This implies $\hat{y} \leq \min_{x\in\mathbb{R}^n} L(x, \hat{\alpha}) \leq \Psi^*$. Now let (α^*, x^*) be a solution of (3), i.e. $\Psi^* = L(x^*, \alpha^*)$. If $\Psi^* = -\infty$ we have obviously $\Psi^* \leq \hat{y}$. If $\Psi^* > -\infty$ it follows $\nabla_x^2 L(x^*, \alpha^*) \succeq 0$ and $\nabla L(x^*, \alpha^*) = 0$ implying $Q(\alpha^*, \Psi^*) = \begin{pmatrix} \frac{1}{2}\nabla_x^2 L(x^*, \alpha^*) & 0 \\ 0 & 0 \end{pmatrix} \succeq 0$. Therefore, $(\alpha^*, \Psi^*) \in \mathcal{S}$ and hence $\Psi^* \leq \hat{y}$. □

The next Lemma gives a further equivalent formulation of Ψ^*.

LEMMA 2. *Let $I_{lin} \subset I$ and $I_q \subset I$ be the index sets of linear constraints and quadratic constraints of problem (Q) respectively. Define the Lagrangian with respect to quadratic constraints*

$$L_q(x, \alpha) := q_0(x) + \sum_{i\in I_q} \alpha_i q_i(x),$$

the feasible set with respect to linear constraints

$$P := \{x \in \mathbb{R}^n : q_i(x) \leq 0, i \in I_{in} \cap I_{lin}, q_j(x) = 0, j \in I_{eq} \cap I_{lin}\}.$$

and the positive semidefinite cone

$$C := \{\alpha \in \mathbb{R}^{I_q} : \alpha_i \geq 0, i \in I_{in} \cap I_q, \nabla_x^2 L_q(x, \alpha) \succeq 0\}.$$

If $C \neq \emptyset$ then

$$\Psi^* = \max\{\min_{x\in P} L_q(x, \alpha) : \alpha \in C\}. \tag{6}$$

Proof. Let $\alpha^* = (\alpha_q^*, \alpha_l^*)$ be a solution of (3) where α_q^* corresponds to quadratic constraints and α_l^* corresponds to linear constraints of (Q). From $C \neq \emptyset$ it follows $\alpha_q^* \in C$. It holds

$$\begin{aligned}\Psi^* &= \{\max \Psi(\alpha) : \alpha_i \geqslant 0, i \in I_{in}\}\\ &= \{\max \Psi(\alpha_q^*, \alpha_l) : \alpha_i \geqslant 0, i \in I_{in} \cap I_{lin}\}\\ &= \{\max \min_{x\in\mathbb{R}^n} L(x, \alpha_q^*, \alpha_l) : \alpha_i \geq 0, i \in I_{in} \cap I_{lin}\}\\ &= \min_{x\in P} L_q(x, \alpha_q^*) \leq \Psi_q^*\end{aligned}$$

where Ψ_q^* is the optimal value of the right-hand side of (6). The last equation follows from strong duality since $L_q(x, \alpha_q^*)$ is convex (see [12]). On the other hand we have

$$\begin{aligned}\Psi_q^* &= \max\{\min_{x\in P} L_q(x, \alpha_q) : \alpha_q \in C\}\\ &= \max\{\min_{x\in\mathbb{R}^n} L(x, \alpha) : \alpha_i \geqslant 0, i \in I_{in}, \nabla_x^2 L(x, \alpha) \succeq 0\} \leq \Psi^*.\end{aligned}$$

The last equation follows from strong duality since $L_q(x, \alpha_q)$ is convex. This proves $\Psi^* = \Psi_q^*$. □

Note that by including quadratic box constraints $(\underline{x}_i - x_i)(\bar{x}_i - x_i) \leqslant 0$, $1 \leqslant i \leqslant n$ into (Q) it can be shown that $C \neq \emptyset$ (see Lemma 4).

3.4. COMPUTING Ψ^*

In the last years many methods for computing Ψ^* have been proposed. These methods can be divided into three classes. The first class are methods for solving a semidefinite program similar as (5), see for example, [9] for solution methods and applications of semidefinite programming. Most of these algorithms are interior point methods which usually converge fast if the size of the problem is moderate. However, for large scale problems the convergence can be slow if it is not possible to exploit problem structure. The second class are methods based on eigenvalue optimization. In [11] the Lagrangian relaxation bound for the max-cut problem is computed by formulating (3) as an eigenvalue optimization problem and solving this problem by the so-called spectral bundle method. In [11, 10], numerical results for large scale structured problems are presented indicating that this bundle method is faster than an interior point method. A third approach for computing Ψ^* is proposed by [24]. This approach is based on maximizing an exact penalty function over $\mathbb{R}^I$. Numerical results on this approach using the so-called r-algorithm for maximizing a nonsmooth penalty function are reported in [24]. An advantage of using nonsmooth optimization methods is that they often can exploit problem structure making them attractive for large scale structured optimization problems.

4. Improving Dual Bounds

Several methods for improving dual bounds are known (see [15]). One of the most promising method is to add redundant constraints to the original problem. We will apply this approach to the following subproblem of (Q).

4.1. THE SUBPROBLEM

Let $S \subset \mathbb{R}^n$ be a partition set such that $\Omega \cap S \neq \emptyset$. We assume in the sequel that S is a polytope defined by some linear inequalities, i.e. $S = \{x \in \mathbb{R}^n : q_i(x) \leq 0, \quad i \in I_s\}$ where q_i for $i \in I_s$ are linear functions. The quadratic program with respect to a partition set S reads

$$\text{(Q(S))} \qquad \begin{array}{rll} \text{global minimize} & q_0(x) & \\ \text{subject to} & q_i(x) \leq 0, & i \in I_{in} \cup I_s \\ & q_i(x) = 0, & i \in I_{eq}. \end{array}$$

We denote by $\Psi^*(S)$ the dual bound of (Q(S)).

4.2. ADDING REDUNDANT CONSTRAINTS

Shor proposed in [23, 24] a method for improving the Lagrangian relaxation bound by introducing redundant constraints. Let $q_i(x)$ be quadratic forms such that $q_i(x) \leq 0$ for $i \in \hat{I}_{in} \setminus (I_{in} \cup I_s)$ and $q_i(x) = 0$ for $i \in \hat{I}_{eq} \setminus I_{eq}$ for all $x \in \Omega \cap S$ where $\hat{I}_{in} \supset I_{in} \cup I_s$ and $\hat{I}_{eq} \supset I_{eq}$. Consider the following extended all-quadratic program:

$$\text{(QE(S))} \qquad \begin{array}{rll} \text{global minimize} & q_0(x) & \\ \text{subject to} & q_i(x) \leq 0, & i \in \hat{I}_{in} \\ & q_i(x) = 0, & i \in \hat{I}_{eq} \end{array}$$

We denote by $q_e^*(S)$, $L_e(x, \alpha)$ and $\Psi_e^*(S)$ the optimal value, the extended Lagrangian and the dual bound of (QE(S)) respectively.

LEMMA 3. *It holds*

$$\Psi^*(S) \leq \Psi_e^*(S) \leq q^*(S) = q_e^*(S).$$

Proof. Since $\Omega \cap S$ is not changed by the redundant constraints we have $q_e^*(S) = q^*(S)$. Denote by Ψ_e the Lagrange function of problem (QE(S)). Since $\operatorname{dom} \Psi \subset \operatorname{dom} \Psi_e$ it follows $\Psi^*(S) \leq \Psi_e^*(S)$. □

REMARK 1. *A solution of the dual problem of (QE(S)) can be used to define a so-called Lagrange heuristic for approximately computing a global minimizer of*

(QP(S)). Let $\alpha^* \in \text{argmax}\{\Psi_e(\alpha) : \alpha_i \geq 0, i \in \hat{I}_{in}\}$*. Then* $\hat{x} \in \underset{x \in S}{\text{argmin}}\, L_e(x, \alpha^*)$ *is an approximation of global minimizer of* q_0 *over* $\Omega \cap S$*. The quality of this heuristic depends strongly on the added redundant constraints (see for example, [7] and [17]).*

4.3. TIGHT LOWER BOUNDS

The B&B-algorithm presented in Section 2 requires tight lower bounds. Including quadratic box-constraints into (QE(S)) it can be shown that $\Psi_e^*(S)$ is a tight lower bound.

LEMMA 4. *Let* $S \subset \mathbb{R}^n$ *be a compact partition set. Assume that the quadratic inequality constraints*

$$(x_k - \bar{x}_k(S))(x_k - \underline{x}_k(S)) \leq 0, \quad 1 \leq k \leq n \tag{7}$$

are included in the lower bounding program (QE(S)) where $\underline{x}_i(S) := \min_{x \in S} x_i$ *and* $\bar{x}_i(S) := \max_{x \in S} x_i$ *for* $1 \leq i \leq n$*. Then* $C \neq \emptyset$ *where* C *is defined as in Lemma 2 and* $\mu(S) := \Psi_e^*(S)$ *is a tight lower bounding method.*

Proof. Choose a proper indexing of the constraints of (QE(S)) such that $\alpha = (\alpha^{(1)}, \alpha^{(2)})$ where $\alpha^{(1)} \in \mathbb{R}^n$ pertains to the constraints (7) and $\alpha^{(2)}$ pertains to the remaining constraints. Define $\gamma_0 := \max\{0, -\lambda_1(\nabla^2 q_0)\}$ and denote by $e \in \mathbb{R}^n$ the vector of ones. Then $\nabla_x^2 L_e(x, (\gamma_0 e, 0)) \succeq 0$ proving $C \neq \emptyset$. Let $\{S_k\}$ be an exhaustive nested sequence with $\bigcap_{k=1}^{\infty} = \{\hat{x}\}$. We consider firstly the case $\hat{x} \in \Omega$. Let $\tilde{x}^k := \arg\min_{x \in S_k} L_e(x, (\gamma_0 e, 0))$. From Lemma 2 it follows

$$\begin{aligned} 0 &\leq q_0(\hat{x}) - \Psi_e^*(S_k) \leq q_0(\hat{x}) - L_e(\tilde{x}^k, (\gamma_0 e, 0)) \\ &\leq |q_0(\hat{x}) - q_0(\tilde{x}^k)| + \gamma_0 \sum_{i=1}^{n} |\underline{x}_i(S_k) - \bar{x}_i(S_k)|^2. \end{aligned}$$

Hence $\lim_{k \to \infty} \Psi_e^*(S_k) = q_0(\hat{x})$ since $\tilde{x}^k$ converges towards $\hat{x}$. Now assume $\hat{x} \notin \Omega$. This implies that there exist an index $l \in I$ such that either $q_l(\hat{x}) > 0$ and $l \in I_{in}$ or $q_l(\hat{x}) \neq 0$ and $l \in I_{eq}$. Choose $\alpha_l = \gamma_1$ and $\alpha^{(1)} = \gamma_2 e$ where $\gamma_1 = \text{sign}\,(q_l(\hat{x}))\,\text{diam}\,(S_k)^{-1}$ and $\gamma_2 = \gamma_0 + \max\{0, -\lambda_1(\nabla^2 q_l)\}\,\text{diam}\,(S_k)^{-1}$ and set the remaining α_i's equal to zero. Let $\tilde{x}^k := \arg\min_{x \in S_k} L_e(x, \alpha)$. Since $\nabla^2 L_e(x, \alpha) \succeq 0$ from Lemma 2 it follows

$$\Psi_e^*(S_k) \geq L_e(\tilde{x}^k, \alpha) = q_0(\tilde{x}^k) + \gamma_1 q_l(\tilde{x}^k) + O(\,\text{diam}\,(S_k)).$$

Hence $\lim_{k\to\infty} \Psi_e^*(S_k) = \infty$ since $\tilde{x}^k$ converges towards $\hat{x}$. This proves that $\mu(S) = \Psi_e^*(S)$ is a tight lower bounding method. □

The result of Lemma 4 applies also to the general problem (Q) with (possibly) nonconvex constraints. Note that the lower bound used in the global optimization method α-BB [1] is obtained by choosing a specific value α such that $\nabla_x^2 L_e(x, \alpha) \succ 0$. >From this it follows that in general $\Psi_e^*(S)$ is more accurate than the α-BB bound.

4.4. CLOSING THE DUALITY GAP

We discuss now if the duality gap of problem (QE(S)) can be closed by adding redundant constraints. The duality gap of problem (QE(S)) is studied in [5] for special cases. If problem (QE(S)) is convex satisfying a Slater condition there is no duality gap. Also for the trust region problem with one ellipsoid constraint it is known that the duality gap is zero. However, in the presence of two ellipsoid constraints a nonzero duality gap can occur. Shor proved in [23] that problem (QE(S)) has a nonzero duality gap if and only if the objective function of an equivalent unconstrained polynomial programming problem can be represented as a sum of squares of other polynomials. However, in practice it is not known in general how to compute the polynomials. The following simple result gives a theoretical answer to the question if it is possible to close the duality gap of problem (QE(S)) by adding a single quadratic constraint.

LEMMA 5. *Assume that the inequality constraint $q^*(S) - q_0(x) \leq 0$ is included in problem (QE(S)). Then $\Psi_e^*(S) = q^*(S)$.*

Proof. Choosing the Lagrange parameter corresponding to the inequality constraint $q^*(S) - q_0(x) \leqslant 0$ equal to one and setting the remaining Lagrange parameters zero gives $L_e(x, \alpha) = q^*(S)$ implying $\Psi_e^*(S) \geq q^*(S)$. Since $\Psi_e^*(S) \leqslant q^*(S)$ the statement is proved. □

Of course, Lemma 5 is not very useful in practice since the optimal value q^* is not known in advance. The following simple global optimality criterion provides a more constructive condition.

LEMMA 6. *Let $\hat{I} := \hat{I}_{in} \cup \hat{I}_{eq}$. It holds $\Psi_e^*(S) = q^*(S)$ if and only if there exist $\hat{\alpha} \in \mathbb{R}^{\hat{I}}$ and $\hat{x} \in \Omega \cap S$ such that*

$$\hat{\alpha}_i \geq 0 \text{ for all } i \in \hat{I}_{in}, \qquad L_e(\hat{x}, \hat{\alpha}) = q_0(\hat{x}), \quad \nabla_x L_e(\hat{x}, \hat{\alpha}) = 0$$
$$\text{and } \nabla_x^2 L_e(\hat{x}, \hat{\alpha}) \succeq 0. \tag{8}$$

Proof. Let α^* be a solution of (3) and let x^* be a global minimizer of problem (Q(S)). If $\Psi_e^*(S) = q^*(S)$ it follows $x^* \in \operatorname{argmin}_{x\in\mathbb{R}^n} L_e(x, \alpha^*)$. Hence (α^*, x^*) fulfills condition (8).

Now let $(\hat{\alpha}, \hat{x})$ be a point satisfying (8). Then $\Psi_e^*(S) \leq q^*(S) \leq q_0(\hat{x}) = \min_{x \in \mathbb{R}^n} L_e(\hat{x}, \alpha) \leq \Psi_e^*(S)$. Hence $\Psi_e^*(S) = q^*(S)$. □

5. Optimality Cuts

Using the previous results we describe now a method for constructing optimality cuts for problem (Q) based on Lagrangian relaxation.

5.1. ASSUMPTIONS

For constructing an optimality cut with respect to a local minimizer x^* of problem (Q) we have to assume that x^* fulfills the following assumption.

ASSUMPTION 2. *Let x^* be a local minimizer of problem (Q). There exist corresponding Lagrange multipliers $\lambda^* \in \mathbb{R}^I$ and the Hessian $\nabla^2 q_0(x)$ is positive semidefinite over $T_{x^*}^+$, i.e.*

$$y^T \nabla^2 q_0(x) y \geq 0 \text{ for all } y \in T_{x^*}^+,$$

where the extended tangent space $T_{x^*}^+$ is defined by

$$T_{x^*}^+ := \{x \in \mathbb{R}^n : \nabla q_i(x^*)^T x = 0 \text{ for } i \in \mathcal{B}(x^*) \cup I_{eq}\}$$

and

$$\mathcal{B}(x^*) := \{i \in I_{in} : \lambda_i^* > 0\}$$

is the index set corresponding to positive Lagrange multipliers.

We give now several conditions implying Assumption 2.

LEMMA 7. *Let x^* be a local minimizer of problem (Q) and $\lambda^* \in \mathbb{R}^I$ be the corresponding Lagrange multiplier fulfilling the strict complementarity condition:*

$$\lambda_i^* > 0 \quad \text{for} \quad i \in \mathcal{A}(x^*)$$

where

$$\mathcal{A}(x^*) := \{i \in I_{in} : q_i(x^*) = 0\}$$

is the index set of active constraints. The following conditions imply Assumption 2: (i) the point x^* fulfills the modified second order optimality condition: the Hessian $\nabla^2 q_0(x)$ is positive semidefinite over the tangent space T_{x^*} defined by

$$T_{x^*} := \{x \in \mathbb{R}^n : \nabla q_i(x^*)^T x = 0 \text{ for } i \in \mathcal{A}(x^*) \cup I_{eq}\};$$

(ii) the constraints of problem (Q) are linear and x^* fulfills the second order optimality condition: the Hessian $\nabla_x^2 L(x^*, \lambda^*)$ is positive semidefinite over the tangent space T_{x^*};
(iii) the constraints of problem (Q) are linear and x^* is a regular point, i.e. the vectors $\{\nabla q_i(x^*) : i \in \mathcal{A}(x^*) \cup I_{eq}\}$ are linearly independent.

Proof. (i) From the strict complementarity condition it follows $\mathcal{A}(x^*) = \mathcal{B}(x^*)$. This implies $T_{x^*}^+ = T_{x^*}$ which proves the assertion.

(ii) Since the constraints of problem (Q) are linear it holds $\nabla_x^2 L(x, \lambda) = \nabla^2 q_0(x)$. Therefore (ii) is equivalent to (i) in this case.

(iii) Since a local minimizer which is a regular point fulfills the second order optimality condition, (iii) implies (ii). □

EXAMPLE 1. *Consider the following example:* $\min\{-x^T x : 0 \le x \le e\}$, *where $x \in \mathbb{R}^n$ and $e \in \mathbb{R}^n$ is the vector of ones. This problem has a unique global minimizer $x^* = e$ fulfilling the strict complementarity condition. From Lemma 7 (iii) it follows that x^* fulfills Assumption 2.*

5.2. THE MAIN THEOREM

In this section we present the main Theorem for deriving optimality cuts. We begin with the following result.

LEMMA 8. *Let $A \in \mathbb{R}^{(n,n)}$ be a symmetric matrix which is positive semidefinite over the linear subspace span $\{w_1, .., w_p\}^\perp$ where $w_i \in \mathbb{R}^n, 1 \le i \le p$. Then there exist $\bar{\tau} \in \mathbb{R}^p$ such that*

$A + \sum_{i=1}^{p} \tau_i w_i w_i^T$ *is positive semidefinite for all $\tau \ge \bar{\tau}$.*

Proof. Let $B := \sum_{i=1}^{p} \rho_i w_i w_i^T$ where $\rho_i > 0, 1 \le i \le p$. Let $V :=$ span $\{w_1, .., w_p\}$, $R :=$ kern (A), $S := V \cap R^\perp$ and $T := V^\perp \cap R^\perp$. Define

$$c_1 := \min_{x \in T \setminus \{0\}} \frac{x^T A x}{x^T x}, \qquad c_2 := \min_{x \in V \setminus \{0\}} \frac{x^T B x}{x^T x}, \qquad c_3 := \|A\|_2.$$

Since A is positive semidefinite on $V^\perp$ we have $c_1 > 0$ and using $x^T B x = \sum_{k=1}^{p} \rho_k (w_k^T x)^2 \ge 0$ we infer that the matrix B is positive semidefinite over $\mathbb{R}^n$ and positive definite over V implying $c_2 > 0$. Given $x \in \mathbb{R}^n$ there exist $r \in R, s \in S$ and $t \in T$ such that $x = r + s + t$ since $\mathbb{R}^n = R \bigoplus S \bigoplus T$. Therefore

$$\begin{aligned} x^T (A + \mu B) x &= (s+t)^T A (s+t) + \mu (r+s)^T B (r+s) \\ &\ge c_1 |t|^2 - 2c_3 \cdot |s| \cdot |t| - c_3 |s|^2 + \mu \cdot c_2 \cdot (|r|^2 + |s|^2) \\ &= (\sqrt{c_1}|t| - c_2/\sqrt{c_1}|s|)^2 + (\mu c_2 - c_3 - c_2^2/c_1)|s|^2. \end{aligned}$$

This implies $A + \mu_0 B \succeq 0$ where $\mu_0 = (c_3 + c_2^2/c_1)/c2$. Setting $\bar{\tau} = \mu_0 \cdot \rho$ we obtain $A + \sum_{i=1}^{p} \tau_i w_i w_i^T = A + \mu B + \sum_{i=1}^{p} (\tau_i - \bar{\tau}_i) w_i w_i^T \succeq 0$. □

A variant of Lemma 8 is presented in [14] (Debreu's Lemma). We now state the main result.

THEOREM 1. *Let x^* be a local minimizer of problem (Q) fulfilling Assumption 2. Given a partition set $S \ni x^*$ define*

$$\delta_i(S) := -\min_{x \in S} \nabla q_i(x^*)^T (x - x^*), \quad i \in \mathcal{B}(x^*). \tag{9}$$

Let $w_i := \nabla q_i(x^*)$ for $i \in \mathcal{B}(x^*) \cup I_{eq}$ and

$$A(\tau, \sigma) := \nabla^2 q_0(x^*) + \sum_{i \in \mathcal{B}(x^*)} \tau_i w_i w_i^T + \sum_{i \in I_{eq}} \sigma_i w_i w_i^T.$$

Let $\hat{\tau} \in \mathbb{R}^{\mathcal{B}(x^*)}$ and $\hat{\sigma} \in \mathbb{R}^{I_{eq}}$ be parameters fulfilling $A(\hat{\tau}, \hat{\sigma}) \succeq 0$ and $\hat{\tau}_i \geq 0$ for $i \in \mathcal{B}(x^*)$ (which exist according to Lemma 8). Assume that the constraints of problem (QE(S)) are defined by

$$w_i^T (x - x^*)(w_i^T (x - x^*) + \delta_i(S)) \leq 0, \quad i \in \mathcal{B}(x^*) \tag{10}$$

$$w_i^T (x - x^*) \leq 0, \quad i \in \mathcal{B}(x^*) \tag{11}$$

$$q_i(x) = 0, \quad i \in I_{eq} \tag{12}$$

$$q_i(x)^2 = 0, \quad i \in I_{eq} \tag{13}$$

$$q_i(x) \leq 0, \quad i \in \hat{I}_{in}. \tag{14}$$

Then

$$\Psi_e^*(S) = q^*(S) \text{ for all } S \subset S_{\hat{\tau}}$$

where

$$S_{\hat{\tau}} := \{x \in \mathbb{R}^n : 0 \geq w_i^T (x - x^*) \geq -\frac{\lambda_i^*}{\hat{\tau}_i}, \quad i \in \mathcal{B}(x^*) \text{ and } \hat{\tau}_i > 0\}.$$

Proof. Choose a proper indexing of the constraints of problem (QE) such that $\alpha = (\alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}, \alpha^{(4)}, \alpha^{(5)})$, where $\alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}, \alpha^{(4)}$ and $\alpha^{(5)}$ pertain to the constraints (10), (11), (12), (13) and (14) respectively. Assume $\alpha^{(5)} = 0$. Then it holds $L_e(x^*, \alpha) = q_0(x^*)$. From the Karush–Kuhn–Tucker condition

$$\nabla q_0(x^*) + \sum_{i \in \mathcal{B}(x^*) \cup I_{eq}} \lambda_i^* \nabla q_i(x^*) = 0$$

and from

$$\nabla L_e(x^*, \alpha) = \nabla q_0(x^*) + \sum_{i \in \mathcal{B}(x^*)} (\alpha_i^{(1)} \delta_i(S) + \alpha_k^{(2)}) \nabla q_i(x^*) + \sum_{i \in I_{eq}} \alpha_i^{(3)} \nabla q_i(x^*)$$

we obtain

$$\nabla L_e(x^*, \alpha) = \sum_{i \in \mathcal{B}(x^*)} (\alpha_i^{(1)} \delta_i(S) + \alpha_i^{(2)} - \lambda_i^*) \nabla q_i(x^*) + \sum_{i \in I_{eq}} (\alpha_i^{(3)} - \lambda_i^*) \nabla q_i(x^*).$$

Choosing $\alpha^{(1)} = \hat{\tau}$, $\alpha^{(4)} = \hat{\sigma}$, $\alpha_i^{(3)} = \lambda_i^*$ for $i \in I_{eq}$ and $\alpha_i^{(2)} = \lambda_i^* - \hat{\tau}_i \delta_i(S)$ for $i \in \mathcal{B}(x^*)$ we have $\nabla L_e(x^*, \alpha) = 0$. If $S \subset S_{\hat{\tau}}$ we have $\delta_k(S) \leq \frac{\lambda_k^*}{\hat{\tau}_k}$ for $k \in \mathcal{B}(x^*)$ and $\hat{\tau}_k > 0$ implying $\alpha_k^{(2)} = \lambda_k^* - \hat{\tau}_k \delta_k(S) \geq 0$. Hence (α, x^*) fulfills (8) and by Lemma 6 we conclude that $\Psi^*(S) = q^*(S)$. □

From Theorem 1 it follows that the dual bound $\Psi_e^*(S)$ defined by the above redundant constraints matches with $q_0(x^*)$ if the partition set S is small enough. To our knowledge, only the linear programming bound of Epperly and Swaney (1996) has this property.

5.3. COMPUTING OPTIMALITY CUTS

Based on Theorem 1 we can construct a cutting plane cutting off a given local minimizer from the feasible set. A consequence of Theorem 1 is:

COROLLARY 1. *Let x^* be a local minimizer of problem (Q) fulfilling Assumption 2 and let $H \subset \mathbb{R}^n$ be a half-space such that $x^* \in \text{int } H$ and $\Omega \cap H \subset S_{\hat{\tau}}$. Then $q^*(H) = q_0(x^*)$, where $S_{\hat{\tau}}$ is defined as in Theorem 1.*

A halfspace which meets the conditions of Corollary 1 defines an optimality cut with respect to x^*. The following Lemma gives a method for constructing H.

PROPOSITION 2. *Let x^* be a local minimizer of problem (Q) fulfilling Assumption 2 and let $\hat{\tau} \in \mathbb{R}^{\mathcal{B}(x^*)}$ and $\hat{\mu} \in \mathbb{R}^{I_{eq}}$ be parameters fulfilling $A(\hat{\tau}, \hat{\mu}) \succeq 0$ and $\hat{\tau}_i \geq 0$ for $i \in \mathcal{B}(x^*)$ where $A(\hat{\tau}, \hat{\mu})$ is defined as in Theorem 1. Then*

$$H = \{x \in \mathbb{R}^n : \eta^T (x - x^*) \leq 1\}$$

defines an optimality cut with respect to x^* where $\eta := \sum_{i \in \mathcal{B}(x^*)} -\frac{\hat{\tau}_i}{\lambda_i^*} w_i$, $w_i := \nabla q_i(x^*)$ and λ_i^* are the Lagrange parameters corresponding to x^*.

Proof. Obviously, it holds $x^* \in \text{int } H$. Let K_{x^*} be the cone defined by

$$K_{x^*} := \{x \in \mathbb{R}^n : w_i^T (x - x^*) \leq 0 \text{ for } i \in \mathcal{B}(x^*)\}.$$

Let

$$V_j = \{x \in \mathbb{R}^n : w_j^T(x - x^*) = -\delta_j^*, w_i^T(x - x^*) = 0, i \in \mathcal{B}(x^*) \setminus \{j\}\}$$

for $j \in \mathcal{B}(x^*)$ and

$$V_0 = \{x \in \mathbb{R}^n : w_i^T(x - x^*) = 0, i \in \mathcal{B}(x^*)\}$$

where $\delta_i^* = \frac{\lambda_i^*}{\hat{\tau}_i}$ if $\hat{\tau}_i > 0$ and $\delta_i^* = \infty$ else. It holds $\eta^T(x - x^*) = 1$ for $x \in V_i$ and $i \in \mathcal{B}(x^*)$ and $\eta^T(x - x^*) = 0$ for $x \in V_0$. Hence $H \cap K_{x^*} =$ conv $\{V_i : i \in \mathcal{B}(x^*) \cup \{0\}\}$ and due to $V_i \subset S_{\hat{\tau}}$ for $i \in \mathcal{B}(x^*) \cup \{0\}$ we have

$$H \cap \Omega \subset H \cap K_{x^*} \subset S_{\hat{\tau}}.$$

From Theorem 1 it follows that $q^*(S_{\hat{\tau}}) = q_0(x^*)$. Using Corollary 1 this proves the assertion. □

The parameter $\hat{\tau}$ should be computed such that diam $(S_{\hat{\tau}})$ is as large as possible. Since $\delta_i^*/|w_i|$ is an upper bound on the diameter of $S_{\hat{\tau}}$ along the direction w_i this is similar to maximizing $\delta_i^*/|w_i|$ for all $i \in \mathcal{B}(x^*)$ or to minimizing $\sum_{i \in \mathcal{B}(x^*)} \frac{1}{\delta_i^*}|w_i| = \sum_{i \in \mathcal{B}(x^*)} \frac{\hat{\tau}_i}{\lambda_i^*}|w_i|$. This motivates to compute $\hat{\tau}$ by the following semidefinite program:

$$\begin{aligned} \hat{\tau} \in \operatorname{argmin} \quad & \sum_{i \in \mathcal{B}(x^*)} \frac{\tau_i}{\lambda_i^*}|w_i| \\ s.t. \quad & A(\tau, \mu) \succeq 0 \\ & \tau_i \geq 0, \quad i \in B(x^*) \\ & \tau \in \mathbb{R}^{\mathcal{B}(x^*)}, \mu \in \mathbb{R}^{I_{eq}}. \end{aligned} \tag{15}$$

From Theorem 1 it follows that $\hat{\tau}$ is well-defined if x^* fulfills Assumption 2. Note that for the construction of an optimality cut it is sufficient to find a feasible point of (15) which is a much simpler problem than solving (15).

6. Dual Bounds with Zero Duality Gap

For special cases of problem (Q) it is possible to define an extended quadratic program which includes the redundant constraints of Theorem 1 with respect to **all** global minimizers. We define such programs for the box-constrained and the standard quadratic program. Using Theorem 1 we derive conditions which lead to a zero duality gap of the corresponding dual bound.

6.1. A DUAL BOUND FOR THE BOX-CONSTRAINED QUADRATIC PROGRAMMING PROBLEM

The box-constrained quadratic program is defined by

$$\text{(BQ)}\quad \begin{array}{ll} \text{global minimize} & q_0(x) \\ \text{subject to} & \underline{x} \le x \le \bar{x}, \end{array}$$

where $\underline{x}, \bar{x} \in \mathbb{R}^n$. Consider the following extended box-constrained quadratic program:

$$\text{(BQE)}\quad \begin{array}{ll} \text{global minimize} & q_0(x) \\ \text{subject to} & \underline{x} \le x \le \bar{x}, \\ & (x_i - \underline{x}_i)(x_i - \bar{x}_i) \le 0, \quad 1 \le i \le n. \end{array}$$

Obviously, problem (BQE) contains the redundant constraints of Theorem 1 with respect to all global minimizers of problem (BQ). From this it follows that under certain assumptions the dual bound of (BQE), denoted by Ψ^*_{bqe}, coincides with the optimal value of (BQ). More precisely, the following holds.

PROPOSITION 3. *Let x^* be a local minimizer of problem (BQ) fulfilling Assumption 2. Define $\mathcal{B}(x^*) := \{i \in \{1, .., n\} : x^*_i = \bar{x}_i$ or $x^*_i = \underline{x}_i$ and the corresponding Lagrange multiplier is greater zero$\}$. Let $g_i := \left|\frac{\partial q_0(x^*)}{\partial x_i}\right|$ for $i \in \mathcal{B}(x^*)$ and let $\hat{\tau} \in \mathbb{R}^n$ be a parameter (which exists according to Lemma 8) such that $\nabla^2 q_0 +$ diag $(\hat{\tau}) \succeq 0$, $\hat{\tau}_i \ge 0$ for $i \in B(x^*)$ and $\hat{\tau}_i = 0$ for $i \in \{1, .., n\} \setminus B(x^*)$. If*

$$g_i \ge (\bar{x}_i - \underline{x}_i)\hat{\tau}_i \quad \textit{for} \quad i \in \mathcal{B}(x^*) \tag{16}$$

then x^* is a global minimizer and $\Psi^*_{bqe} = q_0(x^*)$.

Proof. We can assume that $x^*_i = \bar{x}_i$ for all $i \in \hat{\mathcal{B}}(x^*)$. The set $S_{\hat{\tau}}$ reads $S_{\hat{\tau}} = \{x \in \mathbb{R}^n : 0 \ge e_i^T(x - x^*) \ge -\frac{\lambda^*_i}{\hat{\tau}_i}, \quad i \in \mathcal{B}(x^*)$ and $\hat{\tau}_i > 0\}$. Since $g_i = \lambda^*_i$ and $0 \ge e_i^T(x - x^*) \ge \underline{x}_i - \bar{x}_i$ for all $i \in \hat{\mathcal{B}}(x^*)$ from (16) it follows $[\underline{x}, \bar{x}] \subset S_{\hat{\tau}}$, which proves the statement due to Theorem 1. □

From Lemma 6 it follows:

COROLLARY 2. *Let $Z^* := \operatorname{argmin}_{x \in \mathbb{R}^n} L_{bqe}(x, \alpha^*)$ where L_{bqe} is the Lagrangian corresponding to (BQE) and α^* is a solution of (3). Assume there exists a local minimizer of (BQ) fulfilling the assumption of Proposition 3. Then there exist a global minimizer of (BQ) in Z^*. If (BQ) has a unique solution x^* then $Z^* = \{x^*\}$.*

This shows that all instances of problem (BQ) which fulfill the assumption of Corollary 2 can be solved by simply computing Ψ^*_{bqe}. This can be done in polynomial time and it is not necessary to compute a local minimizer. Note that assuming that Assumption 2 is fulfilled at a point x^* then condition (16) can always be satisfied if diam $([\underline{x}, \bar{x}])$ is diminished sufficiently.

EXAMPLE 2. *Consider again Example 1 :* $\min\{-x^T x : 0 \leq x \leq e\}$*, where* $x \in \mathbb{R}^n$ *and* $e \in \mathbb{R}^n$ *is the vector of ones. The unique global minimizer* $x^* = e$ *fulfills Assumption 2. Since* $\lambda^* = g = 2e$*,* $\mathcal{B}(x^*) = \{1, .., n\}$ *and* $\hat{\tau} = 2e$ *it follows that* x^* *fulfills (16).*

6.2. DUAL BOUNDS FOR THE STANDARD QUADRATIC PROGRAMMING PROBLEM

Another important quadratic program is the standard quadratic program defined by

$$\text{(SQ)}\quad \begin{array}{ll} \text{global minimize} & q_0(x) \\ \text{subject to} & 0 \leq x \leq e, \\ & e^T x - 1 = 0, \end{array}$$

where $e \in \mathbb{R}^n$ is the vector of ones. Consider the extended quadratic program

$$\text{(SQE1)}\quad \begin{array}{ll} \text{global minimize} & q_0(x) \\ \text{subject to} & 0 \leq x \leq e, \\ & x_i(x_i - 1) \leq 0, \quad 1 \leq i \leq n \\ & e^T x - 1 = 0, \\ & (e^T x - 1)^2 = 0. \end{array}$$

A different extended quadratic program for problem (SQ) is

$$\text{(SQE2)}\quad \begin{array}{ll} \text{global minimize} & q_0(x) \\ \text{subject to} & x \geq 0, \\ & x_i x_j \geq 0, \quad ij \in E_c \\ & e^T x - 1 = 0, \\ & (e^T x - 1)^2 = 0 \end{array} \tag{17}$$

where $E_c := \{ij : 1 \leq i < j \leq n, \partial_{ii} q_0(x) - 2\partial_{ij} q_0(x) + \partial_{jj} q_0(x) > 0\}$ and $\partial_{ij} q_0(x)$ denotes the second derivative of $q_0(x)$ with respect to the variables x_i and x_j. Denote by Ψ^*_{sqe1} and by Ψ^*_{sqe2} the dual bounds of (SQE1) and (SQE2) respectively. Problem (SQE1) contains the redundant constraints of Theorem 1 with respect to all global minimizers. Therefore, we can expect that a similar result as in Proposition 3 holds for problem (SQE1) and (SQE2).

PROPOSITION 4. *(i) It holds* $\Psi^*_{sqe1} \leq \Psi^*_{sqe2}$.

(ii) Let x^* *be a local minimizer of problem (SQ) fulfilling Assumption 2 and* $l \in \{1, .., n\}$ *be an index with* $x^*_l > 0$*. Define* $g_i := \frac{\partial q_0(x^*)}{\partial x_i} - \frac{\partial q_0(x^*)}{\partial x_l}$ *for* $i \in \hat{B}(x^*)$ *(where* $\mathcal{B}(x^*)$ *is defined as in Assumption 2). Let* $\hat{\tau} \in \mathbb{R}^n$ *and* $\mu \in \mathbb{R}$ *be parameters (which exist according to Lemma 8) such that* $\nabla^2 q_0 +$ *diag* $(\hat{\tau}) + \mu J \succeq 0$*,* $\hat{\tau}_i \geq 0$

for $i \in \hat{B}(x^*)$ *and* $\hat{\tau}_i = 0$ *for* $i \in \{1, .., n\} \setminus \mathcal{B}(x^*)$ *where* $J \in \mathbb{R}^{(n,n)}$ *is the matrix of ones. If*

$$g_i \geq \hat{\tau}_i \text{ for } i \in \mathcal{B}(x^*) \tag{18}$$

then x^* *is a global minimizer of problem (SQ) and* $\Psi^*_{sqe1} = \Psi^*_{sqe2} = q_0(x^*)$.

Proof. (i) Denote by $L_1(x, \alpha)$ and $L_2(x, \alpha)$ the Lagrange functions of (SQE1) and (SQE2) respectively and let $(\hat{x}, \alpha^*)$ be such that $L_1(\hat{x}, \alpha^*) = \Psi^*_{sqe1}$. >From [18] it follows that the constraints (17) can be replaced by the constraints

$$x_i x_j \geq 0, \quad 1 \leq i, j \leq n.$$

Let $x \in \mathbb{R}^n$ be a point fulfilling $e^T x - 1 = 0$. Then

$$x_i(x_i - 1) = -\sum_{1 \leq k \leq n, k \neq i} x_i x_k, \quad 1 \leq i \leq n + 1.$$

This implies that there exist $\hat{\alpha}$ such that $L_1(x, \alpha^*) = L_2(x, \hat{\alpha})$ for all $x \in \mathbb{R}^n$ with $e^T x = 1$. From Lemma 2 and Lemma 4 it follows

$$\Psi^*_{sqe1} = \min_{e^T x = 1} L_1(x, \alpha^*) = \min_{e^T x = 1} L_2(x, \hat{\alpha}) \leq \Psi^*_{sqe2}.$$

This proves the assertion.

(ii) The set $S_{\hat{\tau}}$ reads $S_{\hat{\tau}} = \{x \in \mathbb{R}^n : 0 \geq e_i^T(x - x^*) \geq -\frac{\lambda_i^*}{\hat{\tau}_i}, \quad i \in \mathcal{B}(x^*)$ and $\hat{\tau}_i > 0\}$. >From $\nabla q_0(x^*) + \sum_{i \in \mathcal{B}(x^*)} -\lambda_i^* e_i + \mu^* e = 0$ we have $\mu^* = -\frac{\partial q_0(x^*)}{\partial x_l}$ and $\lambda_i^* = \frac{\partial q_0(x^*)}{\partial x_i} + \mu^*$ for $i \in \mathcal{B}(x^*)$. Since $g_i = \lambda_i^*$ and $0 \geq e_i^T(x - x^*) \geq -1$ for all $i \in \hat{\mathcal{B}}(x^*)$ from (18) it follows $[0, e] \subset S_{\hat{\tau}}$, which proves the statement. □

Similar as in Corollary 2 it follows from Lemma 6:

COROLLARY 3. *Let* $Z_1^* := \underset{x \in \mathbb{R}^n}{\text{argmin}}\, L_{sqe1}(x, \alpha^*)$ *and* $Z_2^* := \underset{x \in \mathbb{R}^n}{\text{argmin}}\, L_{sqe2}(x, \alpha^*)$ *where* L_{sqe1} *and* L_{sqe2} *are the Lagrangian's corresponding to (SQE1) and (SQE2) respectively and* α^* *is a solution of (3). Assume there exist a local minimizer of (SQ) fulfilling the assumption of Proposition 4. Then there exist a global minimizer of (SQ) in* Z_1^* *and in* Z_2^*. *If (SQ) has a unique solution* x^* *then* $Z_1^* = Z_2^* = \{x^*\}$.

In [16] the lower bound Ψ^*_{sqe2} was computed for random examples up to 30 variables. The numerical results showed that very often Ψ^*_{sqe2} matches with the optimal value. Note that the redundant constraints of (BQE), (SQE) and (7) are also used in [22] for defining so-called RLT-relaxations of nonconvex quadratic programs.

7. Conclusion

We presented an algorithm for solving problem (Q) based on optimality cuts. If all global minimizers of problem (Q) fulfill Assumption 2 and a heuristic is used which provides a local minimizer if a partition set is diminished sufficiently then the algorithm terminates in finite time. For the implementation of the algorithm the following methods are required:

1. A heuristic $F(S)$. One possibility to define $F(S)$ is to use a Lagrange Heuristic (see Remark 1) combined with a local search method. For special cases of problem (Q) a specialized heuristic should be used (for example for quadratic integer problems).
2. A tight lower bounding method $\mu(S)$. These bounds can be computed as in Proposition 4. Very important is a fast method for solving the Lagrangian relaxation problem (3). Various methods for solving this problem exist and are currently under investigation (see Section 3.4).
3. A method for computing an optimality cut. This is based on finding a feasible point of problem (15). This can be done by checking if there exist $\bar{t} \in \mathbb{R}_+$ such that $A(\bar{t} \cdot \tilde{\tau}, \bar{t} \cdot \tilde{\mu}) \succeq 0$ where $\tilde{\tau}$ and $\tilde{\mu}$ are estimates for $\hat{\tau}$ and $\hat{\mu}$ respectively (see Proposition 2), which is equivalent to check if there exist $\bar{t} \in \mathbb{R}_+$ such that $\lambda_1(A + \bar{\tau} B) \geq 0$ where $A, B \in \mathbb{R}^{(n,n)}$ and B is a positive semidefinite matrix. From Lemma 8 it follows that $\bar{t}$ exists if and only if the corresponding local minimizer x^* fulfills Assumption 2.

We are currently implementing the lower bounding technique and the method for obtaining optimality cuts in a B&B algorithm. Numerical results will be published in a subsequent paper.

Acknowledgments

I would like to thank Prof. N. Thoai for fruitful discussions and two anonymous referees for valuable suggestions.

References

1. Adjiman, C.S., Androulakis, I.P. and C.A. Floudas, (1998), 'A global optimization method, αBB, for general twice-differentiable constrained NLPs — I. Theoretical advances'. *Computers and Chemical Engineering*, 1137–1158.
2. Al-Khayyal, F.A., Larsen, C. and van Voorhis, T. (1995), A Relaxation Method for Nonconvex Quadratically Constrained Quadratic Programs. *J. Global Opt*. 6: 215–230.
3. An, L.T.H. and Tao, P.D. (1996), 'Solving a Class of Linearly Constrained Indefinite Quadratic Problems by D.C. Algorithms. *J. of Global Opt.*, 11: 253–285.
4. An, L.T.H. and Tao, P.D. (1998), 'A Branch and Bound Method via d.c. Optimization Algorithms and Ellipsoidal Technique for Box Constrained Nonconvex Quadratic Problems'. *J. of Global Opt.*, 13: 171–206.
5. Anstreicher, K. and Wolkowicz, H. (1998), 'On Lagrangian Relaxation of Quadratic Matrix Constraints'. *Research Report CORR 98-24, Department of Combinatorics and Optimization,University of Waterloo*.

6. Audet, C., Hansen, P., Jaumard, B. and Savard, G. (1997), 'A Branch and Cut Algorithm for Nonconvex Quadratically Constrained Quadratic Programming'. Technical report, Les Cashiers du GERAD G-97-67, Montréal, 1997.
7. Feltenmark, S. and Kiwiel, K.C. (1999), 'Dual Applications of Proximal Bundle Methods, including Lagrangian relaxation of nonconvex problems'. *to appear in SIAM Journal on Optimization*.
8. Floudas, C.A. and Visweswaran, V. (1995), 'Quadratic Optimization'. In R. Horst and P. Pardalos, (eds.), *Handbook of Global Optimization*, pp. 217–269. Kluwer Academic Publishers.
9. Helmberg, C. (1999), *http://www.zib.de/helmberg/semidef.html*.
10. Helmberg, C. and Kiwiel, K.C. (1999), 'A Spectral Bundle Method with Bounds'. *Preprint SC 99-37, Konrad-Zuse-Zentrum für Informationstechnik, Berlin*, 1999.
11. Helmberg, C. and Rendl, F. (1997), 'A Spectral Bundle Method for Semidefinite Programming'. *Preprint SC 97-37, Konrad-Zuse-Zentrum für Informationstechnik, Berlin*.
12. Horst, R., Pardalos, P. and Thoai, N. (1995), *Introduction to Global Optimization*. Kluwer Academic Publishers, 1995.
13. Horst, R. and Thoai, N.V. (1996), 'A New Algorithm for Solving the General Quadratic Programming Problem. *Comp. Opt. Appl.*, 5: 39–48.
14. Lemaréchal, C. and Renaud, A. (1996), Dual-equivalent convex and non-convex problems. Technical report, INRIA, Rocquencourt, 1996.
15. Lemaréchal, C. and Renaud, A. (1999), 'A geometric study of duality gaps, with applications'. *to appear in Mathematical Programming*.
16. Nowak, I. (1998a), 'A Global Optimality Criterion for Nonconvex Quadratic Programming over a Simplex. Technical report, Preprint 98-18, Humboldt University Berlin.
17. Nowak, I. (1998b) 'Some Heuristics and Test Problems for Nonconvex Quadratic Programming over a Simplex'. Technical report, Preprint 98-17, Humboldt University Berlin.
18. Nowak, I. (1999), 'A new semidefinite programming bound for indefinite quadratic forms over a simplex'. *J. Global Opt.*, 14: 357–364.
19. Pardalos, P., Glick, J.H. and Rosen, J.B. (1987), 'Global Minimization of Indefinite Quadratic Problems'. *Computing*, 39: 281–291.
20. Phillips, A.T. and Rosen, J.B. (1990), 'Guaranteed ϵ-approximate Solution for Indefinite Quadratic Global Minimization. *Naval Research Logistics*, 37: 499–514.
21. Raber, U. (1998), 'A Simplicial Branch and Bound Method for Solving Nonconvex All-Quadratic Programs'. *J. Global Opt.*, 13: 417–432.
22. Sherali, H.D. and Tuncbilek, C.H., (1995), 'A Reformulation-Convexification Approach for Solving Nonconvex Quadratic Programming Problems'. *J. of Global Optimization*, 7: 1–31.
23. Shor, N.Z. (1992), 'Dual Estimates in Multiextremal Problems'. *J. of Global Opt.*, 2: 411–418.
24. Shor, N.Z. (1998), *Nondifferentiable Optimization and Polynomial Problems*. Kluwer Academic Publishers.
25. Spellucci, P. (1993), *Numerische Verfahren der nichtlinearen Optimierung*. Birkhäuser Verlag, ISNM, Darmstadt.
26. Thoai, N. (1997), 'A decomposition method using duality bounds for nonconvex optimization'. Technical report, Research Report 97-24, University of Trier, 1997.
27. Thoai, N. (2000), Duality bound method for the general quadratic programming problem with quadratic constraints. *Journal of Optimization Theory and Applications*, 107(2), 2000.