



First- and Second-Order Optimality Conditions for Quadratically Constrained Quadratic Programming Problems

Fabián Flores-Bazán¹ · Giandomenico Mastroeni² 

Received: 1 May 2021 / Accepted: 4 March 2022 / Published online: 31 March 2022
© The Author(s) 2022, corrected publication 2022

Abstract

We consider a quadratic programming problem with quadratic cone constraints and an additional geometric constraint. Under suitable assumptions, we establish necessary and sufficient conditions for optimality of a KKT point and, in particular, we characterize optimality by using strong duality as a regularity condition. We consider in details the case where the feasible set is defined by two quadratic equality constraints and, finally, we analyse simultaneous diagonalizable quadratic problems, where the Hessian matrices of the involved quadratic functions are all diagonalizable by means of the same orthonormal matrix.

Keywords Karush–Kuhn–Tucker conditions · Duality · Quadratic optimization

Mathematics Subject Classification 90C20 · 90C46

1 Introduction

We analyse a quadratic programming problem with general quadratic cone constraints and an additional geometric constraint. This problem has received attention in the literature in the last decades (see, e.g. [5, 7, 18, 20]) since it contains as a particular case several classic optimization problems as trust region problems, the standard quadratic

Communicated by Massimo Pappalardo.

✉ Giandomenico Mastroeni
giandomenico.mastroeni@unipi.it

Fabián Flores-Bazán
fflores@ing-mat.udec.cl

¹ Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Casilla 160-C, Concepción, Chile

² Department of Computer Science, University of Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy

problem and the max cut problem; moreover, it has many applications in robust optimization under matrix norm data uncertainty and in the field of biology and economics [12].

In this paper, we are interested in establishing necessary or sufficient global optimality conditions for a point that fulfils the Karush–Kuhn–Tucker (KKT) conditions or under the assumption of strong duality on the given problem. The general formulation of the considered quadratic programming problem allows us to treat simultaneously quadratic problems with one or more quadratic equality or inequality constraints and possibly additional constraints that can be included in the geometric one, which makes the analysis of the given problem very general, particularly as regards the possibility of providing equivalent formulations and associating a dual problem with the given one. Our approach allows to recover or generalize several known results in the literature [13, 14, 20].

The paper is organized as follows. In Sect. 2 we recall the main definitions and preliminary results that will be used throughout the paper. In Sect. 3, we characterize global optimality for a KKT point or in the presence of the property of strong duality on the given problem and in Sect. 4, we consider in details the case where the feasible set is defined by two quadratic equality constraints. In Sect. 5 we analyse a simultaneous diagonalizable quadratic problem (*SDQP*), where the Hessian matrices of the involved quadratic functions are all diagonalizable by means of the same orthonormal matrix S . The analysis previously developed allows us to provide suitable conditions that guarantee the existence of a convex reformulation of *SDQP* improving some results stated in [15] in the presence of two quadratic inequality constraints.

2 Preliminary Results

Let us recall the basic notations and preliminary results that will be used throughout the paper. Given $C \subseteq \mathbb{R}^n$, $\text{co } C$, $\text{int } C$, $\text{ri } C$, $\text{cl } C$, $\text{span } C$, denote the convex hull of C , the topological interior of C , the relative interior, the closure of C and the smallest vector linear subspace containing C , respectively. C is said to be a cone if $tC \subseteq C, \forall t \geq 0$. A convex cone C is called pointed if $C \cap (-C) = \{0\}$. We define cone $C := \bigcup_{t \geq 0} tC$. We set $\mathbb{R}_+^m := \{x \in \mathbb{R}^m : x \geq 0\}$. If C is a convex set and $x \in C$, the normal cone to C at $\bar{x} \in C$ is defined by $N_C(\bar{x}) := \{\xi \in \mathbb{R}^n : \langle \xi, x - \bar{x} \rangle \leq 0, \forall x \in C\}$. The positive polar of a set $C \subseteq \mathbb{R}^n$ is defined by $C^* := \{y^* \in \mathbb{R}^n : \langle y^*, x \rangle \geq 0, \forall x \in C\}$. It is well known that

$$\begin{aligned} C^* &= (\text{cl } C)^* = (\text{co } C)^* = (\text{cone } C)^*, \quad \text{cl co}(\text{cone } C) \\ &= \text{cl cone}(\text{co } C) = C^{**} := (C^*)^*. \end{aligned} \tag{1}$$

$C^\perp := \{v \in \mathbb{R}^n : v^\top x = 0, \forall x \in C\}$ is the orthogonal subspace to the set C . The contingent cone $T(C; \bar{x})$ of C at $\bar{x} \in C$ is the set of all $v \in \mathbb{R}^n$ such that there exist sequences $(x_k, t_k) \in C \times \mathbb{R}_+$ with $x_k \rightarrow \bar{x}$ and $t_k(x_k - \bar{x}) \rightarrow v$.

Let $P \subseteq \mathbf{R}^m$ be a convex cone and $C \subseteq \mathbf{R}^n$ a convex set. A function $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ is said P -convex on C if for every $x_1, x_2 \in C$ and for every $\lambda \in [0, 1]$,

$$\lambda f(x_1) + (1 - \lambda)f(x_2) - f(\lambda x_1 + (1 - \lambda)x_2) \in P.$$

For $m = 1$ and $P = \mathbf{R}_+$, we recover the classic definition of a convex function. It is known that if f is P -convex on C , then the set $f(C) + P$ is convex.

In the paper we will use the following preliminary results.

Let $C := \{x \in \mathbf{R}^n : g(x) = 0\}$, where $g : \mathbf{R}^n \rightarrow \mathbf{R}$. Then, we get [6]

$$T(C; \bar{x}) = \{v \in \mathbf{R}^n : \nabla g(\bar{x})^\top v = 0\} = \nabla g(\bar{x})^\perp \text{ if } \nabla g(\bar{x}) \neq 0, \quad (2)$$

and so $[T(C; \bar{x})]^* = \mathbf{R} \nabla g(\bar{x})$; whereas if $g(x) \doteq \frac{1}{2}x^\top Bx + b^\top x + \beta$ is a quadratic function, with B being a real symmetric matrix of order n , $b \in \mathbf{R}^n$ and $\beta \in \mathbf{R}$, then

$$T(C; \bar{x}) = \{v \in \mathbf{R}^n : v^\top Bv = 0\} \text{ if } \nabla g(\bar{x}) = 0. \quad (3)$$

A symmetric matrix B is positive semidefinite on C , if $x^\top Bx \geq 0$, $\forall x \in C$.

Lemma 2.1 ([18, Lemma 3.10]) *Assume that B is an indefinite real symmetric matrix and set $Z := \{v \in \mathbf{R}^n : v^\top Bv = 0\}$. Then*

$$\text{co } Z = \mathbf{R}^n = \text{span } Z.$$

3 The General Case with Cone Quadratic Constraints

Let us consider the problem

$$\mu := \inf\{f(x) : g(x) \in -P, x \in C\}, \quad (4)$$

where P is a convex cone in \mathbf{R}^m , $g(x) := (g_1(x), \dots, g_m(x))$ and $f, g_i : \mathbf{R}^n \rightarrow \mathbf{R}$, $i = 1, \dots, m$ are quadratic functions, $C \subseteq \mathbf{R}^n$,

$$f(x) := \frac{1}{2}x^\top Ax + a^\top x + \alpha, \quad g_i(x) := \frac{1}{2}x^\top B_i x + b_i^\top x + \beta_i, \quad i = 1, \dots, m, \quad (5)$$

with A, B_i being real symmetric matrices; a, b_i being vectors in \mathbf{R}^n and $\alpha, \beta_i \in \mathbf{R}$ for $i = 1, \dots, m$. $K := \{x \in C : g(x) \in -P\}$ is the feasible set of (4). We associate

with (4) the Lagrangian function $L(\lambda, x) \doteq f(x) + \sum_{i=1}^m \lambda_i g_i(x)$ and its dual problem

$$v := \sup_{\lambda \in P^*} \inf_{x \in C} L(\lambda, x). \quad (6)$$

We say that strong duality holds for (4), if there exists $\lambda^* \in P^*$ such that

$$\inf_{x \in K} f(x) = \inf_{x \in C} L(\lambda^*, x).$$

In case (4) admits an optimal solution $\bar{x} \in K$, then the previous condition is equivalent to

$$L(\lambda^*, \bar{x}) \leq L(\lambda^*, x), \quad \forall x \in C, \quad \langle \lambda^*, g(\bar{x}) \rangle = 0, \quad g(\bar{x}) \in -P, \quad \bar{x} \in C. \tag{7}$$

Under suitable assumptions on the cone $T(C; \bar{x})$, we first establish three general results: the first and the second consider the case where \bar{x} is a KKT point and provide a sufficient optimality condition and a characterization of its optimality in the case where $P = \{0\}^m$, respectively, while the third one characterizes optimality under the assumption of strong duality.

Proposition 3.1 *Let f, g_1, \dots, g_m be quadratic functions as above. Assume that $\bar{x} \in K$ is a KKT point for (4), i.e. there exists $\lambda^* \in P^*$ such that*

$$\nabla_x L(\lambda^*, \bar{x}) \in [T(C; \bar{x})]^*, \quad \langle \lambda^*, g(\bar{x}) \rangle = 0, \tag{8}$$

and, additionally, $(K - \bar{x}) \subseteq \text{cl co } T(C; \bar{x})$. Then the following assertion holds. If $\nabla_x^2 L(\lambda^, \bar{x})$ is positive semidefinite on $K - \bar{x}$, then \bar{x} is a (global) optimal solution for problem (4).*

Proof By (8), $\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0$, for every $v \in T(C; \bar{x})$, and by (1) we obtain

$$\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0, \quad \forall v \in \text{cl co } T(C; \bar{x}).$$

The assumptions imply that

$$\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0, \quad \forall v \in (K - \bar{x}). \tag{9}$$

We note that, since the involved functions are quadratic, then, the following equality holds:

$$L(\lambda^*, x) - L(\lambda^*, \bar{x}) = \nabla_x L(\lambda^*, \bar{x})^\top (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^\top \nabla_x^2 L(\lambda^*, \bar{x})(x - \bar{x}), \quad \forall x \in \mathbb{R}^n. \tag{10}$$

Exploiting (10) and (9), for every $x \in K$, we get

$$\begin{aligned} f(x) - f(\bar{x}) &\geq f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) - f(\bar{x}) = L(\lambda^*, x) - L(\lambda^*, \bar{x}) \\ &\geq \frac{1}{2}(x - \bar{x})^\top \nabla_x^2 L(\lambda^*, \bar{x})(x - \bar{x}). \end{aligned}$$

By the previous inequalities, the assertion follows. □

Remark 3.2 Proposition 3.1 is related to Theorem 2.1 in [4] when applied to a quadratic problem. Indeed, Theorem 2.1 in [4] requires that K is a convex set and $C := \mathbb{R}^n$, which guarantees that the condition $(K - \bar{x}) \subseteq \text{cl co } T(C; \bar{x})$ is fulfilled.

Proposition 3.3 Let f, g_1, \dots, g_m be quadratic functions as above, let $P := \{0\}^m$ and $\bar{x} \in K$. Assume that

$$(K - \bar{x}) \subseteq \text{cl co } T(C; \bar{x}) \subseteq -\text{cl co } T(C; \bar{x}), \quad (11)$$

and that \bar{x} is a KKT point for (4), i.e. there exists $\lambda^* \in \mathbb{R}^m$ such that

$$\nabla_x L(\lambda^*, \bar{x}) \in [T(C; \bar{x})]^*. \quad (12)$$

Then the following conditions are equivalent:

- (a) \bar{x} is an optimal solution for the problem (4);
- (b) $\nabla_x^2 L(\lambda^*, \bar{x})$ is positive semidefinite on $K - \bar{x}$ and so on $\text{cl cone}(K - \bar{x})$.

Proof By (12), $\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0$, for every $v \in T(C; \bar{x})$ and by (1) we get,

$$\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0, \forall v \in \text{cl co } T(C; \bar{x}).$$

The second inclusion in (11) implies that

$$\nabla_x L(\lambda^*, \bar{x})^\top v = 0, \quad \forall v \in \text{cl co } T(C; \bar{x}), \quad (13)$$

and, by the first inclusion in (11), $\nabla_x L(\lambda^*, \bar{x})^\top v = 0$, for every $v \in (K - \bar{x})$. By (10) and (13), for every $x \in K$, we get

$$\begin{aligned} f(x) - f(\bar{x}) &= f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) - f(\bar{x}) = L(\lambda^*, x) - L(\lambda^*, \bar{x}) \\ &= \frac{1}{2}(x - \bar{x})^\top \nabla_x^2 L(\lambda^*, \bar{x})(x - \bar{x}). \end{aligned} \quad (14)$$

By the previous equalities, the equivalence between (a) and (b) follows. \square

Remark 3.4 Note that the second inclusion in assumption (11) is not needed for proving that (b) implies (a), as shown by Proposition 3.1.

In the following proposition we characterize optimality under the strong duality property that can be considered as a regularity condition in view of the fulfilment of the KKT conditions.

Proposition 3.5 Let f, g_1, \dots, g_m be quadratic functions as above, let $\bar{x} \in K$, and assume that

$$(C - \bar{x}) \subseteq \text{cl co } T(C; \bar{x}) \subseteq -\text{cl co } T(C; \bar{x}). \quad (15)$$

Then the following assertions are equivalent:

- (a) \bar{x} is an optimal solution for the problem (4) and strong duality holds;
- (b) there exists $\lambda^* \in P^*$ such that (8) is fulfilled and $\nabla_x^2 L(\lambda^*, \bar{x})$ is positive semidefinite on $C - \bar{x}$.

Proof Assume that (a) holds, or equivalently there exists $\lambda^* \in P^*$ such that (7) is fulfilled. Then,

$L(\lambda^*, \bar{x}) \leq L(\lambda^*, x)$, for every $x \in C$, implies that $\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0$, for every $v \in T(C; \bar{x})$ and, consequently,

$$\nabla_x L(\lambda^*, \bar{x})^\top v \geq 0, \forall v \in \text{cl co } T(C; \bar{x}). \tag{16}$$

The assumption (15) yields $\nabla_x L(\lambda^*, \bar{x})^\top v = 0$, for every $v \in \text{cl co } T(C; \bar{x})$ and, in turn,

$$\nabla_x L(\lambda^*, \bar{x})^\top v = 0, \quad \forall v \in (C - \bar{x}). \tag{17}$$

From (10) we have

$$0 \leq L(\lambda^*, x) - L(\lambda^*, \bar{x}) = \frac{1}{2}(x - \bar{x})^\top \nabla_x^2 L(\lambda^*, \bar{x})(x - \bar{x}), \quad \forall x \in C,$$

and (b) follows.

Conversely if (b) holds then (8) implies (16) and, consequently, (17).

From (10) we have

$$L(\lambda^*, x) - L(\lambda^*, \bar{x}) = \frac{1}{2}(x - \bar{x})^\top \nabla_x^2 L(\lambda^*, \bar{x})(x - \bar{x}) \geq 0, \quad \forall x \in C,$$

and, taking into account that $\langle \lambda^*, g(\bar{x}) \rangle = 0$, (a) follows. □

Remark 3.6 We note that, for the implication (b) \Rightarrow (a) in Proposition 3.5, the second inclusion in (15) is not needed: indeed, by (8) we have $\nabla_x L(\lambda^*, \bar{x})^\top (x - \bar{x}) \geq 0$, $\forall x \in C$ and $\langle \lambda^*, g(\bar{x}) \rangle = 0$, so that (10) allows us to prove (a).

Remark 3.7 Condition (15) is fulfilled under the following circumstances:

- (i) $\bar{x} \in \text{int } C$;
- (ii) C is defined by linear equalities, i.e. $C := \{x \in \mathbb{R}^n : Hx = d\}$, $H \in \mathbb{R}^{p \times n}$, $d \in \mathbb{R}^p$;
- (iii) $C := \{x \in \mathbb{R}^n : h(x) = 0\}$, where h is a quadratic function with $\nabla h(\bar{x}) = 0$ and $H := \nabla^2 h(\bar{x})$ is indefinite. In this case $T(C; \bar{x}) = C - \bar{x} = \{v \in \mathbb{R}^n : v^\top H v = 0\}$, this is a consequence of Lemma 3.1 proved in what follows. By Lemma 2.1, $\text{cl co } T(C; \bar{x}) = \mathbb{R}^n$.

Lemma 3.1 Let g_i be defined as in (5), for $i = 1, \dots, m$. Assume that $\bar{x} \in A := \{x \in \mathbb{R}^n : g_i(x) = 0, i = 1, \dots, m\}$ and set $Z_i(\bar{x}) := \{v \in \mathbb{R}^n : \nabla g_i(\bar{x})^\top v + \frac{1}{2}v^\top B_i v = 0\}$, for $i = 1, \dots, m$. Then,

$$Z(\bar{x}) := \bigcap_{i=1}^m Z_i(\bar{x}) = A - \bar{x}. \quad (18)$$

Proof Let $i \in [1, \dots, m]$. Let $v \in Z_i(\bar{x})$, then $g_i(v + \bar{x}) = g_i(\bar{x}) + \nabla g_i(\bar{x})^\top v + \frac{1}{2} v^\top B_i v = 0$, proving that $v + \bar{x} \in \{x \in \mathbb{R}^n : g_i(x) = 0\}$. Therefore, $\bigcap_{i=1}^m Z_i(\bar{x}) \subseteq A - \bar{x}$.

For the other inclusion, take any $x \in A$. Then

$$0 = g_i(x) - g_i(\bar{x}) = \nabla g_i(\bar{x})^\top (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^\top B_i (x - \bar{x}), \quad i = 1, \dots, m, \quad (19)$$

which implies $x - \bar{x} \in \bigcap_{i=1}^m Z_i(\bar{x})$. \square

Remark 3.6 leads to the following result.

Corollary 3.8 *Let f, g_1, \dots, g_m be quadratic functions as above, let $\bar{x} \in K$, and assume that $C := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = m + 1, \dots, p\}$, where g_i are convex functions, for $i = m + 1, \dots, p$.*

If there exists $\lambda^ \in P^*$ such that (8) is fulfilled and $\nabla_x^2 L(\lambda^*, \bar{x})$ is positive semidefinite on $C - \bar{x}$, then \bar{x} is an optimal solution for the problem (4) and strong duality holds.*

Proof By Proposition 3.5 and taking into account Remark 3.6, it is enough to prove that $(C - \bar{x}) \subseteq \text{cl co } T(C; \bar{x})$. The convexity of the functions $g_i, i = m + 1, \dots, p$, yields that C is convex.

Since C is convex then $T(C; \bar{x}) = \text{cl cone}(C - \bar{x})$ which implies $(C - \bar{x}) \subseteq \text{cl co } T(C; \bar{x})$ (see, e.g. [2]). \square

All the results so far obtained generalize optimality conditions for classical quadratic programming to a quadratic problem with cone constraints and a geometric constraint set. We now present suitable particular cases where our results allow to recover and generalize known optimality conditions.

We first consider the quadratic programming problem with bivalent constraints (QP1) defined by

$$\inf_{x \in K} f(x) := x^\top A x + 2a^\top x + \alpha,$$

where $K := \{x \in C : g_i(x) := x^\top B_i x + 2b_i^\top x + \beta_i = 0, i = 1, \dots, m, g_{m+j}(x) := x^\top E_{m+j} x - 1 = 0, j = 1, \dots, n\}$, $E_{m+j} = \text{diag}(e_j)$ and e_j is a vector in \mathbb{R}^n whose j th element is equal to 1 and all the other entries are equal to 0.

Let $L(\lambda, \gamma, x) := f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^n \gamma_j g_{m+j}(x)$, be the Lagrangian function associated with (QP1).

By Proposition 3.3 and Lemma 3.1 we recover Lemma 3.1 of [14] which can be stated as follows.

Proposition 3.9 *Let $C := \mathbf{R}^n$ and $\bar{x} \in K$. Assume that there exist $\lambda \in \mathbf{R}^m$ and $\gamma \in \mathbf{R}^n$ such that $\nabla_x L(\lambda, \gamma, \bar{x}) = 0$. Then \bar{x} is an optimal solution for (QP1) if and only if $\nabla_x^2 L(\lambda, \gamma, \bar{x})$ is positive semidefinite on $Z(\bar{x})$ defined by (18).*

Proof It is enough to notice that since $C = \mathbf{R}^n$, then, by Lemma 3.1, $Z(\bar{x}) = K - \bar{x}$ and, moreover, (11) is fulfilled. Proposition 3.3 allows us to complete the proof. \square

By Proposition 3.5 we obtain the following result.

Next result is inspired by Theorem 3.1 of [14] and provides a characterization and a sufficient condition for strong duality for (QP1).

Proposition 3.10 *Let $\bar{x} \in K$ with $C := \mathbf{R}^n$. Consider the following assertions:*

- (a) \bar{x} is an optimal solution for (QP1) and strong duality holds;
- (b) there exist $\lambda \in \mathbf{R}^m$ and $\gamma \in \mathbf{R}^n$ such that $\nabla_x L(\lambda, \gamma, \bar{x}) = 0$ and $\nabla_x^2 L(\lambda, \gamma, \bar{x})$ is positive semidefinite;
- (c) $A - \text{diag}(\bar{X}A\bar{x} + \bar{X}a)$ is positive semidefinite, where $\bar{X} := \text{diag}(\bar{x}_1, \dots, \bar{x}_n)$.

Then (c) \Rightarrow (b) \Leftrightarrow (a).

Proof (b) \Leftrightarrow (a); it follows from Proposition 3.5 with $C := \mathbf{R}^n$, $K := \{x \in \mathbf{R}^n : g_i(x) = 0, i = 1, \dots, m + n\}$, $P := \{0\}^{m+n}$.

(c) \Rightarrow (b); in the proof of Theorem 3.1 of [14] it is shown that, for any feasible point \bar{x} , the condition $\nabla_x L(\lambda, \gamma, \bar{x}) = 0$ is fulfilled with $\lambda := (0, \dots, 0)^\top$ and $\gamma := (\bar{X}A\bar{x} + \bar{X}a)$ and, moreover, for such λ and γ , $\nabla_x^2 L(\lambda, \gamma, \bar{x}) = A - \text{diag}(\bar{X}A\bar{x} + \bar{X}a)$. Therefore, if (c) holds, then (b) is fulfilled and so is (a), by the previous part of the proof. \square

Conditions (11) and (15) in general are not fulfilled for a problem with bivalent constraints.

Example 3.11 Let $C := \{x \in \mathbf{R}^2 : x_1^2 = 1\}$, $K := \{x \in \mathbf{R}^2 : x_1^2 = 1, x_2^2 = 1\}$, $\bar{x} = (1, 1) \in K$. Then, $T(C, \bar{x}) = \{x \in \mathbf{R}^2 : x_1 = 0\} = \text{cl co } T(C; \bar{x})$,

$$K - \bar{x} = \{(0, 0), (0, -2), (-2, -2), (0, -2)\} \not\subseteq \text{cl co } T(C; \bar{x}).$$

This also implies that $C - \bar{x} \not\subseteq \text{cl co } T(C; \bar{x})$ so that Propositions 3.3 and 3.5 in general cannot be applied to problem (QP1).

Let us make some further comparison with the literature; until the end of this section we assume that $f, g_i, i = 1, \dots, m$, are quadratic functions defined as in (5). According to Remark 3.7, the following results are all particular cases of Proposition 3.5.

Corollary 3.12 ([13] Theorem 2.1, [20] Theorem 1) *Consider the problem*

$$\mu := \inf\{f(x) : g_1(x) \leq 0, \dots, g_m(x) \leq 0, x \in C\}, \tag{20}$$

where $C := \{x \in \mathbf{R}^n : Hx = d\}$, H is a $(p \times n)$ matrix, and let \bar{x} be feasible for (20).

The following assertions are equivalent:

- (a) \bar{x} is an optimal solution and strong duality holds for (20);
- (b) there exists $\lambda^* \in \mathbf{R}_+^m$ such that $\nabla_x L(\bar{x}, \lambda^*) \in H^\top(\mathbf{R}^p)$, $\lambda_i^* g_i(\bar{x}) = 0$, $i = 1, \dots, m$, and $\nabla_x^2 L(\bar{x}, \lambda^*)$ is positive semidefinite on $\text{Ker } H$.
Consequently, when $C := \mathbf{R}^n$, then (b) reduces to the following:
- (b') there exists $\lambda^* \in \mathbf{R}_+^m$ such that $\nabla_x L(\bar{x}, \lambda^*) = 0$, $\lambda_i^* g_i(\bar{x}) = 0$, $i = 1, \dots, m$ and $\nabla_x^2 L(\bar{x}, \lambda^*)$ is positive semidefinite.

4 The Case with Two Quadratic Equality Constraints

In this section we analyse in details a quadratic problem with two quadratic equality constraints defined by

$$\mu := \inf \{f(x) : g_1(x) = 0, g_2(x) = 0\}, \tag{21}$$

where $f, g_i, i = 1, 2$ are quadratic functions defined as in (5).

Let $K := \{x \in \mathbf{R}^n : g_1(x) = 0, g_2(x) = 0\}$.

The standard Lagrangian associated with (21) $L_S : \mathbf{R}^2 \times \mathbf{R}^n \mapsto \mathbf{R}$ is given by

$$L_S(\lambda_1, \lambda_2, x) := f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x).$$

The following result is a consequence of Proposition 3.3.

Proposition 4.1 *Let f, g_1, g_2 be defined as above, let $\bar{x} \in K$ be a KKT point for (21), i.e. there exists $\lambda_1, \lambda_2 \in \mathbf{R}$ such that $\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) = 0$. Then the following conditions are equivalent:*

- (a) \bar{x} is an optimal solution for (21);
- (b) $A + \lambda_1 B_1 + \lambda_2 B_2$ is positive semidefinite on $K - \bar{x}$.

If, additionally, $\nabla g_2(\bar{x}) = 0$ then (b) is equivalent to:

- (b1) $A + \lambda_1 B_1$ is positive semidefinite on $K - \bar{x}$.

Proof The equivalence between (a) and (b) follows from Proposition 3.3 where we set $C := \mathbf{R}^n$. Assume now that $\nabla g_2(\bar{x}) = 0$. The equality $g_2(x) - g_2(\bar{x}) = \nabla g_2(\bar{x})(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^\top \nabla^2 g_2(\bar{x})(x - \bar{x})$ yields $(x - \bar{x})^\top B_2(x - \bar{x}) = 0, \forall x \in K$. Therefore, $\nabla_x^2 L_S(\lambda_1, \lambda_2, \bar{x}) = A + \lambda_1 B_1 + \lambda_2 B_2$ is positive semidefinite on $K - \bar{x}$ if and only if (b1) holds. □

In the following we set $C := \{x \in \mathbf{R}^n : g_2(x) = 0\}$, so that $K = \{x \in C : g_1(x) = 0\}$. The dual problem and the standard dual problem associated with (21) are, respectively, defined by:

$$\nu := \sup_{\lambda_1 \in \mathbf{R}} \inf_{x \in C} \{L(\lambda_1, x)\}; \tag{22}$$

$$\nu_S := \sup_{\lambda_1, \lambda_2 \in \mathbf{R}} \inf_{x \in \mathbf{R}^n} \{L_S(\lambda_1, \lambda_2, x)\}. \tag{23}$$

We say that standard strong duality (SSD) holds for problem (21) if $\mu = \nu_S$ and problem (23) admits solution. It easy to check that $\nu_S \leq \nu \leq \mu$.

Theorem 4.1 *Let $\bar{x} \in K$ be feasible for (21) and suppose that $\mu \in \mathbf{R}$.*

- (a) *Assume that $\nabla g_2(\bar{x}) \neq 0$. Then the following assertions are equivalent*
 - (a1) *\bar{x} is an optimal solution and strong duality holds for problem (21);*
 - (a2) *$\exists \lambda_1, \lambda_2 \in \mathbf{R}$ such that $\nabla_x L_S(\lambda_1, \lambda_2, \bar{x}) = 0$ and $A + \lambda_1 B_1 + \lambda_2 B_2$ is positive semidefinite on $C - \bar{x}$ (and so on $\text{cl cone}(C - \bar{x})$).*
- (b) *Assume that $\nabla g_2(\bar{x}) = 0$, and B_2 positive (or negative) semidefinite. Then, (a1) is equivalent to*
 - (b1) *$\exists \lambda_1 \in \mathbf{R}$ and $\exists y \in \mathbf{R}^n$ s.t. $\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) + B_2 y = 0$ and $A + \lambda_1 B_1$ is positive semidefinite on $\ker B_2$.*
- (c) *Assume that $\nabla g_2(\bar{x}) = 0$, and B_2 indefinite. Then, (a1) is equivalent to*
 - (c1) *$\exists \lambda_1 \in \mathbf{R}$ s.t. $\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) = 0$ and $A + \lambda_1 B_1$ is positive semidefinite on $C - \bar{x}$ (and so on $\text{cl cone}(C - \bar{x})$).*

Proof (a): (a1) \Rightarrow (a2). By assumption there exists $\lambda_1 \in \mathbf{R}$ such that

$$f(x) + \lambda_1 g_1(x) \geq f(\bar{x}) = f(\bar{x}) + \lambda_1 g_1(\bar{x}), \forall x \in C. \tag{24}$$

Thus, $\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) \in [T(C, \bar{x})]^*$. Since $\nabla g_2(\bar{x}) \neq 0$, by (2) we get, $[T(C; \bar{x})]^* = \mathbf{R} \nabla g_2(\bar{x})$. Hence, there exists $\lambda_2 \in \mathbf{R}$ satisfying $\nabla_x L_S(\lambda_1, \lambda_2, \bar{x}) = 0$. Then, for every $x \in C$,

$$\begin{aligned} 0 &\leq f(x) + \lambda_1 g_1(x) - f(\bar{x}) = L_S(\lambda_1, \lambda_2, x) - L_S(\lambda_1, \lambda_2, \bar{x}) \\ &= \nabla_x L_S(\lambda_1, \lambda_2, \bar{x})^\top (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^\top \nabla_x^2 L_S(\lambda_1, \lambda_2, \bar{x}) (x - \bar{x}) \\ &= \frac{1}{2} (x - \bar{x})^\top \nabla_x^2 L_S(\lambda_1, \lambda_2, \bar{x}) (x - \bar{x}). \end{aligned}$$

This proves our claim. The previous equalities also show (a2) \Rightarrow (a1). (b): (a1) \Rightarrow (b1). By assumption there exists $\lambda_1 \in \mathbf{R}$ such that (24) holds. Thus, $\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) \in [T(C, \bar{x})]^*$. Since $\nabla g_2(\bar{x}) = 0$, then (3) yields $T(C; \bar{x}) = \{v \in \mathbf{R}^n : v^\top B_2 v = 0\}$ and, since B_2 is positive or negative semidefinite, then $T(C; \bar{x}) = \ker B_2 = Z_2(\bar{x}) = C - \bar{x}$, where the last equality is due to Lemma 3.1. Thus we can choose $y \in \mathbf{R}^n$ such that

$$\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) + B_2 y = 0.$$

Then, from (24) and for all $x \in C$ (which means $g_2(x) = 0$), it follows that

$$0 \leq L(\lambda_1, x) - L(\lambda_1, \bar{x})$$

$$\begin{aligned}
 &= \nabla_x L(\lambda_1, \bar{x})^\top (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^\top \nabla_x^2 L(\lambda_1, \bar{x}) (x - \bar{x}) \\
 &= -(B_2 y)^\top (x - \bar{x}) + \lambda_1 (x - \bar{x})^\top A_1 (x - \bar{x}) \\
 &= \lambda_1 (x - \bar{x})^\top A_1 (x - \bar{x}).
 \end{aligned}
 \tag{25}$$

Notice that $(B_2 y)^\top (x - \bar{x}) = 0$, since $\ker B_2 = C - \bar{x}$. These chains of equalities also show that $(b1) \Rightarrow (a1)$. $(c) : (a1) \Rightarrow (c1)$. By the above discussion, $T(C; \bar{x}) = Z_2(\bar{x})$. Lemma 2.1 yields $[T(C; \bar{x})]^* = (\text{co } Z_2(\bar{x}))^* = \{0\}$, which implies that $\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) = 0$. By using the relation (25), one concludes that $A + \lambda_1 B_1$ is positive semidefinite on $C - \bar{x}$. The same relation allows us to prove that $(c1) \Rightarrow (a1)$. \square

Necessary or sufficient optimality conditions for a quadratic problem with two quadratic inequality constraints have been obtained in [1, 18]. To the best of our knowledge, Theorem 4.1 is a new characterization of strong duality for a quadratic problem with two quadratic equality constraints.

5 Simultaneously Diagonalizable Quadratic Problems

In this section we characterize strong duality for a simultaneously diagonalizable quadratic problem with quadratic cone constraints, providing conditions that guarantee the existence of a convex reformulation. Our results generalize those obtained in [15] where two quadratic inequality constraints are considered under the assumption that the classic Slater condition is fulfilled.

Consider problem (4) and assume that the matrices A and $B_i, i = 1, \dots, m$ are simultaneously diagonalizable, i.e. there exists an orthonormal matrix S order n , such that $S^\top A S = D_0, S^\top B_i S = D_i, S^\top S = I$, where D_i are diagonal; we set $D_i = \text{diag}(\gamma_i), \gamma_i := (\gamma_{i1}, \dots, \gamma_{in})^\top, i = 0, 1, \dots, m$.

We refer to [3] for an extensive description of the applications of this problem.

Setting $y = S^\top x$, then (4) can be written as follows:

$$\tau := \inf \tilde{f}(y) \quad \text{s.t.} \quad y \in K := \{y \in C : \tilde{g}(y) \in -P\},
 \tag{26}$$

where P is a closed and convex cone in $\mathbb{R}^m, \tilde{g}(y) := (\tilde{g}_1(y), \dots, \tilde{g}_m(y))$ and $\tilde{f}, \tilde{g}_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m$ are quadratic functions,

$$\tilde{f}(y) := \frac{1}{2} y^\top D_0 y + a^\top S y + \alpha, \quad \tilde{g}_i(y) := \frac{1}{2} y^\top D_i y + b_i^\top S y + \beta_i, \quad i = 1, \dots, m.$$

We assume that $\alpha = 0$ and $C = \mathbb{R}^n$. Now, set $\frac{1}{2} y_i^2 = z_i, i = 1, \dots, n$, then $\frac{1}{2} y^\top D_i y = \gamma_i^\top z$ and (26) can be rewritten as follows:

$$\inf \left\{ \gamma_0^\top z + a^\top S y : \hat{g}(y, z) \in -P, \frac{1}{2} y_i^2 = z_i, i = 1, \dots, n, \right\}
 \tag{27}$$

where $\hat{g}_i(y, z) := \gamma_i^\top z + b_i^\top Sy + \beta_i, i = 1, \dots, m$.

Replacing the last n equality constraints with the corresponding inequalities, we obtain the following relaxation of (27) (and therefore of (26)):

$$\tau_R := \inf \left\{ \gamma_0^\top z + a^\top Sy : \hat{g}(y, z) \in -P, \frac{1}{2}y_i^2 \leq z_i, i = 1, \dots, n \right\}. \tag{28}$$

Let $L : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ be defined by $L(\lambda, y) := \frac{1}{2}y^\top D_0 y + a^\top Sy + \sum_{i=1}^m \lambda_i \tilde{g}_i(y)$ as the Lagrangian function associated with (26) and let $\sup_{\lambda \in P^*} \inf_{y \in C} L(\lambda, y)$ be the related dual problem. Similarly, let $L_R : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ be defined by

$$L_R(\lambda, \mu, y, z) := \gamma_0^\top z + a^\top Sy + \sum_{i=1}^m \lambda_i \hat{g}_i(y, z) + \sum_{i=1}^n \mu_i \left(\frac{1}{2}y_i^2 - z_i \right)$$

as the Lagrangian function associated with (28) and let

$$\sup_{\substack{\lambda \in P^* \\ \mu \in \mathbb{R}_+^n}} \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda, \mu, y, z),$$

be the corresponding dual problem.

Proposition 5.1 *The dual problems associated with (26) and (28) are equivalent, i.e.*

$$\sup_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = \sup_{\substack{\lambda \in P^* \\ \mu \in \mathbb{R}_+^n}} \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda, \mu, y, z). \tag{29}$$

Moreover, if the supremum in the right-hand side of (29) is attained at (λ^*, μ^*) , then the supremum in the left-hand side is attained at λ^* .

Proof Let us compute $\psi(\lambda, \mu) := \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda, \mu, y, z)$. Note that

$$\begin{aligned} L_R(\lambda, \mu, y, z) &= \gamma_0^\top z + a^\top Sy + \sum_{i=1}^m \lambda_i (\gamma_i^\top z + b_i^\top Sy + \beta_i) + \sum_{j=1}^n \mu_j \left(\frac{1}{2}y_j^2 - z_j \right) \\ &= \sum_{j=1}^n \gamma_{0j} z_j + a^\top Sy + \sum_{i=1}^m \lambda_i \left(\sum_{j=1}^n \gamma_{ij} z_j + b_i^\top Sy + \beta_i \right) + \sum_{j=1}^n \mu_j \left(\frac{1}{2}y_j^2 - z_j \right) \\ &= a^\top Sy + \sum_{i=1}^m \lambda_i (b_i^\top Sy + \beta_i) + \sum_{j=1}^n \frac{1}{2} \mu_j y_j^2 + \sum_{j=1}^n \left[\sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j} - \mu_j \right] z_j. \end{aligned}$$

Then, $\psi(\lambda, \mu) = \inf_{y \in \mathbb{R}^n} [a^\top Sy + \sum_{i=1}^m \lambda_i (b_i^\top Sy + \beta_i) + \sum_{j=1}^n \frac{1}{2} \mu_j y_j^2]$,

$$\text{if } \sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j} - \mu_j = 0, j = 1, \dots, n, \quad \psi(\lambda, \mu) = -\infty, \text{ otherwise.}$$

By eliminating the variables μ_j , we obtain:

$$\psi(\lambda, \mu) = \inf_{y \in \mathbb{R}^n} [a^\top Sy + \sum_{i=1}^m \lambda_i (b_i^\top Sy + \beta_i) + \sum_{j=1}^n \frac{1}{2} (\sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j}) y_j^2],$$

if $\sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j} = \mu_j \geq 0, j = 1, \dots, n, \quad \psi(\lambda, \mu) = -\infty, \text{ otherwise.}$

Now, observe that $L(\lambda, y) = \frac{1}{2} y^\top D_0 y + a^\top Sy + \sum_{i=1}^m \lambda_i [\frac{1}{2} y^\top D_i y + b_i^\top Sy + \beta_i]$

$$\begin{aligned} &= \frac{1}{2} \sum_{j=1}^n \gamma_{0j} y_j^2 + a^\top Sy + \sum_{i=1}^m \lambda_i [\frac{1}{2} \sum_{j=1}^n \gamma_{ij} y_j^2 + b_i^\top Sy + \beta_i] \\ &= a^\top Sy + \sum_{i=1}^m \lambda_i [b_i^\top Sy + \beta_i] + \sum_{j=1}^n \frac{1}{2} (\sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j}) y_j^2. \end{aligned}$$

Therefore,

$$\psi(\lambda, \mu) = \begin{cases} \inf_{y \in \mathbb{R}^n} L(\lambda, y), & \text{if } \sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j} \geq 0, j = 1, \dots, n, \\ -\infty, & \text{otherwise} \end{cases} \quad (30)$$

and

$$\sup_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = \sup_{\substack{\lambda \in P^* \\ \mu \in \mathbb{R}_+^n}} \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(y, z, \lambda, \mu)$$

provided that

$$\sum_{i=1}^m \lambda_i \gamma_{ij} + \gamma_{0j} \geq 0, j = 1, \dots, n, \text{ for some } \lambda \in P^*. \quad (31)$$

Notice that, if (31) does not hold, then $\sup_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = -\infty$, which yields (29).

The final assertion follows from (30). \square

Consider problem (28) and let

$$\begin{aligned} \hat{f}(y, z) &:= \gamma_0^\top z + a^\top Sy, \hat{h}(y, z) := (\frac{1}{2} y_1^2 - z_1, \dots, \frac{1}{2} y_n^2 - z_n), \\ G &:= (\hat{g}, \hat{h}), F := (\hat{f}, \hat{g}, \hat{h}). \end{aligned}$$

Assuming that $\tau_R \in \mathbb{R}$, following the image space approach introduced by Giannessi [10, 11], we define the extended image associated with (28) by:

$$\mathcal{E} := F(\mathbb{R}^n \times \mathbb{R}^n) - \tau_R(1, 0, 0) + (\mathbb{R}_+ \times P \times \mathbb{R}_+^n).$$

It is possible to show that since \hat{f} and \hat{g} are linear and \hat{h} is convex, then \mathcal{E} is a convex set, in fact F turns out to be a $(\mathbb{R}_+ \times P \times \mathbb{R}_+^n)$ -convex function. Many remarkable properties of a constrained extremum problem can be characterized (see [11]) by means of the set \mathcal{E} , as in the next result.

Proposition 5.2 *Assume that $\tau_R \in \mathbb{R}$ and*

$$\text{cl}(\mathcal{E}) \cap -(\mathbb{R}_+ \times \{0\} \times \{0\}) = \emptyset. \tag{32}$$

Then, $\tau = \tau_R$ if and only if $\tau = \sup_{\lambda \in P^} \inf_{y \in \mathbb{R}^n} L(\lambda, y)$, i.e. the duality gap is zero for (26).*

Proof It is known that condition (32) is equivalent to the fact that the duality gap is zero for (28) (see [17] Theorem 4.2, for a proof where it is assumed that the infimum τ_R of (28) is attained, we notice that it is still valid if merely $\tau_R \in \mathbb{R}$). Then, by Proposition 5.1, the following relations hold:

$$\tau \geq \sup_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = \sup_{\substack{\lambda \in P^* \\ \mu \in \mathbb{R}_+^n}} \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda, \mu, y, z) = \tau_R. \tag{33}$$

The proof is now straightforward. □

Condition (32) is not easy to check: next result, based on a well-known constraints qualification, provides the connections with strong duality for (26).

Proposition 5.3 *Assume that $\tau_R \in \mathbb{R}$ and that the following condition holds for (28):*

$$0 \in \text{ri}(G(\mathbb{R}^n \times \mathbb{R}^n) + (P \times \mathbb{R}_+^n)). \tag{34}$$

Then, $\tau = \tau_R$ if and only if strong duality holds for (26).

Proof We first prove that (34) implies that strong duality holds for (28): to this aim we will apply Theorem 3.6 of [8] where (34) is requested as one of the assumptions. The other one is given by the following condition:

$$0 \notin \text{ri}[\text{co}(\mathcal{E} \cup \{0\})], \tag{35}$$

where \mathcal{E} is the extended image associated with (28). We now prove that (35) is fulfilled. We have already observed that \mathcal{E} is a convex set; we claim that

$$\text{ri } \mathcal{E} = \text{ri}[\text{co}(\mathcal{E} \cup \{0\})].$$

Let us prove our claim. Notice that, since F is a continuous function then $0 \in \text{cl } \mathcal{E}$ and since \mathcal{E} is convex so is $\text{cl } \mathcal{E}$, so that

$$\text{cl } \text{co}(\mathcal{E} \cup \{0\}) \subseteq \text{cl } \mathcal{E}.$$

The reverse inclusion is obvious, so that $\text{cl } \text{co}(\mathcal{E} \cup \{0\}) = \text{cl } \mathcal{E}$; by Theorem 6.3 of [19] we prove our claim. Now, since $\tau_R \in \mathbb{R}$, by Proposition 3.1 of [8] we have

$$\mathcal{E} \cap -(\mathbb{R}_+ \times P \times \mathbb{R}_+^n) = \emptyset,$$

which implies

$$\text{ri } \mathcal{E} \cap -\text{ri } (\mathbb{R}_+ \times P \times \mathbb{R}_+^n) = \emptyset,$$

or, equivalently,

$$0 \notin \text{ri } [\mathcal{E} + (\mathbb{R}_+ \times P \times \mathbb{R}_+^n)] = \text{ri } \mathcal{E}.$$

This proves that (35) is fulfilled and that strong duality holds for (28).

Finally, Proposition 5.1 leads to the following relations:

$$\tau \geq \sup_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = \max_{\substack{\lambda \in P^* \\ \mu \in \mathbb{R}_+^n}} \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda, \mu, y, z) = \tau_R. \quad (36)$$

Assume that $\tau = \tau_R$; then the first inequality in (36) is fulfilled as equality and because of the second equality, the supremum is attained (see Proposition 5.1), i.e. strong duality holds for (26).

Conversely, if strong duality holds for (26), then $\tau = \max_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y)$ and (36) yields $\tau = \tau_R$. \square

We note that, when $\text{int } P \neq \emptyset$ the (34) collapses to the classic Slater condition.

Corollary 5.4 *Assume that $\tau_R \in \mathbb{R}$, (34) holds and \bar{y} is an optimal solution of (26). Then $\tau = \tau_R$ if and only if there exist $\lambda_i^* \in \mathbb{R}_+$, $i = 1, \dots, m$, such that:*

- (i) $D_0 \bar{y} + Sa + \sum_{i=1}^m \lambda_i^* (Sb_i + D_i \bar{y}) = 0$;
- (ii) $D_0 + \sum_{i=1}^m \lambda_i^* D_i$ is positive semidefinite.

Proof It is a direct consequence of Proposition 5.3 and Proposition 3.5. \square

Proposition 5.5 *Assume that (\bar{y}, \bar{z}) is a KKT point for (28) with $(\bar{\lambda}, \bar{\mu})$ the associated multipliers. If $\bar{\mu} > 0$, then \bar{y} is an optimal solution and strong duality holds for (26).*

Proof We first note that, since (28) is a convex problem, then the KKT conditions guarantee the optimality of (\bar{y}, \bar{z}) and $(\bar{\lambda}, \bar{\mu}, \bar{y}, \bar{z})$ is a saddle point of the Lagrangian function L_R . Moreover, if $\bar{\mu} > 0$, then the constraints $\frac{1}{2}y_j^2 - z_j \leq 0$ are active for $j = 1, \dots, n$, which yields that \bar{y} is feasible for (27) and therefore for (26), which proves that $\tau = \tau_R$ and \bar{y} is a global optimal solution for (26).

By Proposition 5.1 the following relations hold:

$$\tau \geq \sup_{\lambda \in P^*} \inf_{y \in \mathbf{R}^n} L(\lambda, y) = \sup_{\lambda \in P^*} \inf_{\substack{y \in \mathbf{R}^n \\ \mu \in \mathbf{R}_+^n \\ z \in \mathbf{R}^n}} L_R(\lambda, \mu, y, z) = L_R(\bar{\lambda}, \bar{\mu}, \bar{y}, \bar{z}) = \tau_R, \tag{37}$$

where the last two equalities follow from the fact that $(\bar{\lambda}, \bar{\mu}, \bar{y}, \bar{z})$ is a saddle point of L_R . Since $\tau = \tau_R$ then

$$\tau = \sup_{\lambda \in P^*} \inf_{y \in \mathbf{R}^n} L(\lambda, y) = \inf_{y \in \mathbf{R}^n} L(\bar{\lambda}, y),$$

where the last equality is due to Proposition 5.1, which proves that strong duality holds for (26). □

We provide a sufficient condition for (34) to be fulfilled.

Proposition 5.6 *Assume that*

- (i) $\text{cl cone}(\hat{g}(\mathbf{R}^n \times \mathbf{R}^n) + P) = \mathbf{R}^m$;
- (ii) *There exists (\hat{y}, \hat{z}) such that $\hat{g}(\hat{y}, \hat{z}) \in -P$ and $\frac{1}{2}\hat{y}_j^2 - \hat{z}_j < 0, j = 1, \dots, n$.*

Then (34) is fulfilled.

Proof Assume that (34) does not hold, i.e. $0 \notin \text{ri}(G(\mathbf{R}^n \times \mathbf{R}^n) + (P \times \mathbf{R}_+^n))$. Since $G(\mathbf{R}^n \times \mathbf{R}^n) + (P \times \mathbf{R}_+^n)$ is a convex set, by the separation theorem for convex sets (see, e.g. [19]), there exists $(\lambda^*, \mu^*) \in (\mathbf{R}^m \times \mathbf{R}^n) \setminus \{(0, 0)\}$ such that

$$\begin{aligned} \langle \lambda^*, \hat{g}(y, z) + v \rangle + \sum_{j=1}^n \mu_j^* (\frac{1}{2}y_j^2 - z_j + w_j) &\leq 0, \\ \forall (y, z) \in \mathbf{R}^n \times \mathbf{R}^n, \forall v \in P, \forall w \geq 0, \end{aligned} \tag{38}$$

where $w := (w_1, \dots, w_n)$.

Note that, since (38) must be fulfilled for every $v \in P$ and $w \geq 0$, it follows that $\lambda^* \in -P^*$ and $\mu^* \leq 0$. Moreover, by condition (i), we can easily prove that $\mu^* \neq 0$. Indeed, if $\mu^* = 0$, then $\lambda^* \neq 0$ and (38) becomes

$$\langle \lambda^*, \hat{g}(y, z) + v \rangle \leq 0, \quad \forall (y, z) \in \mathbf{R}^n \times \mathbf{R}^n, \forall v \in P,$$

which implies

$$\langle \lambda^*, t(\hat{g}(y, z) + v) \rangle \leq 0, \quad \forall t \geq 0, \forall (y, z) \in \mathbf{R}^n \times \mathbf{R}^n, \forall v \in P,$$

i.e.

$$\langle \lambda^*, v \rangle \leq 0, \quad \forall t \geq 0, \quad \forall v \in \text{cl cone}(\hat{g}(\mathbf{R}^n \times \mathbf{R}^n) + P),$$

but the previous inequality cannot hold, since $\text{cl cone}(\hat{g}(\mathbf{R}^n \times \mathbf{R}^n) + P) = \mathbf{R}^m$. Finally, because of condition (ii), setting $y := \hat{y}$, $z := \hat{z}$, $v := 0$, $w := 0$ in (38), yields

$$0 < \langle \lambda^*, \hat{g}(\hat{y}, \hat{z}) \rangle + \sum_{j=1}^n \mu_j^* \left(\frac{1}{2} \hat{y}_j^2 - \hat{z}_j \right) \leq 0,$$

a contradiction, which completes the proof. \square

In the particular case where the feasible set of (28) is defined by explicit equality and inequality constraints, i.e. $P := \{0\}_s \times \mathbf{R}_+^{m-s}$, for $0 \leq s \leq m$, we obtain a refinement of Proposition 5.3.

Proposition 5.7 *Let $P := \{0\}_s \times \mathbf{R}_+^{m-s}$, let (\bar{y}, \bar{z}) be an optimal solution of (28), $I(\bar{y}, \bar{z}) := \{i \in [s+1, \dots, m] : \hat{g}_i(\bar{y}, \bar{z}) = 0\}$, $J(\bar{y}, \bar{z}) := \{i \in [1, \dots, n] : \hat{h}_i(\bar{y}, \bar{z}) = 0\}$. Assume that there exists $d \in \mathbf{R}^n \times \mathbf{R}^n$ such that*

- (i) $\nabla \hat{g}_i(\bar{y}, \bar{z})^\top d = 0, i = 1, \dots, s, \nabla \hat{g}_i(\bar{y}, \bar{z})^\top d \leq 0, i \in I(\bar{y}, \bar{z})$;
- (ii) $\nabla \hat{h}_i(\bar{y}, \bar{z})^\top d < 0, i \in J(\bar{y}, \bar{z})$.

Then, $\tau = \tau_R$ if and only if strong duality holds for (26).

Proof We first prove that there exist $(\lambda^*, \mu^*) \in P^* \times \mathbf{R}_+^n$ such that

$$L_R(\lambda^*, \mu^*, y, z) \geq L_R(\lambda^*, \mu^*, \bar{y}, \bar{z}), \quad \forall (y, z) \in \mathbf{R}^n \times \mathbf{R}^n. \quad (39)$$

Denote by Q the feasible set of (28) and set $w := (y, z)$. Since $\bar{w} = (\bar{y}, \bar{z})$ is an optimal solution of (28), then $\langle c, d \rangle \geq 0, \forall d \in T(Q; \bar{w})$, where $c^\top := \nabla \hat{f}(\bar{y}, \bar{z}) = (a^\top S, \gamma_0^\top)$. Consider the set

$$\Gamma := \{d \in \mathbf{R}^n \times \mathbf{R}^n : \nabla \hat{g}_i(\bar{w})^\top d = 0, i = 1, \dots, s, \nabla \hat{g}_i(\bar{w})^\top d \leq 0, i \in I(\bar{w}), \nabla \hat{h}_i(\bar{w})^\top d < 0, i \in J(\bar{w})\}.$$

Note that, since $\Gamma \neq \emptyset$, then

$$\text{cl } \Gamma = \{d \in \mathbf{R}^n \times \mathbf{R}^n : \nabla \hat{g}_i(\bar{w})^\top d = 0, i = 1, \dots, s, \nabla \hat{g}_i(\bar{w})^\top d \leq 0, i \in I(\bar{w}), \nabla \hat{h}_i(\bar{w})^\top d \leq 0, i \in J(\bar{w})\}.$$

We show that $\text{cl } \Gamma = T(Q; \bar{w})$. We first prove that $\Gamma \subseteq T(Q; \bar{w})$. Let $d \in \Gamma$, $\{\alpha_k\} > 0, \alpha_k \downarrow 0$, then

$$\begin{aligned} \hat{g}_i(\bar{w} + \alpha_k d) &= \hat{g}_i(\bar{w}) + \alpha_k \nabla \hat{g}_i(\bar{w})^\top d = 0, \quad i = 1, \dots, s \\ \hat{g}_i(\bar{w} + \alpha_k d) &= \hat{g}_i(\bar{w}) + \alpha_k \nabla \hat{g}_i(\bar{w})^\top d \leq 0, \quad i \in I(\bar{w}), \end{aligned}$$

$$\hat{h}_i(\bar{w} + \alpha_k d) = \hat{h}_i(\bar{w}) + \alpha_k \nabla \hat{h}_i(\bar{w})^\top d + o(\alpha_k d), \quad i \in J(\bar{w}).$$

The third relation may be written as

$$\frac{1}{\alpha_k} [\hat{h}_i(\bar{w} + \alpha_k d)] = \nabla \hat{h}_i(\bar{w})^\top d + \frac{o(\alpha_k d)}{\alpha_k}, \quad i \in J(\bar{w}).$$

Since $\nabla \hat{h}_i(\bar{w})^\top d < 0, i \in J(\bar{w})$, then $\hat{h}_i(\bar{w} + \alpha_k d) < 0$, for k sufficiently large. Therefore, $w_k := \bar{w} + \alpha_k d \in Q$, for k sufficiently large, $w_k \rightarrow \bar{w}$ and $\frac{1}{\alpha_k}[w_k - \bar{w}] = d, \forall k$, which implies that $d \in T(Q; \bar{w})$. Since $T(Q; \bar{w})$ is closed, then $\text{cl } \Gamma \subseteq T(Q; \bar{w})$. We now prove that $T(Q; \bar{w}) \subseteq \text{cl } \Gamma$. Let $d \in T(Q; \bar{w})$, then $\exists \alpha_k > 0, \exists w_k \in Q, w_k \rightarrow \bar{w}, \alpha_k(w_k - \bar{w}) \rightarrow d$. Then, recalling that \hat{g} is linear, we have

$$\begin{aligned} 0 &= \hat{g}_i(w_k) = \nabla \hat{g}_i(\bar{w})^\top [w_k - \bar{w}], \quad i = 1, \dots, s \\ 0 &\geq \hat{g}_i(w_k) = \nabla \hat{g}_i(\bar{w})^\top [w_k - \bar{w}], \quad i \in I(\bar{w}), \\ 0 &\geq \hat{h}_i(w_k) \geq \nabla \hat{h}_i(\bar{w})^\top [w_k - \bar{w}] \quad i \in J(\bar{w}), \end{aligned}$$

where the last inequality is due to the convexity of \hat{h} . Multiplying the previous relations by α_k and taking the limit for $k \rightarrow \infty$ yields $d \in \text{cl } \Gamma$, which proves that $T(Q; \bar{w}) \subseteq \text{cl } \Gamma$. Since $T(Q; \bar{w}) = \text{cl } \Gamma$ and \bar{w} is an optimal solution of (28), then the following system is impossible:

$$\begin{aligned} \langle c, d \rangle &< 0 \\ \nabla \hat{g}_i(\bar{w})^\top d &= 0, \quad i = 1, \dots, s, \\ \nabla \hat{g}_i(\bar{w})^\top d &\leq 0, \quad i \in I(\bar{w}), \\ \nabla \hat{h}_i(\bar{w})^\top d &\leq 0, \quad i \in J(\bar{w}). \end{aligned} \tag{40}$$

Applying the Motzkin’s alternative theorem (see, e.g. [16]), we obtain that there exists a solution $(\lambda^*, \mu^*) \in P^* \times \mathbb{R}_+^n$ of the following system:

$$\begin{aligned} c + \sum_{i=1}^m \lambda_i^* \nabla \hat{g}_i(\bar{w}) + \sum_{i=1}^n \mu_i^* \nabla \tilde{h}_i(\bar{w}) &= 0 \\ \langle \lambda^*, \hat{g}(\bar{w}) \rangle = 0, \quad \langle \mu^*, \hat{h}(\bar{w}) \rangle &= 0. \end{aligned} \tag{41}$$

Finally, note that $L_R(\lambda^*, \mu^*, y, z)$ is a convex function such that $\nabla L_R(\lambda^*, \mu^*, \bar{y}, \bar{z}) = 0$, because of (41), where, we recall $\bar{w} = (\bar{y}, \bar{z})$. This implies that (\bar{y}, \bar{z}) is a global minimum point of $L_R(\lambda^*, \mu^*, y, z)$ on $\mathbb{R}^n \times \mathbb{R}^n$, which proves (39).

Since (39) and the complementarity conditions in (41) are fulfilled, then strong duality holds for (28). With the same arguments used in Proposition 5.5, we have that Proposition 5.1 leads to the relations:

$$\tau \geq \sup_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = \sup_{\lambda \in P^*} \inf_{\substack{y \in \mathbb{R}^n \\ \mu \in \mathbb{R}_+^n}} L_R(\lambda, \mu, y, z) = \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda^*, \mu^*, y, z) = \tau_R. \tag{42}$$

Assume that $\tau = \tau_R$; then, the first inequality in (42) is fulfilled as equality and because of the second equality, the supremum is attained at λ^* (see Proposition 5.1), i.e. strong duality holds for (26).

Conversely, if strong duality holds for (26), then

$$\tau = \max_{\lambda \in P^*} \inf_{y \in \mathbb{R}^n} L(\lambda, y) = \sup_{\substack{\lambda \in P^* \\ \mu \in \mathbb{R}_+^n}} \inf_{\substack{y \in \mathbb{R}^n \\ z \in \mathbb{R}^n}} L_R(\lambda, \mu, y, z) = \tau_R.$$

□

The proof is complete.

Remark 5.8 Computing explicitly the gradients of \hat{g} and \hat{h} , then (i) and (ii) of Proposition 5.7 can be written as

- (i') $(b_i^\top S, \gamma_i^\top)d = 0, i = 1, \dots, s, (b_i^\top S, \gamma_i^\top)d \leq 0, i \in I(\bar{y}, \bar{z})$;
- (ii') $(\bar{y}_i e_i^\top, -e_i^\top)d < 0, i \in J(\bar{y}, \bar{z})$, where e_i denotes the i -th unit vector in \mathbb{R}^n .

Next result relates condition (ii) of Proposition 5.6 with the assumptions of the previous proposition.

Proposition 5.9 Let $P := \{0\}_s \times \mathbb{R}_+^{m-s}$. If there exists $(\hat{y}, \hat{z}) \in \mathbb{R}^n \times \mathbb{R}^n$ such that $\hat{g}(\hat{y}, \hat{z}) \in -P$ and $\hat{h}(\hat{y}, \hat{z}) < 0$, then the assumptions (i) and (ii) of Proposition 5.7 are fulfilled.

Proof Set $d \doteq (\hat{y}, \hat{z}) - (\bar{y}, \bar{z})$. Since \hat{g} is an affine function then

$$\nabla \hat{g}_i(\bar{y}, \bar{z})^\top d = \hat{g}_i(\hat{y}, \hat{z}) - \hat{g}_i(\bar{y}, \bar{z}), \quad i = 1, \dots, m,$$

which yields (i), because $\hat{g}(\hat{y}, \hat{z}) \in -P$. Moreover, since \hat{h}_i is convex, then

$$0 > \hat{h}_i(\hat{y}, \hat{z}) - \hat{h}_i(\bar{y}, \bar{z}) \geq \nabla \hat{h}_i(\bar{y}, \bar{z})^\top d, \quad \forall i \in J(\bar{y}, \bar{z}),$$

and (ii) follows. □

Next example shows that the conditions of the previous proposition are weaker than (34).

Example 5.10 Set $n := 1, P := \mathbb{R}_+^2, \hat{g}_1(y, z) := -y - z, \hat{g}_2(y, z) := y + z, \hat{h}(y, z) := \frac{1}{2}y^2 - z$. Then,
 $G(\mathbb{R} \times \mathbb{R}) + \mathbb{R}_+^3 = \{(u, v, w) \in \mathbb{R}^3 : u \geq -y - z, v \geq y + z, w \geq \frac{1}{2}y^2 - z, (y, z) \in \mathbb{R}^2\}$

$$\subseteq \{(u, v, w) \in \mathbb{R}^3 : u + v \geq 0\}.$$

This implies that $(0, 0, 0) \notin \text{int}[G(\mathbb{R} \times \mathbb{R}) + \mathbb{R}_+^3]$, i.e. (34) is not fulfilled.

Nevertheless, the assumptions of Proposition 5.9 are fulfilled. Indeed, $(y^*, z^*) := (-1, 1)$ fulfils the inequalities:

$$\hat{g}_i(-1, 1) \leq 0, \quad i = 1, 2, \quad \hat{h}(-1, 1) < 0.$$

We note that in [15] the Slater-type condition (34) has been considered as a blanket assumption. Finally, we provide a refinement of Corollary 5.4.

Corollary 5.11 *Let $P := \{0\}_s \times \mathbf{R}_+^{m-s}$, let \bar{y} be an optimal solution of (26), $\bar{z} := (\frac{1}{2}\bar{y}_1^2, \dots, \frac{1}{2}\bar{y}_n^2)$ and assume that the assumptions (i) and (ii) of Proposition 5.7 hold. Then $\tau = \tau_R$ if and only if there exist $\lambda_i^* \in \mathbf{R}_+$, $i = 1, \dots, m$, such that:*

$$(i) \quad D_0\bar{y} + Sa + \sum_{i=1}^m \lambda_i^* (Sb_i + D_i\bar{y}) = 0;$$

$$(ii) \quad D_0 + \sum_{i=1}^m \lambda_i^* D_i \text{ is positive semidefinite.}$$

Proof Assume that $\tau = \tau_R$. Let us prove that (\bar{y}, \bar{z}) is an optimal solution of (28). Indeed, $\tilde{f}(\bar{y}) = \tau = \tau_R$ and (\bar{y}, \bar{z}) is an optimal solution of (27). Since (\bar{y}, \bar{z}) is feasible for (28) and $\gamma_0^T \bar{z} + a^T S\bar{y} = \tau = \tau_R$, then (\bar{y}, \bar{z}) is an optimal solution of (28). By Proposition 5.7, strong duality holds for (26). Conversely, if strong duality holds for (26), then $\tau = \tau_R$, as proved in Proposition 5.3. Recalling that here $C = \mathbf{R}^n$, applying Proposition 3.5 we complete the proof. \square

6 Conclusions

We have considered a quadratic programming problem with general quadratic cone constraints and an additional geometric constraint. We have established necessary and sufficient conditions for global optimality for a KKT point or in the presence of the property of strong duality, considering in details the case where the feasible set is defined by two quadratic equality constraints. As a further application, we have obtained conditions that guarantee the existence of a convex reformulation of a simultaneous diagonalizable quadratic problem.

Acknowledgements The research, for the first author, was supported in part by ANID-Chile through FONDECYT 1212004, ACE210010 and Basal FB210005.

Funding Open access funding provided by Università di Pisa within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Ai, W., Zhang, S.: Strong duality for the CDT subproblem: a necessary and sufficient condition. *SIAM J. Optim.* **19**, 1735–1756 (2009)

2. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: *Nonlinear programming, theory and algorithms*. Wiley, New Jersey (2006)
3. Ben-Tal, A., den Hertog, D.: Hidden conic quadratic representation of some nonconvex quadratic optimization problems. *Math. Program.* **143**, 1–29 (2014)
4. Bomze, I.M.: Copositivity for second-order optimality conditions in general smooth optimization problems. *Optimization* **65**(4), 779–795 (2015)
5. Bomze, I., Jeyakumar, V., Li, G.: Extended trust-region problems with one or two balls: exact copositive and Lagrangian relaxations. *J. Glob. Optim.* **56**(2), 551–569 (2018)
6. Di, S., Poliquin, R.: Contingent cone to a set defined by equality and inequality constraints at a Fréchet differentiable point. *J. Optim. Theory Appl.* **81**(3), 469–478 (1994)
7. Flores-Bazán, F., Cárcamo, G.: Strong duality and KKT conditions in nonconvex optimization with a single equality constraint and geometric constraints. *Math. Program.* **168**, 369–400 (2018)
8. Flores-Bazán, F., Mastroeni, G.: Strong duality in cone constrained nonconvex optimization, *SIAM J. Optim.* **23**, 153–169 (2013)
9. Flores-Bazán, F., Mastroeni, G.: Characterizing FJ and KKT points in nonconvex mathematical programming with applications, *SIAM J. Optim.* **25**, 647–676 (2015)
10. Giannessi, F.: Theorems of the alternative and optimality conditions. *J. Optim. Theory Appl.* **42**, 331–365 (1984)
11. Giannessi, F.: *Constrained Optimization and Image Space Analysis*. Springer, Berlin (2005)
12. Horst, R., Pardalos, P.M. (eds.): *Handbook of Global Optimization, Nonconvex Optimization and Its Applications*. Kluwer Academic, Dordrecht (1995)
13. Jeyakumar, V., Li, G.: Regularized Lagrangian duality for linearly constrained quadratic optimization and trust-region problems. *J. Glob. Optim.* **49**, 1–14 (2011)
14. Li, G.: Global quadratic optimization over bivalent constraints: necessary and sufficient global optimality conditions. *J. Optim. Theory Appl.* **152**, 710–726 (2012)
15. Locatelli, M.: Some results for quadratic problems with one or two quadratic constraints. *Op. Res. Letters* **43**, 126–131 (2015)
16. Mangasarian O.: *Nonlinear programming*, SIAM, *Classics in Applied Mathematics*, Philadelphia (1994)
17. Mastroeni, G.: Some applications of the image space analysis to the duality theory for constrained extremum problems. *J. Global Optim.* **46**, 603–614 (2010)
18. Peng, J.M., Yuan, Y.X.: Optimality conditions for the minimization of a quadratic with two quadratic constraints, *SIAM J. Optim.* **7**, 579–594 (1997)
19. Rockafellar, R.T.: *Convex analysis*. Princeton University Press, Princeton (1970)
20. Zheng, X.J., Sun, X.L., Li, D., Xu, Y.F.: On zero duality gap in nonconvex quadratic programming problems. *J. Global Optim.* **52**, 229–242 (2011)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.