Open access • Journal Article • DOI:10.1137/080736090

Alternative Theorems for Quadratic Inequality Systems and Global Quadratic Optimization — Source link

Vaithilingam Jeyakumar, G. M. Lee, Guoyin Li

Institutions: University of New South Wales

Published on: 01 Jul 2009 - Siam Journal on Optimization (Society for Industrial and Applied Mathematics)

Topics: Quadratically constrained quadratic program, Quadratic programming, Quadratic function, Binary quadratic form and Quadratic form

Alternative Theorems for Quadratic Inequality Systems and Global Quadratic Optimization*

V. Jeyakumar†, G. M. Lee‡ and G. Y. Li§

Revised Version: May 15, 2009

Abstract

We establish alternative theorems for quadratic inequality systems. Consequently, we obtain Lagrange multiplier characterizations of global optimality for classes of non-convex quadratic optimization problems. We present a generalization of Dine's theorem to a system of two homogeneous quadratic functions with a regular cone. The class of regular cones are cones K for which $(K \cup -K)$ is a subspace. As a consequence, we establish a generalization of the powerful S-lemma, which paves the way to obtain a complete characterization of global optimality for a general quadratic optimization model problem involving also a system of equality constraints in addition to a single quadratic inequality constraint. We then present an alternative theorem for a system of three non-homogeneous inequalities by way of establishing the convexity of the joint-range of three homogeneous quadratic functions using a regular cone. This yields Lagrange multiplier characterizations of global optimality for classes of trust-region type problems with two inequality constraints. Finally, we establish an alternative theorem for systems involving an arbitrary finite number of quadratic inequalities involving Z-matrices, which are matrices with non-positive off diagonal elements, and present necessary and sufficient conditions for global optimality for classes of non-convex inequality constrained quadratic optimization problems.

*Research of the first author was supported by the Australian Research Council. The second author was supported by the Korea Science and Engineering Foundation NRL Program grant funded by the Korea government (No. ROA-2008-000-20010-0).

†Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: v.jeyakumar@unsw.edu.au

‡Department of Applied Mathematics, Pukyong National University, Pusan 608-737, Korea. E-mail:gmlee@pknu.ac.kr

§Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au

1 Introduction

Mathematical problems involving quadratic inequalities arise naturally in many areas of optimization, control and engineering. For instance, they often appear in the form of trust region problems or quadratic programming problems in mathematical programming [23, 17, 18, 33]. Such problems also emerge in stability analysis of control systems in engineering [4]. Moreover, solvability problems of quadratic inequalities arise in the form of alternative theorems in continuous optimization [25].

Alternative theorems for arbitrary finite systems of linear or convex inequalities have played key roles in the development of optimality conditions for continuous optimization problems [6, 16]. Although these theorems are generally not valid for an arbitrary finite system of quadratic inequalities, recent research has established alternative theorems for quadratic systems involving two or three inequalities [9, 26, 33]. For instance, an alternative theorem of Gordan form for a strict inequality system of two homogeneous quadratic functions has been given in [33], where it was used in convergence analysis of trust-region algorithms. The S-lemma, which is a version of Farkas' lemma [15] for an inequality system of two non-homogeneous quadratic functions, has been given in [4, 25]. For an excellent recent survey of S-lemma and its applications, see [25]. These theorems mainly rely on the convexity of joint-range of homogeneous quadratic functions [3] even though the functions may be non-convex. The well known such joint-range convexity theorems are probably Dine's theorem [12, 25] and Brickman's theorem [7, 26, 20].

This paper makes the following three key contributions by exploiting hidden convexity of certain quadratic systems. (i) We generalize Dine's theorem to a system of two homogeneous quadratic functions with a regular cone. The class of regular cones are cones K for which $(K \cup -K)$ is a subspace, and it includes convex cones such as half-spaces and rays. As a consequence, we obtain a generalization of Gordan's theorem for two non-homogeneous quadratic functions, extending the corresponding theorem of [9] for two homogeneous quadratic functions. We also establish a generalization of S-lemma, which paves the way to obtain a complete characterization of global optimality for a general quadratic optimization model problem involving now a system of equality constraints in addition to a single quadratic inequality constraint. (ii) We present a Gordan type alternative theorem for a system of three non-homogeneous inequalities by way of generalizing Polyak's result on the convexity of the joint-range of three homogeneous quadratic functions [26] using a regular cone. This yields traditional (combined first and second-order) Lagrange multiplier characterizations of global optimality for classes of trust-region type problems with two inequality constraints. (iii) We also establish an alternative theorem for systems involving an arbitrary finite number of quadratic inequalities involving Z-matrices, which are matrices with non-positive off diagonal elements, and present necessary and sufficient conditions for global optimality for a class of non-convex inequality constrained quadratic optimization problems.

The outline of the paper is as follows. Section 2 presents basic results on the joint-range convexity of quadratic functions and its implication to quadratic inequality systems. Section 3 develops alternative theorems for systems of two inequalities and obtains a complete characterization of global optimality for a quadratic optimization model problem with a single inequality as well as a system of equality constraints. Section 4 establishes alternative theorems for systems of three quadratic inequalities and obtains global optimality conditions for classes of trust-region type quadratic optimization problems. Finally, Section 5 provides a theorem of the alternative of Gordan form for a system of finitely many quadratic inequalities involving Z-matrices and presents global optimality conditions for quadratic optimization problems with finitely many quadratic inequality constraints.

2 Preliminaries on Quadratic Functions

In this section, we fix the notation and recall some basic facts on quadratic functions that will be used throughout this paper. The real line is denoted by $\mathbb{R}$ and the n-dimensional Euclidean space is denoted by $\mathbb{R}^n$. Let $C \subseteq \mathbb{R}^n$. The dimension of C is denoted by $\dim(C)$. The set of all non-negative vectors of $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$, and the interior of $\mathbb{R}^n_+$ is denoted by $\mathrm{int}\mathbb{R}^n_+$. The space of all $(n \times n)$ symmetric matrices is denoted by S^n. The $(n \times n)$ identity matrix is denoted by I_n. The notation $A \succeq B$ means that the matrix $A - B$ is positive semidefinite. Moreover, the notation $A \succ B$ means the matrix $A - B$ is positive definite. The positive semidefinite cone is defined by $S^n_+ := \{M \in S^n : M \succeq 0\}$. Let $A, B \in S^n$. Denote the (trace) inner product of A and B is defined by $A \cdot B = Tr[AB] = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji}$ where a_{ij} is the (i,j) element of A and b_{ji} is the (j,i) element of B. A useful fact about the trace inner product is $A \cdot (xx^T) = x^T A x$ for all $x \in \mathbb{R}^n$ and $A \in S^n$. The set $K \subset \mathbb{R}^n$ is a cone if $\lambda K \subset K$, for each $\lambda \geq 0$. Clearly, subspaces, half-spaces, the set of all non-negative numbers, $\mathbb{R}^n_+$, and rays, $\{td : t \geq 0, d \in \mathbb{R}^n\}$ are examples of cones.

Definition 2.1. *The set $K \subset \mathbb{R}^n$ is a* ***regular cone*** *if $K \cup (-K)$ is a subspace of $\mathbb{R}^n$.*

Note that subspaces, half-spaces, rays are, in fact, regular cones. More generally, the first order cone [1], $K = S + \mathbb{R}_+ d$, where S is a subspace and $d \in \mathbb{R}^n$, is a regular cone. Moreover, a regular cone is not necessarily convex. For instance, the cones $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq 0\} \cup \mathbb{R}^2_+$ and $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 x_2 x_3 \leq 0\}$ are examples of nonconvex regular cones.

The basic and probably the most useful result on the joint-range convexity of homogeneous quadratic functions is given as follows.

Lemma 2.1. (**Dine's Theorem** [12, 25]) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2} x^T A_f x$ and $g(x) = \frac{1}{2} x^T A_g x$, where $A_f, A_g \in S^n$. Then the set $\{(\frac{1}{2} x^T A_f x, \frac{1}{2} x^T A_g x) : x \in \mathbb{R}^n\}$ is convex.*

Dine's theorem is known to fail for more than two homogeneous quadratic functions. Polyak [26] established the following joint-range convexity result for three homogeneous quadratic functions under a positive definite condition on the matrices involved.

Lemma 2.2. (**Polyak's Lemma** [26, Theorem 2.1]) *Let $n \geq 3$ and let $f, g, h : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x$, $g(x) = \frac{1}{2}x^T A_g x$ and $h(x) = \frac{1}{2}x^T A_h x$, where $A_f, A_g, A_h \in S^n$. Suppose that there exist $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ such that*

$$\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0. \tag{2.1}$$

Then the set $\{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x, \frac{1}{2}x^T A_h x) : x \in \mathbb{R}^n\}$ is convex.

Using Dine's Theorem, Yakubovich (cf [25]) obtained the following fundamental S-lemma which has played a key role in many areas of control and optimization. Note that the S-lemma is a form of the celebrated Farkas lemma [15] for a system of two quadratic inequalities.

Lemma 2.3. (**S-lemma** [25]) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$ and $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$, where $A_f, A_g \in S^n, b_f, b_g \in \mathbb{R}^n, c_f, c_g \in \mathbb{R}$. Suppose that there exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$. Then the following statements are equivalent:*
(i) $g(x) \leq 0 \Rightarrow f(x) \geq 0$.
(ii) $(\exists \lambda \geq 0)\ (\forall x \in \mathbb{R}^n)\ f(x) + \lambda g(x) \geq 0$.

The following alternative theorem of Yuan [9] for two strict inequalities can be viewed as a generalization of Gordan's theorem for linear systems to quadratic systems. This theorem turned out to be useful in the study of eigenvalue problems and convergence analysis of trust-region algorithms.

Lemma 2.4. (**Yuan's Alternative Theorem** [9, Lemma 2.3]) *Let $A_1, A_2 \in S^n$. Then, exactly one of the following two statements holds.*
(i) $(\exists x \in \mathbb{R}^n)\ \ \frac{1}{2}x^T A_1 x < 0,\ \frac{1}{2}x^T A_2 x < 0$.
(ii) $(\exists (\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\})\ (\forall x \in \mathbb{R}^n)\ \ x^T(\lambda_1 A_1 + \lambda_2 A_2)x \geq 0$.

3 Systems of Two Quadratic Inequalities

In this section we derive theorems of the alternative for inequality systems involving two quadratic functions and establish a complete characterization of global optimality of quadratic optimization problems. We obtain these results by generalizing Dine's theorem. We begin by examining homogeneous quadratic functions.

3.1 Homogeneous Quadratic Systems

We show that the joint-range of two homogeneous quadratic functions over a regular cone is convex.

Theorem 3.1. (**Generalized Dine's Theorem**) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x$ and $g(x) = \frac{1}{2}x^T A_g x$, where $A_f, A_g \in S^n$. Let K be a regular cone of $\mathbb{R}^n$. Then,*

$$\{(f(x), g(x)) : x \in K\} \text{ is convex.}$$

Proof. Let $\Omega = \{(f(x), g(x)) : x \in K\}$ and let $S = K \cup (-K)$. Then, by regularity, S is a subspace of dimension m, where $m \leq n$. Also, $\{(f(x), g(x)) : x \in -K\} = \{(f(x), g(x)) : x \in K\}$ as $(f(x), g(x)) = (f(-x), g(-x))$. So, $\Omega = \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x) : x \in S\}$. Since the dimension of the subspace S is m, we can find a matrix $Q \in \mathbb{R}^{n\times m}$ of full rank such that $\{Qa : a \in \mathbb{R}^m\} = S$. This gives us that

$$\begin{aligned} \Omega &= \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x) : x \in S\} \\ &= \{(\frac{1}{2}a^T(Q^T A_f Q)a, \frac{1}{2}a^T(Q^T A_g Q)a) : a \in \mathbb{R}^m\}. \end{aligned}$$

So, from Dine's theorem, Ω is a convex set in $\mathbb{R}^2$. □

The following example shows that Theorem 3.1 is, in general, not true for a homogeneous quadratic system with a non-regular convex cone.

Example 3.1. *Let $K = \mathbb{R}^2_+$ and let $f, g : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x) = x_1^2 + 2x_1x_2 - x_2^2$ and $g(x) = -2x_1^2 + 2x_1x_2 + x_2^2$. Then, $f(x) = \frac{1}{2}x^T A_f x$ and $g(x) = \frac{1}{2}x^T A_g x$, where*

$$A_f = \begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix} \text{ and } A_g = \begin{pmatrix} -4 & 2 \\ 2 & 2 \end{pmatrix}.$$

Clearly, $K \cup (-K) = \mathbb{R}^2_+ \cup (-\mathbb{R}^2_+)$ is not a subspace, and so, the convex cone K is not regular. Next, we show that $\Omega := \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x) : x \in K\}$ is not convex. To see this, note that $f(1, 0) = 1$, $g(1, 0) = -2$, $f(0, 1) = -1$ and $g(0, 1) = 1$. It follows that $a := (1, -2) \in \Omega$ and $b := (-1, 1) \in \Omega$. However, $\frac{a+b}{2} = (0, -1/2) \notin \Omega$. Otherwise, there exist $x_1, x_2 \geq 0$ such that $x_1^2 + 2x_1x_2 - x_2^2 = (x_1 + x_2)^2 - 2x_2^2 = 0$ and $-2x_1^2 + 2x_1x_2 + x_2^2 = -1/2$. Then $x_1 = (\sqrt{2}-1)x_2$ and hence $-1/2 = -2x_1^2 + 2x_1x_2 + x_2^2 = [1 + 2(\sqrt{2} - 1) - 2(\sqrt{2} - 1)^2]x_2^2$. Note that $1 + 2(\sqrt{2} - 1) - 2(\sqrt{2} - 1)^2 > 0$. This is a contradiction. Thus, Ω is not convex in this example.

As a consequence of Theorem 3.1, we derive a form of Gordan's theorem of the alternative for quadratic functions, extending the corresponding result of Yuan [9].

Theorem 3.2. (**Generalized Yuan's Theorem**) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x$ and $g(x) = \frac{1}{2}x^T A_g x$, where $A_f, A_g \in S^n$. Let K be a regular cone. Then exactly one of the following statements holds.*

(i) $(\exists x \in K)$ $f(x) < 0$, $g(x) < 0$.

(ii) $(\exists(\lambda_1, \lambda_2) \in \mathbb{R}^2_+\backslash\{(0, 0)\})$ $(\forall x \in K)$ $\lambda_1 f(x) + \lambda_2 g(x) \geq 0$.

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) does not hold. Then

$$\Omega \cap (-\text{int}\mathbb{R}^2_+) = \emptyset,$$

where $\Omega := \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x) : x \in K\}$. Since Ω is a convex set in $\mathbb{R}^2$, it follows from the convex separation theorem, there exists $(\lambda_1, \lambda_2) \neq (0,0)$ such that for each $(y_1, y_2) \in -\text{int}\mathbb{R}^2_+$ and for each $x \in K$,

$$\lambda_1 y_1 + \lambda_2 y_2 \leq 0, \ \lambda_1(\frac{1}{2}x^T A_f x) + \lambda_2(\frac{1}{2}x^T A_g x) \geq 0.$$

These inequalities give us $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+\backslash\{(0,0)\}$ and (ii) holds. $\square$

Theorem 3.2 was well known in the following particular cases: it was established by Yuan [9] when $K = \mathbb{R}^n$; whereas the theorem was given in [1, Corollary 3.2] when K is a first order cone.

Corollary 3.1. (**Homogeneous** S**-lemma**) *Let K be a regular cone and let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x$ and $g(x) = \frac{1}{2}x^T A_g x$, where $A_f, A_g \in S^n$. Suppose that there exists $x_0 \in K$ such that $g(x_0) < 0$. Then the following statements are equivalent:*
(i) *$g(x) \leq 0, x \in K \Rightarrow f(x) \geq 0$.*
(ii) *$(\exists \lambda \geq 0)$ $(\forall x \in K)$ $f(x) + \lambda g(x) \geq 0$.*

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) does not hold. Then, the system $g(x) \leq 0$, $f(x) < 0$, $x \in K$ has no solution, which, in turn, gives us that the system $g(x) < 0$, $f(x) < 0$, $x \in K$ also has no solution. Now, it follows from Theorem 3.2 that there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+\backslash\{(0,0)\}$ such that for all $x \in K$, $\lambda_1 f(x) + \lambda_2 g(x) \geq 0$. If $\lambda_1 = 0$ then $\lambda_2 g(x) \geq 0$ for all $x \in K$. This contradicts the assumption that there exists $x_0 \in K$ such that $g(x_0) < 0$. So, $\lambda_1 \neq 0$ and hence (ii) follows. $\square$

As another corollary, we obtain a Motzkin type alternative theorem for a system involving two homogeneous quadratic strict inequalities and a single linear inequality.

Corollary 3.2. *Let $a \in \mathbb{R}^n$ and let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x$ and $g(x) = \frac{1}{2}x^T A_g x$, where $A_f, A_g \in S^n$. Then, exactly one the following statements holds.*
(i) *$(\exists x \in \mathbb{R}^n)$ $f(x) < 0$, $g(x) < 0$, $a^T x \leq 0$.*
(ii) *$(\exists (\lambda_1, \lambda_2) \in \mathbb{R}^2_+\backslash\{(0,0)\}, \mu \geq 0)$ $(\forall x \in \mathbb{R}^n)$ $\lambda_1 f(x) + \lambda_2 g(x) + \mu a^T x \geq 0$.*

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) does not hold. If $a = 0$ then the (ii) holds by applying Theorem 3.2 with $K = \mathbb{R}^n$. Without loss of generality, we may assume that $a \neq 0$. Then, there exists $\overline{x}$ such that $a^T \overline{x} = -1 < 0$. Let $K = \{x \in \mathbb{R}^n : a^T x \leq 0\}$. Then K is a regular cone of $\mathbb{R}^n$. So, it follows from Theorem 3.2 that there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+\backslash\{(0,0)\}$ such that for all $x \in K = \{x \in \mathbb{R}^n : a^T x \leq 0\}$, $\lambda_1 f(x) + \lambda_2 g(x) \geq 0$. Thus, $[a^T x \leq 0 \Rightarrow \lambda_1 f(x) + \lambda_2 g(x) \geq 0]$. Note that $a^T \overline{x} = -1 < 0$. Now, by the S-lemma (Corollary 3.1), there exists $\mu \geq 0$ such that for all $x \in \mathbb{R}^n$ $\lambda_1 f(x) + \lambda_2 g(x) + \mu a^T x \geq 0$, i.e., (ii) holds. $\square$

3.2 Non-Homogeneous Systems

We now derive alternative theorems for systems of inequalities involving non-homogeneous quadratic functions.

Theorem 3.3. (**Nonhomogenous Yuan's Theorem**) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$ and $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$, where $A_f, A_g \in S^n$, $b_f, b_g \in \mathbb{R}^n, c_f, c_g \in \mathbb{R}$. Let $a_0 \in \mathbb{R}^n$ and let S_0 be a subspace of $\mathbb{R}^n$. Then exactly one of the following statements holds:*
(i) $(\exists x \in a_0 + S_0)\ f(x) < 0,\ g(x) < 0.$
(ii) $(\exists(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\})\ (\forall x \in a_0 + S_0)\ \lambda_1 f(x) + \lambda_2 g(x) \geq 0.$

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) does not hold. Then, the following system has no solution: $f_1(x) < 0$, $g_1(x) < 0$, $x \in S_0$, where

$$f_1(x) = f(x + a_0) = \frac{1}{2}x^T A_f x + (b_f + A_f a_0)^T x + (c_f + \frac{1}{2}a_0^T A_f a_0 + b_f^T a_0)$$

and

$$g_1(x) = g(x + a_0) = \frac{1}{2}x^T A_g x + (b_g + A_g a_0)^T x + (c_g + \frac{1}{2}a_0^T A_g a_0 + b_g^T a_0).$$

Note that $f_1(x) = \frac{1}{2}x^T \bar{A}_f x + \bar{b}_f^T x + \bar{c}_f$ and $g(x) = \frac{1}{2}x^T \bar{A}_g x + \bar{b}_g^T x + \bar{c}_g$ where $\bar{A}_f = A_f$, $\bar{b}_f = b_f + A_f a_0$, $\bar{c}_f = c_f + \frac{1}{2}a_0^T A_f a_0 + b_f^T a_0$, $\bar{A}_g = A_g$, $\bar{b}_g = b_g + A_g a_0$ and $\bar{c}_g = c_g + \frac{1}{2}a_0^T A_g a_0 + b_g^T a_0$. Define two homogeneous functions $\tilde{f}_1, \tilde{g}_1 : \mathbb{R}^{n+1} \to \mathbb{R}$ by

$$\tilde{f}_1(x,t) = \frac{1}{2}x^T \bar{A}_f x + \bar{b}_f^T xt + \bar{c}_f t^2 \text{ and } \tilde{g}_1(x,t) = \frac{1}{2}x^T \bar{A}_g x + \bar{b}_g^T xt + \bar{c}_g t^2.$$

Then, the system $\tilde{f}_1(x,t) < 0$, $\tilde{g}_1(x,t) < 0$, $(x,t) \in S_0 \times \mathbb{R}_+$ has no solution. Otherwise, there exists $(x_0, t_0) \in S_0 \times \mathbb{R}_+$ such that $\tilde{f}_1(x_0, t_0) < 0$ and $\tilde{g}_1(x_0, t_0) < 0$. If $t_0 \neq 0$, then $f(\frac{x_0}{t_0} + a_0) = f_1(\frac{x_0}{t_0}) = t_0^{-2}\tilde{f}_1(x_0, t_0) < 0$ and $g(\frac{x_0}{t_0} + a_0) = g_1(\frac{x_0}{t_0}) = t_0^{-2}\tilde{g}_1(x_0, t_0) < 0$. This contradicts the fact that the system, $f(x) < 0$, $g(x) < 0$, $x \in a_0 + S_0$, has no solution. If $t_0 = 0$, then $\frac{1}{2}x_0^T A_f x_0 = \tilde{f}_1(x_0, t_0) < 0$ and $\frac{1}{2}x_0^T A_g x_0 = \tilde{g}(x_0, t_0) < 0$. This implies that $\lim_{\alpha \to +\infty} f_1(\alpha x_0) = -\infty$ and $\lim_{\alpha \to +\infty} g_1(\alpha x_0) = -\infty$. Thus, there exists $\gamma > 0$ such that $f(\gamma x_0 + a_0) = f_1(\gamma x_0) < 0$ and $f(\gamma x_0 + a_0) = g(\gamma x_0) < 0$. This is a contradiction.

As $S_0 \times \mathbb{R}_+$ is a regular convex cone and $\tilde{f}_1, \tilde{g}_1$ are homogeneous functions. it follows from Theorem 3.2 that there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ such that for all $(x,t) \in S_0 \times \mathbb{R}_+$,

$$\lambda_1 \tilde{f}_1(x,t) + \lambda_2 \tilde{g}_1(x,t) \geq 0. \tag{3.2}$$

Thus, by setting $t = 1$ in (3.2), we see that for each $x \in S_0$ $\lambda_1 \tilde{f}_1(x,1) + \lambda_2 \tilde{g}_1(x,1) = \lambda_1 f_1(x + a_0) + \lambda_2 f_2(x + a_0) \geq 0$. Therefore, (ii) holds. $\square$

The following example illustrates that theorem 3.3 may fail for the system (i) even with a first order cone $\mathbb{R}_+ a_0 + S_0$, replacing the set $a_0 + S_0$.

Example 3.2. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x - x^2$ and $g(x) = x^2 - 1$. Let $a_0 = 1$ and $S_0 = \{0\}$. Consider the following two statements:*
(i') $(\exists x \in \mathbb{R}_+ a_0 + S_0)\ f(x) < 0,\ g(x) < 0.$
(ii') $(\exists (\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\})\ (\forall x \in \mathbb{R}_+ a_0 + S_0).\ \lambda_1 f(x) + \lambda_2 g(x) \geq 0.$
Observe that $\mathbb{R}_+ a_0 + S_0 = [0, +\infty)$ and that $x - x^2 \geq 0$ for all $x \in [0, 1]$. So, the following system has no solution: $f(x) < 0,\ g(x) < 0,\ \ x \in \mathbb{R}_+ a_0 + S_0$. Thus (i') *fails. But we see that* (ii') *also does not hold. Otherwise, there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ such that $\lambda_1 f(x) + \lambda_2 g(x) \geq 0$, for all $x \geq 0$. In particular, $0 \leq \lambda_1 f(0) + \lambda_2 g(0) = -\lambda_2$. Thus, $\lambda_2 = 0$ and $\lambda_1 > 0$ and so, $f(x) \geq 0$, for all $x \geq 0$ which is impossible.*

We now derive a generalization of S-lemma which allows us to characterize global optimality of a quadratic optimization model problem with a single quadratic constraint and a system of linear equality constraints.

Corollary 3.3. (**Generalized S-lemma**) *Let S_0 be a subspace and let $a_0 \in \mathbb{R}^n$. Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$ and $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$, where $A_f, A_g \in S^n$, $b_f, b_g \in \mathbb{R}^n, c_f, c_g \in \mathbb{R}$. Suppose that there exists $x_0 \in a_0 + S_0$ such that $g(x_0) < 0$. Then the following statements are equivalent:*
(i) $g(x) \leq 0,\ x \in a_0 + S_0 \Rightarrow f(x) \geq 0$.
(ii) $(\exists \lambda \geq 0)\ (\forall x \in a_0 + S_0)\ f(x) + \lambda g(x) \geq 0.$

Proof. As [(Ii) $\Rightarrow$ (i)] holds always, we only show [(i) $\Rightarrow$ (ii)]. Suppose that (i) holds. Then, the system $g(x) \leq 0,\ f(x) < 0,\ x \in a_0 + S_0$ has no solution, which, in turn, gives us that the system $g(x) < 0,\ f(x) < 0,\ x \in a_0 + S_0$ also has no solution. Now, it follows from Theorem 3.3 that there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ such that for all $x \in a_0 + S_0$

$$\lambda_1 f(x) + \lambda_2 g(x) \geq 0. \tag{3.3}$$

If $\lambda_1 = 0$ then $\lambda_2 g(x) \geq 0$ for all $x \in a_0 + S_0$. By assumption, there exists $x_0 \in a_0 + S_0$ such that $g(x_0) < 0$. This forces $\lambda_2 = 0$ which contradicts the fact that $(\lambda_1, \lambda_2) \neq (0,0)$. Hence, $\lambda_1 \neq 0$ and (ii) follows by dividing (3.3) by λ_1. □

3.3 Necessary and Sufficient Global Optimality Conditions

Consider the following quadratic optimization problem.

$$(QP) \qquad \min \frac{1}{2}x^T A_f x + b_f^T x + c_f$$
$$\text{s.t. } \frac{1}{2}x^T A_g x + b_g^T x + c_g \leq 0,\ \ Hx = b,$$

where $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$ and $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$, where $A_f, A_g \in S^n$, $H \in \mathbb{R}^{m \times n}$, $b_f, b_g \in \mathbb{R}^n, c_f, c_g, \in \mathbb{R}$ and $b \in \mathbb{R}^m$. We assume that a global minimizer of (QP) exists.

We now derive a complete characterization of global optimality for (QP) under the Slater condition, extending the corresponding results for (QP) without the equality constraints [22, 29].

Corollary 3.4. *For (QP), suppose that there exists x_0 such that $Hx_0 = b$ and $g(x_0) < 0$. Then a feasible point $\overline{x}$ is a global minimizer of (QP) if and only if there exist $\lambda \geq 0$ and $\mu \in \mathbb{R}^m$ such that $\nabla(f + \lambda g)(\overline{x}) + H^T\mu = 0$, $\lambda g(\overline{x}) = 0$ and $d^T(A_f + \lambda A_g)d \geq 0$ whenever $Hd = 0$.*

Proof. Let $\overline{x}$ be a global minimizer of (QP) and let $S_0 = \{x : Hx = 0\}$. Then $[g(x) \leq 0, x \in \overline{x} + S_0 \Rightarrow f(x) - f(\overline{x}) \geq 0]$. So, by the generalized S-lemma (Corollary 3.3), there exists $\lambda \geq 0$ such that for each $x \in \overline{x} + S_0$

$$f(x) + \lambda g(x) \geq f(\overline{x}). \tag{3.4}$$

This implies that $\lambda g(\overline{x}) = 0$ and so $\overline{x}$ is a global minimizer of $f + \lambda g$ over $Hx = b$. Now, by the necessary optimality conditions at $\overline{x}$, there exists $\mu \in \mathbb{R}^m$ such that $\nabla(f + \lambda g)(\overline{x}) + H^T\mu = 0$ and $d^T(A_f + \lambda A_g)^T d \geq 0$ whenever $Hd = 0$.

Conversely, suppose that there exist $\lambda \geq 0$ and $\mu \in \mathbb{R}^m$ such that $\nabla(f + \lambda g)(\overline{x}) + H^T\mu = 0$, $\lambda g(\overline{x}) = 0$ and, for each $d \in \mathbb{R}^n$ with $Hd = 0$, $d^T(A_f + \lambda A_g)d \geq 0$. Then for each feasible x of (QP),

$$\begin{aligned} & (f + \lambda g)(x) - (f + \lambda g)(\overline{x}) \\ = \; & \nabla(f + \lambda g)(\overline{x})^T(x - \overline{x}) + \frac{1}{2}(x - \overline{x})^T(A_f + \lambda A_g)(x - \overline{x}) \geq 0, \end{aligned}$$

as $\nabla(f + \lambda g)(\overline{x})^T(x - \overline{x}) = -\mu^T H(x - \overline{x}) = 0$ and $H(x - \overline{x}) = 0$. This gives us that for each feasible x of (QP),

$$f(x) + \lambda g(x) \geq f(\overline{x}) + \lambda g(\overline{x}) = f(\overline{x}),$$

which, in turn, yields $f(x) \geq f(\overline{x}) - \lambda g(x) \geq f(\overline{x})$. Hence, $f(\overline{x})$ is a global minimizer of (QP). $\square$

Below, we present an example verfying Corollary 3.4 where the Hessian of the corresponding Lagrangian function is not positive semi-definite.

Example 3.3. *Consider the following quadratic optimization problem:*

$$\begin{aligned} (E1) \qquad & \min \frac{1}{2}x^T A_f x + b_f^T x + c_f \\ & s.t. \;\; \frac{1}{2}x^T A_g x + b_g^T x + c_g \leq 0, \;\; Hx = b, \end{aligned}$$

where $n = 2$, $m = 1$, $b_f = b_g = (0,0)^T$, $c_f = 0$, $c_g = 1$, $b = 1$,

$$A_f = \begin{pmatrix} 0 & -2 \\ -2 & -2 \end{pmatrix}, A_g = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } H = (1,1).$$

This problem can be equivalently rewritten as

$$\begin{aligned} \min_{(x_1,x_2)\in\mathbb{R}^2} \quad & -x_2^2 - 2x_1x_2 \\ s.t. \quad & 1 - x_1^2 \leq 0 \\ & x_1 + x_2 = 1. \end{aligned}$$

It can be verify that global minimizers of (E1) are $\overline{x} = (1,0)$ and $\overline{z} = (-1,2)$. Let $\mu = 2$, $\lambda = 1$ and $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$ and $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$. Then

$$A_f + \lambda A_g = \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix},$$

and so, for any $d = (d_1, d_2) \in \mathbb{R}^2$ satisfying $d_1 + d_2 = 0$, one has

$$d^T(A_f + \lambda A_g)d = 0.$$

Moreover, since $g(\overline{x}) = 0$ and

$$\nabla(f + \lambda g)(\overline{x}) + H^T\mu = (A_f + \lambda A_g)\overline{x} + H^T\mu = (0,0)^T,$$

the global optimiality condition is satisfied at $\overline{x}$. Similarly, one can also verify that $g(\overline{z}) = 0$ and

$$\nabla(f + \lambda g)(\overline{z}) + H^T\mu = (A_f + \lambda A_g)\overline{z} + H^T\mu = (0,0)^T.$$

Thus, global optimiality condition is also satisfied at $\overline{z}$ and our Corollary 3.4 is verified. Finally, we note that the matrix $A_f + \lambda A_g$ is not positive semidefinte since its eigenvalues are -4 and 0.

4 Systems of Three Quadratic Inequalities

In this section we derive theorems of the alternative for systems involving three quadratic inequalities and obtain global optimality conditions for trust-region type problems. We obtain these results by generalizing Polyak's Lemma (Lemma 2.2).

4.1 Homogeneous Systems

Using the similar line of arguments as in the proof of Theorem 3.1, we derive the following generalization of Polyak's Lemma.

Theorem 4.1. (**Generalized Polyak's Lemma**) *Let $f, g, h : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x$, $g(x) = \frac{1}{2}x^T A_g x$ and $h(x) = \frac{1}{2}x^T A_h x$, where $A_f, A_g, A_h \in S^n$. Let $K \subseteq \mathbb{R}^n$ be a regular cone with $\dim(K \cup -K) \geq 3$. Suppose that there exist $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ such that $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$. Then the set $\{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x, \frac{1}{2}x^T A_h x) : x \in K\}$ is convex.*

Proof. Let $S := K \cup (-K)$. Then S is a subspace of dimension m, where $3 \leq m \leq n$. Let

$$\Omega \;=\; \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x, \frac{1}{2}x^T A_h x) : x \in K\}.$$

Using the same line of arguments as in the proof of Theorem 3.1, we can show that $\Omega = \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x, \frac{1}{2}x^T A_h x) : x \in S\}$. So, we can find a matrix $Q \in \mathbb{R}^{n\times m}$ of full rank such that $\{Qa : a \in \mathbb{R}^m\} = S$. This gives us that

$$\begin{aligned}\Omega &= \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x, \frac{1}{2}x^T A_h x) : x \in S\} \\ &= \{(\frac{1}{2}a^T(Q^T A_f Q)a, \frac{1}{2}a^T(Q^T A_g Q)a, \frac{1}{2}a^T(Q^T A_h Q)a) : a \in \mathbb{R}^m\}.\end{aligned}$$

Since there exist $\gamma_1, \gamma_2, \gamma_3$ such that $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$, for each $v \in \mathbb{R}^m\backslash\{0\}$, $Qv \neq 0$ and so,

$$v^T(\gamma_1 Q^T A_f Q + \gamma_2 Q^T A_g Q + \gamma_3 Q^T A_h Q)v = (Qv)^T(\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h)(Qv) > 0.$$

That is, $\gamma_1 Q^T A_f Q + \gamma_2 Q^T A_g Q + \gamma_3 Q^T A_h Q \succ 0$. As $m \geq 3$, it follows by Polyak's lemma that Ω is a convex set in $\mathbb{R}^3$. $\square$

As a consequence of Theorem 4.1, we derive a form of Gordan's theorem of the alternative for three quadratic functions involving a regular cone K, extending the corresponding result of Polyak [26] where $K = \mathbb{R}^n$.

Theorem 4.2. *Let $f, g, h : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x + c_f$, $g(x) = \frac{1}{2}x^T A_g x + c_g$, $h(x) = \frac{1}{2}x^T A_h x + c_h$, where $A_f, A_g, A_h \in S^n$ and $c_f, c_g, c_h \in \mathbb{R}$. Let $K \subseteq \mathbb{R}^n$ be a regular cone with $\dim(K \cup -K) \geq 3$. Suppose that there exist $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ such that $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$. Then, exactly one of the following two statements holds.*
(i) $(\exists x \in K)\ \ f(x) < 0,\ g(x) < 0,\ h(x) < 0.$
(ii) $(\exists(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3_+\backslash\{(0,0,0)\})\ \ (\forall x \in K)$

$$\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) \geq 0. \tag{4.1}$$

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that the following system has no solution: $f(x) < 0$, $g(x) < 0$, $h(x) < 0$, $x \in K$. Define $\Omega := \{(f(x), g(x), h(x)) : x \in K\} = \{(\frac{1}{2}x^T A_f x, \frac{1}{2}x^T A_g x, \frac{1}{2}x^T A_h x) : x \in K\} + \{(c_f, c_g, c_h)\}$. Then, by Theorem 4.1, Ω is a convex set and $\Omega \cap (-\mathrm{int}\mathbb{R}^3_+) = \emptyset$. Now, by the convex separation theorem, there exists $(\lambda_1, \lambda_2, \lambda_3) \neq (0,0,0)$ such that for all $(y_1, y_2, y_3) \in -\mathrm{int}\mathbb{R}^3_+$ and for all $x \in K$

$$\lambda_1 y_1 + \lambda_2 y_2 + \lambda_3 y_3 \leq 0,\ \lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) \geq 0.$$

Thus, $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3_+\backslash\{(0,0,0)\}$ and (4.1) holds. $\square$

In passing, note that it was shown in [9], without condition (2.1), that if $c_f = c_g = c_h = 0$, $K = \mathbb{R}^n$ and if the system $f(x) < 0$, $g(x) < 0$, $h(x) < 0$, $x \in K$, has no solution, then there exist $\lambda_1, \lambda_2, \lambda_3 \geq 0$ such that $\lambda_1 A_f + \lambda_2 A_g + \lambda_3 A_h$ has at most one negative eigenvalue. However, by imposing condition (2.1), we have obtained, in

Theorem 4.2, a stronger conclusion that $\lambda_1 A_f + \lambda_2 A_g + \lambda_3 A_h$ is positive semidefinite whenever the system has no solution with $K = \mathbb{R}^n$.

We now present two examples where the first one shows the condition "$\dim(K \cup -K) \geq 3$" cannot be dropped and the second one shows that the condition (2.1) cannot be dropped.

Example 4.1. *Consider $K = \mathbb{R}^2$. Let $f, g, h \in \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x_1, x_2) = x_1^2 - x_2^2$, $g(x_1, x_2) = -x_1^2 - x_1 x_2$ and $h(x_1, x_2) = -x_1^2 + x_1 x_2$. Then,*

$$A_f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, A_g = \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix} \text{ and } A_h = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now,

$$-A_f - 2A_g - 2A_h = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} \succ 0.$$

Clearly, the system $f(x) < 0$, $g(x) < 0$ and $h(x) < 0$ has no solution. We show that (ii) *of Theorem 4.2 fails. To see this, we proceed by contradiction and suppose that there exists $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3_+ \backslash \{(0,0,0)\}$ such that for all $x = (x_1, x_2) \in \mathbb{R}^2$*

$$\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) = (\lambda_1 - \lambda_2 - \lambda_3) x_1^2 + (\lambda_3 - \lambda_2) x_1 x_2 - \lambda_1 x_2^2 \geq 0.$$

If $\lambda_1 = 0$, then $\lambda_2 = \lambda_3 = 0$ which is impossible. On the other hand, if $\lambda_1 \neq 0$, then by fixing x_1 and letting $x_2 \to +\infty$, we see that $\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) \to -\infty$. This is a contradiction.

Example 4.2. *Let $n = 3$, $K = \mathbb{R}^3$ and let $f(x_1, x_2, x_3) = 2x_1 x_2 + x_1^2 - 2x_2^2 + 2x_3^2$, $g(x_1, x_2, x_3) = -2x_1^2 + 2x_2^2$ and $h(x_1, x_2, x_3) = -2x_1 x_2$. Then,*

$$A_f = \begin{pmatrix} 2 & 2 & 0 \\ 2 & -4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, A_g = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } A_h = \begin{pmatrix} 0 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

and, for each $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$,

$$\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h = \begin{pmatrix} 2\gamma_1 - 4\gamma_2 & 2\gamma_1 - 2\gamma_3 & 0 \\ 2\gamma_1 - 2\gamma_3 & -4\gamma_1 + 4\gamma_2 & 0 \\ 0 & 0 & 4\gamma_1 \end{pmatrix}.$$

Thus, positive definiteness of the matrix $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h$ implies that $\gamma_1 > 0$, $\gamma_1 > 2\gamma_2$ and $\gamma_1 < \gamma_2$ which is impossible. Hence, there do not exist $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ such that $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$. Note that the system $f(x) < 0$, $g(x) < 0$, $h(x) < 0$ has no solution. Otherwise, there exist $x_1, x_2, x_3 \in \mathbb{R}$ such that $-2x_1 x_2 < 0$, $-2x_1^2 + 2x_2^2 < 0$ and $2x_1 x_2 + x_1^2 - 2x_2^2 + 2x_3^2 < 0$. This implies that either $0 < x_2 < x_1$ or $0 > x_2 > x_1$ holds. It follows that $2x_2(x_1 - x_2) > 0$. Thus, $2x_1 x_2 + x_1^2 - 2x_2^2 + 2x_3^2 = 2x_2(x_1 - x_2) + x_1^2 + 2x_3^2 > 0$ which is impossible. Finally, we show that (ii) *of Theorem 4.2 fails. To see this,* *we proceed by contradiction and suppose that there exists* $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3_+ \backslash \{(0,0,0)\}$ *such that for each* $x = (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) = (\lambda_1 - 2\lambda_2)x_1^2 + 2(-\lambda_1 + \lambda_2)x_2^2 + 2(\lambda_1 - \lambda_3)x_1x_2 + 2\lambda_1 x_3^2 \geq 0.$$

Thus, we must have $\lambda_1 = \lambda_2 = \lambda_3 = 0$ *which is impossible.*

As a corollary of Theorem 4.2, we obtain a Motzkin type alternative theorem for three homogeneous quadratic strict inequalities and a single linear inequality.

Corollary 4.1. *Let* $a \in \mathbb{R}^n$ *and* $n \geq 3$. *Let* $f, g, h : \mathbb{R}^n \to \mathbb{R}$ *be defined by* $f(x) = \frac{1}{2}x^T A_f x + c_f$, $g(x) = \frac{1}{2}x^T A_g x + c_g$, $h(x) = \frac{1}{2}x^T A_h x + c_h$. *Suppose that there exist* $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ *such that* $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$. *Then, exactly one of the following two statements holds.*
(i) $(\exists x \in \mathbb{R}^n)$ $\;f(x) < 0,\; g(x) < 0,\; h(x) < 0,\; a^T x \leq 0.$
(ii) $(\exists (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}_3^+ \backslash \{(0,0,0)\}, \mu \geq 0)$ $(\forall x \in \mathbb{R}^n)$

$$\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) + \mu a^T x \geq 0.$$

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) does not hold. If $a = 0$, then the conclusion follows by applying Theorem 4.2 with $K = \mathbb{R}^n$. Without loss of generality, we may assume that $a \neq 0$. Thus, there exists $\overline{x} \in \mathbb{R}^n$ such that $a^T \overline{x} = -1$. Let $K = \{x \in \mathbb{R}^n : a^T x \leq 0\}$. Then, K is a regular convex cone of $\mathbb{R}^n$ with $\dim(K \cup (-K)) = \dim(\mathbb{R}^n) \geq 3$. Since (i) fails, it follows from Theorem 4.2 that there exists $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^2_+ \backslash \{(0,0,0)\}$ such that for all $x \in K = \{x \in \mathbb{R}^n : a^T x \leq 0\}$

$$\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) \geq 0.$$

Thus, $[a^T x \leq 0 \Rightarrow \lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) \geq 0]$. Note that $a^T \overline{x} = -1 < 0$. It follows from the S-lemma that there exists $\mu \geq 0$ such that for each $x \in \mathbb{R}^n$ $\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) + \mu a^T x \geq 0$. $\square$

Optimality Conditions for Homogeneous Programming Problems

Consider the following homogeneous optimization problem

$$(HOP) \quad \min \frac{1}{2}x^T A_f x \;\text{ s.t. }\; \frac{1}{2}x^T A_g x \leq 1,\; \frac{1}{2}x^T A_h x \leq 1,$$

where $f, g, h : \mathbb{R}^n \to \mathbb{R}$ $(n \geq 3)$ are defined by $f(x) = \frac{1}{2}x^T A_f x$, $g(x) = \frac{1}{2}x^T A_g x - 1$ and $h(x) = \frac{1}{2}x^T A_h x - 1$. Model problems of the form (HOP) arise in telecommunications and robust control (cf. [21, 28]). We now derive necessary and sufficient global optimality condition for problems (HOP) satisfying $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$.

Corollary 4.2. *For (HOP), suppose that there exist* $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ *such that* $\gamma_1 A_f + \gamma_2 A_g + \gamma_3 A_h \succ 0$. *Then, a feasible point* $\overline{x}$ *is a global minimizer of (HOP) if and only if there exist* $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ *such that* $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ *and* $A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0$.

Proof. Suppose that $\overline{x}$ is a global minimizer of (HOP). Define $\tilde{f}(x) = f(x) - f(\overline{x})$. Then the following system $\tilde{f}(x) < 0$, $g(x) < 0$ and $h(x) < 0$ has no solution. Now, it follows from Theorem 4.2 (replacing (f, g, h) by $(\tilde{f}, g, h)$) that there exists $(\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3_+ \backslash \{(0, 0, 0)\}$ such that for all $x \in \mathbb{R}^n$ $\mu_1 \tilde{f}(x) + \mu_2 g(x) + \mu_3 h(x) = \mu_1(f(x) - f(\overline{x})) + \mu_2 g(x) + \mu_3 h(x) \geq 0$. In particular,

$$\mu_2 g(\overline{x}) = \mu_3 h(\overline{x}) = 0. \tag{4.2}$$

So, $\mu_1 f + \mu_2 g + \mu_3 h$ attains its minimum at $\overline{x}$ over $\mathbb{R}^n$. We now show that $\mu_1 > 0$. Otherwise, $\mu_2 g(x) + \mu_3 h(x) \geq 0$ for all $x \in \mathbb{R}^n$. As $g(0) < 0$ and $h(0) < 0$, it follows that $\mu_2 = \mu_3 = 0$. This contradicts the fact that $(\mu_1, \mu_2, \mu_3) \neq (0, 0, 0)$. Hence

$$f(x) + \lambda_1 g(x) + \lambda_2 h(x) \geq f(\overline{x}),$$

where $\lambda_1 = \mu_2/\mu_1$ and $\lambda_2 = \mu_3/\mu_1$. This implies that $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$. Therefore, $\overline{x}$ is a global minimizer of $f + \lambda_1 g + \lambda_2 h$ over $\mathbb{R}^n$. This gives us that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$ and $\nabla^2(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0$.

Conversely, suppose that there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0, 0)\}$ such that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ and $A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0$. Let the function $L : \mathbb{R}^n \to \mathbb{R}$ be defined by $L(x) = f(x) + \lambda_1 g(x) + \lambda_2 h(x)$. Then $L(\cdot)$ is a convex function on $\mathbb{R}^n$ as $\nabla^2 L(x) = A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0$ for all $x \in \mathbb{R}^n$. Now, it follows from $\nabla L(\overline{x}) = \nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$ that $\overline{x}$ is a global minimizer of L. Therefore, one has for each $x \in \mathbb{R}^n$,

$$f(x) + \lambda_1 g(x) + \lambda_2 h(x) = L(x) \geq L(\overline{x}) = f(\overline{x}) + \lambda_1 g(\bar{x}) + \lambda_2 h(\bar{x}) = f(\overline{x}).$$

Therefore, $f(\overline{x}) \leq f(x)$ for all x such that $g(x) \leq 0$ and $h(x) \leq 0$, i.e., $\overline{x}$ is a global minimizer of (HOP). □

It is worthy noting that Corollary 4.2 was established in [26] under the following slightly stronger condition: there exist $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $\gamma_1 A_g + \gamma_2 A_h \succ 0$.

4.2 Non-Homogeneous Systems

Let $f, g, h : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$, $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$ and $h(x) = \frac{1}{2}x^T A_h x + b_h^T x + c_h$, where $A_f, A_g, A_h \in S^n$, $b_f, b_g, b_h \in \mathbb{R}^n$, $c_f, c_g, c_h \in \mathbb{R}$. We define H_f, H_g, H_h by

$$H_f = \begin{pmatrix} A_f & b_f \\ b_f^T & 2c_f \end{pmatrix}, H_g = \begin{pmatrix} A_g & b_g \\ b_g^T & 2c_g \end{pmatrix} \text{ and } H_h = \begin{pmatrix} A_h & b_h \\ b_h^T & 2c_h \end{pmatrix}. \tag{4.3}$$

We derive a nonconvex Gordan-type alternative theorem for three nonhomogeneous quadratic functions by way of homogenization.

Theorem 4.3. *Let $n \geq 3$ and let f, g, h and H_f, H_g, H_h be defined as above. Suppose that there exist $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ such that*

$$\gamma_1 H_f + \gamma_2 H_g + \gamma_3 H_h \succ 0. \tag{4.4}$$

Then exactly one of the following statements holds:
(i) $(\exists x \in \mathbb{R}^n)$ $f(x) < 0,\ g(x) < 0,\ h(x) < 0.$
(ii) $(\exists (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3_+ \backslash \{(0,0,0)\})$ $(\forall x \in \mathbb{R}^n)$

$$\lambda_1 f(x) + \lambda_2 g(x) + \lambda_3 h(x) \geq 0. \tag{4.5}$$

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) does not hold. Then the system $f(x) < 0,\ g(x) < 0,\ h(x) < 0$ has no solution. Define three homogeneous functions $\tilde{f}, \tilde{g}, \tilde{h} : \mathbb{R}^{n+1} \to \mathbb{R}$ by

$$\tilde{f}(x,t) = \frac{1}{2} x^T A_f x + b_f^T x t + c_f t^2,\ \tilde{g}(x,t) = \frac{1}{2} x^T A_g x + b_g^T x t + c_g t^2$$

and

$$\tilde{h}(x,t) = \frac{1}{2} x^T A_h x + b_h^T x t + c_h t^2.$$

Then, $\tilde{f}(x,t) = \frac{1}{2}(x,t)^T H_f (x,t)$, $\tilde{g}(x,t) = \frac{1}{2}(x,t)^T H_g (x,t)$ and $\tilde{h}(x,t) = \frac{1}{2}(x,t)^T H_h (x,t)$. Moreover, the system $\tilde{f}(x,t) < 0$ $\tilde{g}(x,t) < 0$ and $\tilde{h}(x,t) < 0$ has no solution. Otherwise, there exists (x_0, t_0) such that $\tilde{f}(x_0,t_0) < 0$, $\tilde{g}(x_0,t_0) < 0$ and $\tilde{h}(x_0,t_0) < 0$. If $t_0 \neq 0$, then $f(x_0/t_0) = t_0^{-2} \tilde{f}(x_0,t_0) < 0$, $g(x_0/t_0) = t_0^{-2} \tilde{g}(x_0,t_0) < 0$ and $h(x_0/t_0) = t_0^{-2} \tilde{h}(x_0,t_0) < 0$. This contradicts to the fact that the system ($f(x) < 0$, $g(x) < 0$, $h(x) < 0$) has no solution. If $t_0 = 0$, then $\frac{1}{2} x_0^T A_f x_0 = \tilde{f}(x_0,t_0) < 0$, $\frac{1}{2} x_0^T A_g x_0 = \tilde{g}(x_0,t_0) < 0$ and $\frac{1}{2} x_0^T A_h x_0 = \tilde{h}(x_0,t_0) < 0$. This implies that

$$\lim_{\alpha \to +\infty} f(\alpha x_0) = -\infty,\ \lim_{\alpha \to +\infty} g(\alpha x_0) = -\infty \text{ and } \lim_{\alpha \to +\infty} h(\alpha x_0) = -\infty.$$

Thus, there exists $\alpha_0 > 0$ such that $f(\alpha_0 x_0) < 0, g(\alpha_0 x_0) < 0$ and $h(\alpha_0 x_0) < 0$. This is a contradiction. So, it follows from Theorem 4.2 (replacing (f, g, K) by $(\tilde{f}, \tilde{g}, \mathbb{R}^{n+1})$) that there exists $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3_+ \backslash \{(0,0,0)\}$ such that for all $(x,t) \in \mathbb{R}^{n+1}$

$$\lambda_1 \tilde{f}(x,t) + \lambda_2 \tilde{g}(x,t) + \lambda_3 \tilde{h}(x,t) \geq 0. \tag{4.6}$$

Thus, by setting $t = 1$ in (4.6), we see that (ii) holds. $\square$

4.3 Optimality Conditions for Trust Region type Problems

Consider the following trust region type problems

$$(TR) \quad \min f(x) \ \text{ s.t. } g(x) \leq 0,\ h(x) \leq 0$$

where $f, g, h : \mathbb{R}^n \to \mathbb{R}$ $(n \geq 3)$ are defined by $f(x) = \frac{1}{2}x^T A_f x + b_f^T x + c_f$, $g(x) = \frac{1}{2}x^T A_g x + b_g^T x + c_g$ and $h(x) = \frac{1}{2}x^T A_h x + b_h^T x + c_h$. The problem (TR) is said to be regular whenever there exist $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\gamma_1 H_g + \gamma_2 H_h \succ 0, \tag{4.7}$$

where H_g, H_h are defined as in (4.3). Let us now derive necessary, and sufficient global optimality conditions for regular (TR) problems. Such conditions for possibly non-regular problems are given in Section 5.2.

Theorem 4.4. *Suppose that problem (TR) is regular and that $\overline{x}$ is a global minimizer. Then, the following Fritz-John type necessary condition holds, that is, there exists $(\mu, \lambda_1, \lambda_2) \in \mathbb{R}^3_+ \backslash \{(0,0,0)\}$ such that $\nabla(\mu f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ and*

$$\mu A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0.$$

Moreover, if the Slater condition holds, i.e., there exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$ and $h(x_0) < 0$, then a feasible point $\overline{x}$ is a global minimizer if and only if there exist $\lambda_1, \lambda_2 \geq 0$ such that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ and

$$A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0.$$

Proof. Suppose that $\overline{x}$ is a global minimizer of (TR). Then the following system $f(x) - f(\overline{x}) < 0$, $g(x) < 0$ and $h(x) < 0$ has no solution. Since (TR) is regular, (4.4) holds with $\gamma_1 = 0$ and so, from Theorem 4.3, there exists $(\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3_+ \backslash \{(0,0,0)\}$ such that for all $x \in \mathbb{R}^n$ $\mu_1(f(x) - f(\overline{x})) + \mu_2 g(x) + \mu_3 h(x) \geq 0$. In particular,

$$\mu_2 g(\overline{x}) = \mu_3 h(\overline{x}) = 0. \tag{4.8}$$

So, $\mu_1 f + \mu_2 g + \mu_3 h$ attains its minimum at $\overline{x}$ over $\mathbb{R}^n$. Thus the Fritz-John type necessary condition holds with $\mu = \mu_1$ and $\lambda_i = \mu_i$, $i = 1, 2$.

Suppose further that the Slater condition holds. Then $\mu_1 > 0$. Otherwise, $\mu_2 g(x) + \mu_3 h(x) \geq 0$ for all $x \in \mathbb{R}^n$. Note that $g(x_0) < 0$ and $h(x_0) < 0$. It follows that $\mu_2 = \mu_3 = 0$. This contradicts the fact that $(\mu_1, \mu_2, \mu_3) \neq (0,0,0)$. Hence

$$f(x) + \lambda_1 g(x) + \lambda_2 h(x) \geq f(\overline{x}),$$

where $\lambda_1 = \mu_2/\mu_1$ and $\lambda_2 = \mu_3/\mu_1$. This implies that $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$. Therefore, $\overline{x}$ is a global minimizer of $f + \lambda_1 g + \lambda_2 h$ over $\mathbb{R}^n$. This gives us that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$ and

$$\nabla^2(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0.$$

Conversely, suppose that there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ such that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ and $A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0$. Consider the function $L : \mathbb{R}^n \to \mathbb{R}$ defined by $L(x) = f(x) + \lambda_1 g(x) + \lambda_2 h(x)$. Since $\nabla^2 L(x) = A_f + \lambda_1 A_g + \lambda_2 A_h \succeq 0$ for all $x \in \mathbb{R}^n$, $L(\cdot)$ is a convex function on $\mathbb{R}^n$. It follows from $\nabla L(\overline{x}) =$

$\nabla(f+\lambda_1 g+\lambda_2 h)(\overline{x}) = 0$ that $\overline{x}$ is a global minimizer of L. Therefore, one has for each $x \in \mathbb{R}^n$,

$$f(x) + \lambda_1 g(x) + \lambda_2 h(x) = L(x) \geq L(\overline{x}) = f(\overline{x}) + \lambda_1 g(\bar{x}) + \lambda_2 h(\bar{x}) = f(\overline{x}).$$

Therefore, $f(\overline{x}) \leq f(x)$ for all x such that $g(x) \leq 0$ and $h(x) \leq 0$, i.e., $\overline{x}$ is a global minimizer of (TR). □

We note that the following positive independent condition on $\nabla g(\overline{x})$ and $\nabla h(\overline{x})$ at a feasible point $\overline{x}$,

$$[\alpha \nabla g(\overline{x}) + \beta \nabla h(\overline{x}) = 0, \alpha g(\overline{x}) = \beta h(\overline{x}) = 0, \alpha \geq 0, \beta \geq 0] \;\Rightarrow\; [\alpha = \beta = 0], \qquad (4.9)$$

implies that the Slater condition holds. To see this, let $\overline{x}$ be a feasible point satisfying (4.9). Suppose that the Slater condition fails. Then, by Theorem 3.3, there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ such that for all $x \in \mathbb{R}^n$ $\lambda_1 g(x) + \lambda_2 h(x) \geq 0$. Note that $g(\overline{x}) \leq 0$ and $h(\overline{x}) \leq 0$. It follows that $\overline{x}$ is a global minimizer of $\lambda_1 f + \lambda_2 g$ over $\mathbb{R}^n$. Thus, one has $\lambda_1 \nabla g(\overline{x}) + \lambda_2 \nabla h(\overline{x}) = 0, \lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$. This together with (4.9) yields that $\lambda_1 = \lambda_2 = 0$ which is impossible. We also note that the condition (4.9) includes several interesting cases which were studied in [24]. (For example, (1) $\nabla g(\overline{x})$ and $\nabla h(\overline{x})$ are linear independent; (2) $\nabla g(\overline{x}) = t \nabla h(\overline{x})$ with $t > 0$ and $\nabla h(\overline{x}) \neq 0$; (3) $g(\overline{x}) = 0$, $\nabla g(\overline{x}) \neq 0$ and $h(\overline{x}) < 0$). In [24], without assuming (4.7), it was shown that at a global minimizer, there exist $\lambda_1, \lambda_2 \geq 0$ such that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ and $A_f + \lambda_1 A_g + \lambda_2 A_h$ has at most one negative eigenvalue in the above 3 cases. Our optimality condition is stronger than the ones in [24] for regular trust region type problems (TR).

Consider the following trust-region problem, considered in [30, 32]:

$$(TR1) \quad \min \quad x^T A_f x + 2b_f^T x + c_f \;\; \text{s.t. } \alpha \leq x^T A_g x \leq \beta,$$

where $A_f, A_g \in S^n, b_f \in \mathbb{R}^n, c_f, \alpha, \beta \in \mathbb{R}$ ($n \geq 3$), $A_g \succ 0$ and $0 < \alpha < \beta$. Let $f(x) = x^T A_f x + 2b_f^T x + c_f$, $g(x) = -x^T A_g x + \alpha$ and $h(x) = x^T A_g x - \beta$. Let $\overline{x}$ be a feasible point of (TR1). Then $\overline{x}$ is a global minimizer of (TR1) if and only if there exist $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ such that $\nabla(f + \lambda_1 g + \lambda_2 h)(\overline{x}) = 0$, $\lambda_1 g(\overline{x}) = \lambda_2 h(\overline{x}) = 0$ and

$$A_f - \lambda_1 A_g + \lambda_2 A_g \succeq 0.$$

To see this, let $\gamma_1, \gamma_2 < 0$ be such that $\frac{\alpha}{\beta} < \frac{\gamma_2}{\gamma_1} < 1$. Since $A_g \succ 0$, it follows that

$$\gamma_1 H_g + \gamma_2 H_h = \begin{pmatrix} 2(-\gamma_1 + \gamma_2) A_g & 0 \\ 0 & 2(\gamma_1 \alpha - \gamma_2 \beta) \end{pmatrix} \succ 0.$$

This gives us that the problem (TR1) is regular and so, the optimality condition follows from Theorem 4.4.

5 Systems of Finitely Many Quadratic Inequalities

In this section, we present an alternative theorem for a system of finitely many quadratic inequalities, and obtain necessary and sufficient global optimality conditions for a class of quadratic programming problem involving Z-matrices.

Recall that a matrix $A = (A_{ij})_{1\le i,j\le n} \in S^n$ is called a Z-matrix if $A_{ij} \le 0$ for all $i \ne j$ (S^n is the set consisting of all real $n \times n$ symmetric matrix). From the definition, any diagonal matrix is a Z-matrix. The Z-matrix arises naturally in solving Dirichlet problem numerically, and play an important role in the theory of linear complementary problem (cf. [5, 13, 14]). Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be defined by $f_i(x) = \frac{1}{2}x^T A_i x + b_i^T x + c_i$, $i = 1, \dots, m$, where $A_i \in S^n, b_i \in \mathbb{R}^n, c_i \in \mathbb{R}$. We define H_i by

$$H_i = \begin{pmatrix} A_i & b_i \\ b_i^T & 2c_i \end{pmatrix}. \tag{5.1}$$

Define a set Ω_0 by

$$\Omega_0 := \{(\frac{1}{2}a^T H_1 a, \dots, , \frac{1}{2}a^T H_m a) : a \in \mathbb{R}^{n+1}\} + \mathrm{int}\mathbb{R}^m_+.$$

In the following, we give a sufficient condition for the convexity of Ω_0 in terms of Z-matrices.

Theorem 5.1. *Let H_i be Z-matrices $i = 1, \dots, m$. Then Ω_0 is a convex set.*

Proof. Since $x^T H_i x = H_i \cdot (xx^T)$ $i = 1, \dots, m$, it can be verified that

$$\begin{aligned} \Omega_0 : \ &= \ \{(\frac{1}{2}x^T H_1 x, \dots, \frac{1}{2}x^T H_m x) : x \in \mathbb{R}^{n+1}\} + \mathrm{int}\mathbb{R}^m_+ \\ &= \ \{(\frac{1}{2}H_1 \cdot X, \dots, \frac{1}{2}H_m \cdot X) : X = xx^T,\, x \in \mathbb{R}^{n+1}\} + \mathrm{int}\mathbb{R}^m_+ \\ &\subseteq \ \{(\frac{1}{2}H_1 \cdot X, \dots, \frac{1}{2}H_m \cdot X) : X \in S^{n+1}_+\} + \mathrm{int}\mathbb{R}^m_+. \end{aligned}$$

Note that $\{(\frac{1}{2}H_1\cdot X, \dots, \frac{1}{2}H_m\cdot X) : X \in S^{n+1}_+\}$ is convex (and hence $\{(\frac{1}{2}H_1\cdot X, \dots, \frac{1}{2}H_m\cdot X) : X \in S^{n+1}_+\} + \mathrm{int}\mathbb{R}^m_+$ is also convex). To conclude the proof, it suffices to show that, if each H_i is a Z-matrix, then

$$\begin{aligned} &\{(\frac{1}{2}H_1 \cdot X, \dots, \frac{1}{2}H_m \cdot X) : X \in S^{n+1}_+\} + \mathrm{int}\mathbb{R}^m_+ \\ \subseteq \ &\{(\frac{1}{2}H_1 \cdot X, \dots, \frac{1}{2}H_m \cdot X) : X = xx^T,\, x \in \mathbb{R}^{n+1}\} + \mathrm{int}\mathbb{R}^m_+. \end{aligned}$$

To see this, take $(z_1, \dots, z_m) \in \{(\frac{1}{2}H_1\cdot X, \dots, \frac{1}{2}H_m\cdot X) : X \in S^{n+1}_+\} + \mathrm{int}\mathbb{R}^m_+$. Then, there exists a $X_0 \in S^{n+1}_+$ such that $H_k \cdot X_0 < 2z_k$, $k = 1, \dots, m$. We now show that there exists a vector u_0 such that for each $k = 1, \dots, m$

$$H_k \cdot X_0 \ge u_0^T H_k u_0 = H_k \cdot (u_0 u_0^T).$$

To see this, we use x_{ij} to denote the element of X_0 which lies at the i^{th} row and j^{th} column. Since $X_0 \in S_+^{n+1}$, one has $x_{ii} \geq 0$ $(i = 1, \ldots, n+1)$ and

$$x_{jj}x_{ii} - x_{ji}^2 \geq 0 \quad i, j \in \{1, \ldots, n+1\}. \tag{5.2}$$

Now, define $u_0 = (\sqrt{x_{11}}, \ldots, \sqrt{x_{n+1n+1}})$. Then, the (j,i) element of $u_0u_0^T$ is $\sqrt{x_{jj}x_{ii}}$, and hence for all $k = 1, \ldots, m$, one has

$$\begin{aligned} u_0^T H_k u_0 - H_k \cdot X_0 = H_k \cdot (u_0 u_0^T) - H_k \cdot X_0 &= H_k \cdot (u_0 u_0^T - X_0) \\ &= \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} a_{ij}^k (\sqrt{x_{jj}x_{ii}} - x_{ji}) \\ &= \sum_{i,j=1, i \neq j}^{n+1} a_{ij}^k (\sqrt{x_{jj}x_{ii}} - x_{ji}) \\ &\leq 0 \end{aligned}$$

where a_{ij}^k is the (i,j) element of H_k and the last inequality follows from $a_{ij}^k \leq 0$ for all $i \neq j$ (since H_k is a Z-matrix) and (5.2). Hence, $H_k \cdot (u_0 u_0^T) < 2z_k$. Therefore, $(z_1, \ldots, z_m) \in \{(\frac{1}{2}H_1 \cdot X, \ldots, \frac{1}{2}H_m \cdot X) : X = xx^T,\ x \in \mathbb{R}^n\} + \mathrm{int}\mathbb{R}_+^m$. □

In the following we establish a Gordan type alternative theorem for systems of arbitrary finite number of quadratic inequalities involving Z-matrices.

Theorem 5.2. *Let f_i be defined by $f_i(x) = \frac{1}{2}x^T A_i x + b_i^T x + c_i$, $i = 1, \ldots, m$. Suppose that H_i, $i = 1, \ldots, m$ are all Z-matrices. Then, exactly one of the following two statements holds.*
(i) $(\exists x \in \mathbb{R}^n)$ $f_i(x) < 0$, $i = 1, \ldots, m$.
(ii) $(\exists \lambda \in \mathbb{R}_+^m \backslash \{0\})$ $(\forall x \in \mathbb{R}^n)$ $\sum_{i=1}^m \lambda_i f_i(x) \geq 0$.

Proof. It suffices to show [Not(i) $\Rightarrow$ (ii)]. Suppose that (i) fails. That is, the following system has no solution: $x \in \mathbb{R}^n$, $f_i(x) < 0$, $i = 1, \ldots, m$. Define m homogeneous functions $\tilde{f}_i : \mathbb{R}^{n+1} \to \mathbb{R}$ by

$$\tilde{f}_i(x,t) = \frac{1}{2}(x,t)H_i(x,t)^T = \frac{1}{2}x^T A_i x + b_i^T x t + c_i t^2.$$

Then $0 \notin \Omega_0$. Otherwise, there exists (x_0, t_0) such that $\tilde{f}_i(x_0, t_0) < 0$. If $t_0 \neq 0$, then $f_i(x_0/t_0) = t_0^{-2}\tilde{f}_i(x_0, t_0) < 0$, $i = 1, \ldots, m$. This contradicts our assumption (i). If $t_0 = 0$, then $\frac{1}{2}x_0^T A_i x_0 = \tilde{f}_i(x_0, t_0) < 0$, $i = 1, \ldots, m$. This implies that

$$\lim_{\alpha \to +\infty} f_i(\alpha x_0) = -\infty.$$

Thus, there exists $\alpha_0 > 0$ such that $f_i(\alpha_0 x_0) < 0, i = 1, \ldots, m$ which means assumption (i) holds. This is a contradiction.

The set Ω_0, is convex, and so, by the convex separation theorem, there exists $\lambda \in \mathbb{R}^m \backslash \{0\}$ such that for all $(y_1, \ldots, , y_m) \in \Omega_0$, $\sum_{i=1}^m \lambda_i y_i \geq 0$. This implies that $\lambda \in \mathbb{R}_+^m \backslash \{0\}$ and for all $(x,t) \in \mathbb{R}^{n+1}$, $\sum_{i=1}^m \lambda_i (\frac{1}{2}x^T A_i x + b_i^T x t + c_i t^2) \geq 0$. Thus, by setting $t = 1$, we see that (ii) holds. □

In passing, we observe that Theorem 5.2 extends the corresponding result in [4, Exercise 4.57], where A_i is a diagonal matrix and $b_i = 0$.

5.1 Optimality Conditions for General Quadratic Programs

Consider the following general quadratic optimization problem

$$(QOP) \quad \min f_0(x) \text{ s.t. } f_i(x) \leq 0, i = 1, \dots, m,$$

where each $f_i : \mathbb{R}^n \to \mathbb{R}$ is defined by $f_i(x) = \frac{1}{2}x^T A_i x + b_i^T x + c_i$. Let H_i be defined as in (5.1), $i = 0, 1, \dots, m$.

Corollary 5.1. *For (QOP), suppose that each H_i is a Z-matrix, $i = 0, \dots, m$. Let $\overline{x}$ be a global minimizer. Then, the following Fritz-John type necessary condition holds, that is, there exists $(\lambda_0, \dots, \lambda_m) \in \mathbb{R}^{m+1}_+ \backslash \{0\}$ such that $\nabla(\sum_{i=0}^m \lambda_i f_i)(\overline{x}) = 0$, $\lambda_i f_i(\overline{x}) = 0$, $i = 1, \dots, m$ and $\sum_{i=0}^m \lambda_i A_i \succeq 0$. Moreover, if the Slater condition holds, i.e., there exists $x_0 \in \mathbb{R}^n$ such that $f_i(x_0) < 0$, $i = 1, \dots, m$, then, a feasible point $\overline{x}$ is a global minimizer if and only if there exists $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m_+ \backslash \{0\}$ such that $\nabla(f_0 + \sum_{i=1}^m \lambda_i f_i)(\overline{x}) = 0$, $\lambda_i f_i(\overline{x}) = 0$, $i = 1, \dots, m$ and $A_0 + \sum_{i=1}^m \lambda_i A_i \succeq 0$.*

Proof. Suppose that $\overline{x}$ is a global minimizer of (QOP). Let $\tilde{f}(x) := f_0(x) - f_0(\overline{x})$ and

$$H_{\tilde{f}} = \begin{pmatrix} A_0 & b_0 \\ b_0^T & -\overline{x}^T A_0 \overline{x} - 2b_0^T \overline{x} \end{pmatrix}.$$

Then the following system $\tilde{f}(x) < 0$, $f_i(x) < 0$, $i = 1, \dots, m$, has no solution. Note that $H_{\tilde{f}}$ is a Z-matrix if and only if H_0 is a Z-matrix. Thus, from Theorem 5.2, there exists $(\lambda_0, \dots, \lambda_m) \in \mathbb{R}^{m+1}_+ \backslash \{0\}$ such that for all $x \in \mathbb{R}^n$ $\lambda_0 \tilde{f}(x) + \sum_{i=1}^m \lambda_i f_i(x) = \lambda_0(f(x) - f(\overline{x})) + \sum_{i=1}^m \lambda_i f_i(x) \geq 0$. In particular, one has $\lambda_i f_i(\overline{x}) = 0$, $i = 1, \dots, m$. Thus, $\sum_{i=0}^m \lambda_i f_i$ attains its minimum at $\overline{x}$ over $\mathbb{R}^n$. Thus the Fritz-John type necessary condition holds. If the Slater condition holds, then, using similar arguments as the proof of Theorem 4.4, we see that the sufficient and necessary condition holds. $\square$

5.2 Applications to CDT Problems

Finally, we see that Corollary 5.1 can be used to examine possibly non-regular trust-region type problems. For instance, consider the following CDT problem which plays an important role in the trust region algorithms for nonlinear programming in [8, 27].

$$(CDT) \quad \min \frac{1}{2}x^T Bx + b^T x \text{ s.t. } \frac{1}{2}\|A^T x + a\|^2 \leq \xi^2/2, \ \frac{1}{2}\|x\|^2 \leq \Delta^2/2,$$

where $A \in \mathbb{R}^{n \times m}$, $B \in S^n$, $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$ and $\Delta, \xi \in [0, +\infty)$. Let $f(x) = \frac{1}{2}x^T Bx + b^T x$, $g(x) = \frac{1}{2}(\|A^T x + a\|^2 - \xi^2)$, $h(x) = \frac{1}{2}(\|x\|^2 - \Delta^2)$. Define

$$H_f = \begin{pmatrix} B & b \\ b^T & 0 \end{pmatrix}, H_g = \begin{pmatrix} AA^T & Aa \\ (Aa)^T & \|a\|^2 - \xi^2 \end{pmatrix} \text{ and } H_h = \begin{pmatrix} I_n & 0 \\ 0 & -\Delta^2 \end{pmatrix}. \quad (5.3)$$

Corollary 5.2. *For (CDT), suppose that B, AA^T are Z-matrices, $Aa, b \in -\mathbb{R}^n_+$ and the Slater condition holds. Then, a feasible point $\overline{x}$ is a global minimizer if and only if there exists $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+\backslash\{0\}$ such that $(B + \lambda_1 AA^T + \lambda_2 I_n)\overline{x} + (b + \lambda_1 Aa) = 0$, $\lambda_1(\|A\overline{x} + a\|^2 - \xi^2) = 0$, $\lambda_2(\|\overline{x}\|^2 - \Delta^2) = 0$, and $B + \lambda_1 AA^T + \lambda_2 I_n \succeq 0$.*

Proof. Define $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, 3$, by $f_1(x) = f(x)$, $f_2(x) = g(x)$ and $f_3(x) = h(x)$. Since B, AA^T are Z-matrices and $Aa, b \in -\mathbb{R}^n_+$, H_f, H_g are Z-matrices. Note that H_h is a diagonal matrix and hence is also a Z-matrix. Thus, the conclusion follows by Corollary 5.1. $\square$

The following example illustrates an application of Corollary 5.2 to a two-dimensional CDT problem.

Example 5.1. *Consider the following CDT problem*

$$\min_{x\in\mathbb{R}^2} \frac{1}{2}x^T Bx + b^T x \quad s.t. \quad \frac{1}{2}\|Ax + a\|^2 \le \xi^2/2, \ \frac{1}{2}\|x\|^2 \le \Delta^2/2,$$

where $n = m = 2$, $a = (0, -6)^T$, $b = (0, -6)^T$, $\Delta = 5$, $\xi = 5$,

$$B = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix} \text{ and } A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This problem can be equivalently rewritten as

$$\min_{x=(x_1,x_2)\in\mathbb{R}^2} f(x) \quad s.t. \ g(x) \le 0, \ h(x) \le 0,$$

where $f(x) = -x_1^2 + x_2^2 - 6x_2$, $g(x) = \frac{1}{2}(x_1^2 + (x_2 - 6)^2 - 25)$ and $h(x) = \frac{1}{2}(x_1^2 + x_2^2 - 25)$ with global optimizer $\overline{x} = (\pm 4, 3)$. We can easily verify that

$$H_f = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & -6 \\ 0 & -6 & 0 \end{pmatrix}, H_g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -6 \\ 0 & -6 & 11 \end{pmatrix} \text{ and } H_h = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -25 \end{pmatrix}.$$

and, for any $\mu_1, \mu_2 \in \mathbb{R}$,

$$\mu_1 H_g + \mu_2 H_h = \begin{pmatrix} \mu_1 + \mu_2 & 0 & 0 \\ 0 & \mu_1 + \mu_2 & -6\mu_1 \\ 0 & -6\mu_1 & 11\mu_1 - 25\mu_2 \end{pmatrix}.$$

Note that $(\mu_1 + \mu_2)(11\mu_1 - 25\mu_2) - 36\mu_1^2 = -25\mu_1^2 - 25\mu_2^2 - 14\mu_1\mu_2 = -7(\mu_1 + \mu_2)^2 - 18(\mu_1^2 + \mu_2^2) \le 0$. Thus it is not a regular (TR) problem. However, it is easy to verify that B, AA^T are Z-matrices, $Aa, b \in -\mathbb{R}^n_+$ and the Slater condition holds. Letting $\lambda_1 = \lambda_2 = 1$, we see that

$$B + \lambda_1 AA^T + \lambda_2 I_n = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} \succeq 0,$$

$g(\overline{x}) = h(\overline{x}) = 0$ and

$$(B + \lambda_1 AA^T + \lambda_2 I_n)\overline{x} + (b + \lambda_1 Aa) = (0, 12)^T + (0, -6)^T + (0, -6)^T = (0, 0)^T.$$

Thus, the global optimality conditions of Corollary 5.2 holds.

References

[1] A. Baccari, A. Trad, On the classical necessary second-order optimality conditions in the presence of equality and inequality constraints. *SIAM J. Optim.* **15** (2005), 394-408.

[2] A. Beck and Y. C. Eldar, Strong duality in nonconvex quadratic optimization with two quadratic constraints, *SIAM J. Optim.*, **17** (2006), 844-860 .

[3] A. Beck, On the convexity of a class of quadratic mappings and its application to the problem of finding the smallest ball enclosing a given intersection of ball, *J. Global Optim.*, 39(2007), 113–126.

[4] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications*, SIAM-MPS, Philadelphia, 2000.

[5] R. B. Bapat and T. E. S. Raghavan, *Nonnegative Matrices and Applications*, Cambridge University Press, United Kingdom, 1997.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[7] L. Brickman, On the field of values of a matrix, *Proc. Amer. Math. Soc.* 12 (1961), 61-66.

[8] M. R. Celis, J. E. Dennis, and R. A. Tapia, *A trust region strategy for nonlinear equality constrained optimization, in Numerical Optimization*, P. T. Boggs, R. H. Byrd, and R. B. Schnabel, eds., SIAM, Philadelphia, 1985, pp. 71-82.

[9] X. Chen and Y. X. Yuan, A note on quadratic forms, *Math. Program.*, **86**, (1999) 187-197.

[10] X. D. Chen, and Y. X. Yuan, On local solutions of the Celis-Dennis-Tapia subproblem. *SIAM J. Optim.* **10** (2000), 359-383.

[11] J. P. Crouzeix, J. E. Martínez-Legaz and A. Seeger, An alternative theorem for quadratic forms and extensions. *Linear Algebra Appl.* **215** (1995), 121-134.

[12] L. L. Dines, On the mapping of quadratic forms, *Bull. Amer. Math. Soc.*, **47** (1941), 494-498.

[13] M. Fiedler, *Special Matrices and their Applications in Numerical Mathematics,* 1986, Martinus Nijhoff Publishers, Dordrecht.

[14] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, 1991, Cambridge University Press, United Kingdom.

[15] V. Jeyakumar, Farkas Lemma: Generalizations, *Encylopedia of optimization*, **2**, 87-91, Kluwer Boston, USA, 2000.

[16] V. Jeyakumar and D. T. Luc, *Nonsmooth vector functions and continuous optimization*, Springer Optimization and Its Applications, 10, Springer, New York, 2008.

[17] V. Jeyakumar, A. M. Rubinov and Z. Y. Wu, Non-convex quadratic minimization problems with quadratic constraints: Global optimality conditions, *Math. Program.*, Ser. A, **110** (2007), 521-541.

[18] G. M. Lee, N. N. Tam and N. D. Yen, *Quadratic Programming and Affine Variational Inequalities. A Qualitative Study*, Nonconvex Optimization and its Applications, 78. Springer-Verlag, New York, 2005.

[19] J. E. Martínez-Legaz and A. Seeger, Yuan's alternative theorem and the maximization of the minimum eigenvalue function. *J. Optim. Theory Appl.* **82** (1994), no. 1, 159-167.

[20] J. E. Martinez-Legaz, On Brickman's theorem, *J. Convex Anal.* 12 (2005), 139–143.

[21] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, Convex approximation techniques for joint multiuser downlink beamforming and admission control, *IEEE Trans. Wireless Communication* (to appear).

[22] J.J. Moré, and D.C. Sorensen, Computing a trust region step. *SIAM J. Sci. Statist. Comput.* **4** (1983), no. 3, 553–572.

[23] P. Pardalos and H. Romeijn, *Handbook in Global Optimization*, **2**, Kluwer Academic Publishers, (2002).

[24] J. M. Peng, and Y. X. Yuan, Optimality conditions for the minimization of a quadratic with two quadratic constraints. *SIAM J. Optim.* **7**, (1997), 579-594.

[25] I. Pólik and T. Terlaky, A survey of the S-Lemma, *SIAM Rev.*, **49** (2007), 371-418.

[26] B.T. Polyak, Convexity of quadratic transformation and its use in control and optimization, *J. Optim. Theory Appl.* **99**, (1998) 563-583.

[27] M. J. D. Powell and Y. Yuan, A trust region algorithm for equality constrained optimization, *Math. Program.*, **49** (1991), pp. 189-211.

[28] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, Transmit beamforming for physical layer multicasting, *IEEE Trans. Signal Process.*, **54** (2006), 2239–2251.

[29] D.C. Sorensen, Newton's method with a model trust region modification. *SIAM J. Numer. Anal.* **19** (1982), no. 2, 409–426.

[30] R. J. Stern and H. Wolkowicz, Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations, *SIAM J. Optim.*, **5**, (1995), 286-313.

[31] J. F. Sturm and S. Z. Zhang, On cones of nonnegative quadratic functions, *Math. Oper. Res.,* **28**, (2003) , 246-267.

[32] Y. Y. Ye and S. Z. Zhang, New results of quadratic minimization, *SIAM J. Optim.*, **14**, (2003), 245-267.

[33] Y. X. Yuan, On a subproblem of trust region algorithms for constrained optimization, *Math. Program.*, **47** (1990) 53-63.