The Hong Kong Polytechnic University
Department of Applied Mathematics

On The Equivalence of Global Quadratic Growth Condition and Second-order Sufficient Condition

by

Zhangyou Chen

A thesis submitted in partial fulfilment of
the requirements for the degree of Master of Philosophy

November 2009

CERTIFICATE OF ORIGINALITY

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institute of higher learning, except where due acknowledgment is made in the text.

ZHANGYOU CHEN

Abstract

The main purpose of this thesis is to study various properties of quadratically constrained quadratic programming problems. We concentrate on the existence of solutions and global second-order sufficient conditions for the quadratic problems.

First, we study the existence of global solutions for general quadratic programming problems. With the tool of asymptotical directions of sets, we provide an alternative proof for the nonemptiness of intersection of a sequence of nested sets defined by a finite number of convex quadratic functions, which can be applied to show the existence results of the corresponding optimization problems. We also prove the existence of global solutions of the convex quadratic program with convex quadratic constraints by an analytic approach.

Next, we study the global quadratic growth condition and the global second-order sufficient condition for the quadratic programming problems. By formulating the problem in the form of minimizing a maximum function of a finite number of quadratic functions, we study the relation between the global quadratic growth condition and the global second-order sufficient condition for the maximum function. As is shown, the global second-order sufficient condition implies the global quadratic growth condition. But the reverse implication is in general not true. We show that, when the solution set is a singleton and the number of quadratic term is 2, the reverse implication holds. For the homogeneous quadratic case, these two are equivalent. We then apply the corresponding results to the constrained optimization problem in standard form.

Finally, we investigate the fractional programming problems. We present the S-lemma for the fractional functions. We also obtain the attainability property of various types of fractional programming problems.

Acknowledgments

I would like to express my deep gratitude to my supervisor, Professor Xiaoqi Yang, for his guidance and support throughout this research. It has been a precious experience for me to study under his supervision in the Department of Applied Mathematics at The Hong Kong Polytechnic University. Without him, completion of the task would have been impossible. Thanks are also due to the Research Committee of The Hong Kong Polytechnic University for the financial support during the course of my research study.

Contents

Chapter 1

Introduction

1.1 Quadratic Problem

Quadratically constrained quadratic program, which we refer to as the quadratic problem for simplicity, represents a special and also an important class of nonlinear programming problems. Any twice differentiable function can be approximated by a quadratic function in the neighborhood of a given point, so the quadratic problems are the most natural.

In a narrow sense, as most of the optimization books show, the quadratic problem refers to the problem of minimizing a quadratic function over a polyhedral constraint, i.e. over a system of linear inequalities and/or equalities constraints. There are several classes of problems that can be naturally expressed as this kind of quadratic problems. Examples of such problems can be found in planning and scheduling, game theory, problems involving economies of scale, facility allocation and location problems, quadratic assignment problems, problems in engineering design, and a number of problems in microeconomics. Besides, quadratic problems with linear constraints can be viewed as a generalization of the linear programming problem with a quadratic objective function. Therefore, it contains all linear programming problems.

The more general class of quadratic problems is the one with quadratic constraints. It arises from various practical applications including facility location, product planning, optimal design of water distribution networks, and most problems in chemical engineering design. In early years since it was first introduced in the seminal paper of Kuhn and Tucker [42], the only case considered was when there was only one quadratic constraint in the problem.

In spite of its simple and clear structure, quadratic problem is difficult to solve. Even if the Hessian of the objective function has a single negative eigenvalue, as is shown by Pardalos and Vavasis [54], quadratic problem is NP-hard. Other complexity results may also be found in [47, 51, 65, 71].

To characterize the global solutions and to design algorithms estimating the global solutions of the quadratic problems, the standard nonlinear programming theories and algorithms generally don't work except for a few cases, such as under some convexity assumptions. Their deficiency is due to the intrinsic multiextremality of the formulation. One can observe that local tools such as gradients, subgradients, and the second order constructions such as Hessians, cannot be expected to yield more than local solutions. One finds, for example, that a stationary point is often detected for which there is no guarantee of local minimality. Moreover, determining the local minimality of such a point is known to be NP-hard in the sense of computational complexity even in relatively simple cases. Apart from this deficiency in the local situation, classical methods do not recognize conditions for global optimality. For these reseasons global solution methods must be significantly different from standard nonlinear programming techniques. However, with its special structure, we expect more promising results for quadratic problems.

A popular and effective method of global optimization is successive approximation. Outer or inner approximation of the constraint set by a sequence of simpler sets, such as polyhedron, is a basic method of this kind. Also successive underestimation of the objective function by convex or polyhedral functions is commonly used. Another method is successive partition method, dividing the feasible set into smaller pieces and refining the partition as needed, such as the branch and bound method. Details of these methods may refer to the works of Horst and Tuy [37] and the references therein.

One particular quadratic problem that has received considerable attention is the trust region problem, which involves minimizing a quadratic function over a sphere.

Details may refer to [30, 68, 50, 28, 41, 33, 77, 78, 45, 60, 21].

Finally, we end this subsection by giving a practical model of quadratic problems [26].

Consider a case when n products are being produced, with x_i being the number of units of product i and c_i being the cost of production per unit of product i. Usually, as the number of units produced increases, the unit cost decreases. Often this can be correlated by a linear functional:

$$c_i = c_{i_0} + e_i x_i$$

where e_i is a negative quantity and c_{i_0} is a constant. Then, given constraints on the production demands and availabilities of each product, the problem of minimizing the total cost can be cast as

$$\min\{\sum_{i=1}^{m} x_i(c_{i_0} + e_i x_i) : x \in D\}$$

where D represents the demand and availability constraints, which are normally represented by quadratic functions.

1.2 Existence of Optimal Solution

When dealing with optimization problems, first and foremost, one cares about the existence of solutions or the solvability of the problems. Even under the convex assumptions, the optimization problems may have no optimal solutions, such as the problem of minimizing an exponential function without constraints.

Baiocchi et al. [20] studied the following general optimization problem

$$\min_{x \in X} f(x),$$

where $f(x)$ is an extended real-valued functional defined on a topological vector space (X, σ). Under some coerciveness assumptions on f, we can easily get an existence theorem. To relax the condition, one need to investigate the behavior at infinity of the functional f. Therefore, by introducing the recession functionals (for convex cases) or topological recession functionals (for general cases), Baiocchi et al. derived the sufficient conditions for the existence of solutions of the problem and gave many applications in mechanics. Motivated by the works of Baiocchi et al., Auslender [4] refined the conditions given by them and made it being necessary conditions for the existence of optimal solutions of the problem. As an application, a new class of functions, termed as asymptotically multipolyhedral functions, was introduced by Auslender and related existence results were obtained. More details may also be found in [5, 6, 7].

In the finite dimension case, if $f(x)$ is convex, we may refer to Rockefellar [63] where various existence results are given and some duality characterizations of optimal solutions also are obtained.

In quadratic case, Frank and Wolfe [27] first considered the existence of a global solution of the quadratic problem with linear constraints. They showed that a quadratic function bounded below over a polyhedron always attains its global optimal solution. This is also known as the Frank-Wolfe theorem.

Several alternative proofs have been proposed by other authors. For example, Blum and Oettli [13] offered an elementary analytical proof for the Frank-Wolfe theorem. Eaves [23] gave a simplified proof of the Frank-Wolfe theorem via the linear complementarity theory and Eaves improved on the theorem by showing that if the problem does attain its minimum over the polyhedron, it is unbounded from below on some halfline contained in the polyhedron. Meanwhile, Eaves obtained the necessary and sufficiency conditions for the existence of solutions of the quadratic problem. This provides us a criteria to check whether a quadratic function is bounded on a polyhedral set or not. More details of the Frank-Wolfe theorem can be found in the book of Lee et al. [43].

Many authors have discussed the generalizations of the Frank-Wolfe theorem. Perold [56] gave sufficient conditions under which a function either attains its minimum over a convex polyhedral set or is unbounded from below on some halfline of that set. Luo and Zhang [44] studied various variants of the quadratic problem and obtained a series of existence theorems for the general quadratically constrained quadratic programs. Belousov and Klatte [9] generalized the theorem to convex polynomial programming. Recently, Obuchowska [52] generalized the Frank-Wolfe theorem to the programs with faithfully convex or quasi-convex polynomial objective functions and the feasible set defined by a system of faithfully convex inequalities and/or quasi-convex polynomial inequalities. Some other references for generalizations of the Frank-Wolfe theorem can be found in the article of Belousov and Klatte [9] and the references therein.

Only requiring the related functions to recede or retract along the asymptotical nonpositive directions of the functions, Ozdaglar and Tseng [53] presented a unified approach to establishing the existence of global minima of a constrained optimization problem. Their results generalize some results presented by Auslender and Luo and Zhang and also generalize the Frank-Wolfe theorem. For example, instead of requiring the objective function to be quadratic, they only require it to be a polynomial over a polyhedral set.

Bertsekas and Tseng [12] studied the nonemptiness property of the intersection of a nested sequence of closed sets and applied it directly to obtain some existence results of the optimization problems.

Some other conditions for global optimality may refer to the articles of Hiriart-Urruty [34, 35, 36].

The Frank-Wolfe theorem also has various applications. For example, Cottle et al. [22] used it as a main tool for obtaining the existence results for linear complementarity problems.

1.3 Second Order Analysis

Second order sufficient conditions of optimization problems are important for sensitivity analysis and numerical optimization, see Robinson [61], Fiacco and McCormick [25], Ben-Tal and Zowe [10] and Polak [57]. In fact various results concerning optimality conditions were obtained as byproducts of research on sensitivity analysis. Many authors have studied the perturbation properties of the optimal solution and the value function under various conditions. The differentiability properties of the optimal solutions were first obtained by applying the classical implicit function theorem to the first order optimality conditions written in the form of equations see Fiacco [24, 25]. The hypothesis of linear independence of the gradients of active constraints, strict complementarity, and second order sufficient condition were needed. Jittorntrum [40], by relaxing the strict complementarity hypothesis and using a strong second order sufficient condition, obtained the directional differentiability of a local solution. Shapiro [66], using the Mangasarian-Fromovitz constraint qualification and some second order sufficient condition, gave a second order analysis of the value function. The Mangasarian-Fromovitz constraint qualification can be relaxed by using the hypothesis of Gollan [31, 32].

As we have seen in the traditional approaches for sensitivity analysis, the problems need to satisfy some non-degeneracy assumptions, such as the assumption that the unperturbed problem has a unique solution or finite set of isolated points, some constraint qualification, and a second order sufficient condition. And therefore, when dealing with problems with non-isolated minima, the sufficient conditions cannot be applied. To study a wider class of problems and to overcome these defects, some other tools are needed. More attentive analysis of existing proofs, see, e.g. [1, 3, 18, 39, 57, 64, 67], shows, that, at least as far as sensitivity analysis is concerned, what is needed and efficient is the following quadratic growth condition:

$$f(x) \geq c + \alpha \mathrm{dist}^2(S, x),$$

where f is the cost function, S is a set on which f has constant value c, and α is a positive parameter. Also, it turns out that stability properties of locally optimal solutions, in a neighborhood of a feasible point, are closely related to a uniform version of a quadratic growth condition.

In the case where S is a singleton, the standard second-order sufficient conditions for optimality are sufficient for the quadratic growth condition; in fact, they often characterize the quadratic growth condition in the presence of some constraint qualification, such as the Mangasarian-Fromovitz constraint qualification, see Alt [2], Robinson [62].

This quadratic growth condition is also referred to as weak sharp minima of order two by some authors. More generally, this type property is called weak sharp minima of order m, for m being positive integers. For the characterizations of weak sharp minima of order one, one may refer to the works of Burke and Ferris [19]. Characterization and sufficient conditions for the quadratic growth condition are given by Bonnans and Ioffe [16]. In [16] they studied the relations between the general second order sufficient condition and the quadratic growth condition for the unconstrained optimization of a simple composite function (maximum of a finite collection of smooth functions) and some sufficient conditions for the quadratic growth condition. In [15] they obtained the characterization of the growth condition for the optimization problem in which the objective and constraint functions are convex and smooth functions. Using the tools of the Mordukhovich normal cone and the generalized directional derivatives, Ward gave some necessary conditions for the quadratic growth condition not requiring the data being twice differentiable in [72] and gave some sufficient conditions in [73], respectively. In addition, Studniarski and Ward [70] proposed sufficient conditions and characterizations of the quadratic growth condition for nonsmooth programming.

Shapiro [66] studied the perturbed mathematical programming problem with the optimal solution set being not a singleton and showed that the optimal set-valued solution function is upper-Lipschitzian and the optimal value function possesses a second order derivative under some regularity assumption. Under the quadratic growth condition assumption and other regularity conditions, Shapiro [67] studied the perturbed optimization problem in Banach spaces and derived the Lipschitz continuity and directional differentiability properties of the optimal solution. In addition, without a priori regularity assumption, Ioffe [39] studied in detail the perturbation properties of the value function and the optimal solution set for unconstrained optimization involving a simple composite function, under the growth-like condition hypothesis. The quadratic growth condition determines the rates of convergence of the optimal solutions as the perturbation goes to zero. This role of quadratic growth condition can be also found in many classical analysis of convergence rate of various optimization algorithms. For example, Polak [57] recurrently assumed the quadratic growth condition, although not explicitly stated it, in the convergence rate analysis, such as of the Armijo gradient method for unconstrained optimization, of the Newton method for min-max problems and of algorithms for optimal controls, etc.. More researches on perturbation analysis of optimization problems may refer to the review by Bonnans and Shapiro [17], the book of Bonnans and Shapiro [18], and the references therein.

Finally, we review some results on the global second order optimality conditions of optimization problems. Yang [75, 76] proposed a second order sufficient condition of a global solution by introducing a generalized representation condition. Quadratic functions and linear fractional functions satisfy this representation condition. For the quadratic problems, second order optimality conditions for global solutions are also obtained for some problems of specific structures. For example, Gay [30] and Sorenson [68] characterized the global solution of the trust region subproblem; Moré [49] studied the quadratic problem with one quadratic constraint and obtained the necessary and sufficiency optimality conditions for the global solution. For the quadratic problem with a two-sided constraint, Stern and Wolkowicz [69] obtained the optimality conditions for a global solution. They also showed that this problem is an implicit convex problem by the duality argument. For all the above three special cases of quadratic problems, we can see that there is no gap between the necessary and sufficient optimality conditions for the global solution. However, for the quadratic problems with two quadratic constraints, Peng and Yuan [55] showed that the Hessian of the Lagrangian has at most one negative eigenvalue at a global solution and this condition is not sufficient. More results may be found in Hiriart-Urruty [35, 36].

In this thesis, we will mainly concentrate on the studying of the quadratic programming problems, minimizing a quadratic function over a system of quadratic inequalities and/or equalities constraints. Specifically, we will study the existence of global solutions of a given quadratic problem and also study the global second-order optimality conditions and its relations with the quadratic growth condition and finally we extend some results to the quadratic fractional program problems.

Chapter 2

Existence of solutions of QP

2.1 Introduction

Existence of solutions is always a central issue concerning to optimization problems. In this chapter, we will consider the existence of solutions of the general quadratic programs or the extensions of Frank-Wolfe type theorems. As we all know, a fundamental theorem of linear programming is that a bounded feasible problem always has a solution. As for the general nonlinear optimization problem, we can hardly say so. For quadratic programs, Frank and Wolfe [27] stated that a quadratic function (not necessary convex) bounded from below over a polyhedron attains its minimum. This result is known as the Frank-Wolfe theorem. A short and elegant proof was also proposed by Frank and Wolfe. Alternative proofs have been proposed by Blum and Oettli [13] and Eaves [23].

Luo and Zhang [44] considered generalizations of the Frank-Wolfe theorem. They obtained the following results for the quadratical programming problem with quadratic constraints:
(1) If the objective function is convex and at least one of the constraint functions is nonlinear and non-convex, the optimal value is in general not attainable.
(2) If the objective function is non-convex and at least two or more constraint functions are nonlinear, the optimal value is in general not attainable.
(3) If the objective function is non-convex and at most one of the constraint functions

are nonlinear but convex, the optimal value is always attained.
(4) If the objective function is quasi-convex over the feasible region and all the constraint functions are convex, the optimal value is always attained.

In this chapter, we will make some complements to Luo and Zhang's results. We will provide a new proof for the continuity of the sequence of nested sets defined by a system of convex quadratic inequalities which can be applied to prove the attainability of the convex quadratic program with convex quadratic constraints. We will also prove the attainability of the convex quadratic program with convex quadratic constraints following the scheme proposed by Blum and Oettli.

2.2 Extensions of Frank-Wolfe Theorem

We begin with some elementary definitions used throughout this section.

We will confine our study in the finite dimensional Euclidean space R^n. For a nonempty closed convex set $C \subset R^n$, the recession cone C^∞ of C is defined by $C^\infty = \{d \in R^n | x + \tau d, \forall \tau \geq 0, x \in C\}$. $L_C := C^\infty \cap -C^\infty$ denotes the linearity space of C. For a proper lower semi-continuous (lsc) convex function $f : R^n \to R$, the recession function f^∞ of f is defined by

$$\text{epi } f^\infty = \text{epi f}^\infty.$$

For convex quadratic function $f(x) = \frac{1}{2}x^t Qx + q^T x + c$, we have

$$f^\infty(d) = q^T d + \delta(d|Qd = 0), \tag{2.2.1}$$

where δ denotes the indicator function.

First, we take a look at a known result and its proof by Frank and Wolfe [27].

Theorem 2.2.1 (Frank-Wolfe theorem). *If the problem*

$$\min\{f(x) = x^t Qx + q^T x : Ax \leq b, x \in R^n\}$$

is bounded from below, then the minimizer exists.

Proof. We may assume that the feasible set is unbounded, since a continuous function always attains its extremum over a compact set.

We use induction on the dimension m of the polyhedron.

For $m = 1$, let the feasible set F be of the form $x + tv$ for some $x, v \in R^n$ and $t \geq 0$ (if t takes all the value in R, the problem is unconstrained, a trivial case). Then, since $f(x+tv) = f(x) + t(q^T v + 2x^T Qv) + t^2 v^T Qv$ for all $t \geq 0$ is bounded from below, $v^T Qv > 0$ or $v^T Qv = 0$ and $q^T v + 2x^T Qv \geq 0$. In either case, the minimum is attained.

Assume that the result is true for the case $m = k$. Now, consider the case $m = k+1$. Assume the feasible set F of the form $\{s + \mu t | s \in S, t \in T, \mu \geq 0\}$ with S being a bounded convex polyhedron and T an intersection of a certain convex cone with the unit sphere. Since for $s \in S, t \in T$, $f(s+\mu t) = f(s) + \mu(q + 2s^T Q)t + \mu^2 t^T Qt$ is bounded for $\mu \geq 0$, $t^T Qt \geq 0$ for all $t \in T$.

If, on the one hand, $t^T Qt > 0$ for all $t \in T$, then there exist $\delta > 0$ and $D < 0$ such that $t^T Qt \geq \delta$ and $(q + 2s^T Q)t \geq D$ for all $t \in T$ and $s \in S$, so that the minimum of f is assumed on the compact set $S + \frac{-D}{\delta} T$.

Suppose, on the other hand, some $t_0^T Qt_0 = 0$. If for all $x \in F$ we have $x + \mu t_0 \in F$ for all μ (all the real numbers), then the boundedness of $f(x+\mu t)$ implies $(q + 2x^T Q)t_0 = 0$ for all x, so that the values of f on F are unchanged by projection into the k-dimensional subspace normal to t_0, to which the induction hypothesis may be applied.
Otherwise, if for some $x_0 \in F$, $x_0 + \mu t_0 \notin F$ for some $\mu (< 0)$, then for all $x \in F$, $x + \mu t_0 \notin F$ for some μ. So, for each $x \in F$, $b_x := x + \min\{\mu | x + \mu t_0 \in F\} t_0$ lies on the boundary of F, and $f(b_x) \leq f(x)$, since $(q + 2x^T Q)t_0 \geq 0$. Since the minimum of f on each k-dimensional bounding hyperplane of F is assumed, so is it on F. $\square$

Lemma 2.2.1. *[11] Let C a nonempty convex subset of R^n. Then for every subspace S that is contained in the lineality space of C, we have*

$$C = S + (C \cap S^{\perp}) \tag{2.2.2}$$

Lemma 2.2.2. *[7] For any proper function $f : R^n \to R \cup \{+\infty\}$ and any $\alpha \in R$ such that the level set $\{x \in R^n | f(x) \leq \alpha\} \neq \emptyset$ one has*

$$\{x \in R^n | f(x) \leq \alpha\}^{\infty} \subset \{d \in R^n | f^{\infty}(d) \leq 0\} \tag{2.2.3}$$

Equality holds in the inclusion when f is lsc, proper, and convex.

From this lemma, we have that any nonempty level sets of a convex quadratic function have the same recession cone.

Lemma 2.2.3. *[7] Let $f_i : R^n \to R \cup \{+\infty\}, i \in I$ be a collection of proper functions and let $S \subset R^n$ with $S \neq \emptyset$. Define $C := \{x \in S | f_i(x) \leq 0, \forall i \in I\}$. Then*

$$C^\infty \subset \{d \in S^\infty | (f_i)^\infty(d) \leq 0, \forall i \in I\}. \tag{2.2.4}$$

The inclusion holds as an equation when $C \neq \emptyset$, S is closed and convex, and each f_i is lsc convex.

Lemma 2.2.4. *[7] For any nonempty closed convex set $C \subset R^n$ that contains no lines one has $C = \mathrm{conv}(\mathrm{ext}\ C) + C^\infty$.*

Lemma 2.2.5. *[63](Helly's Theorem) Let $\{C_i | i \in I\}$ be a collection of nonempty closed convex sets in R^n, where I is an arbitrary index set. Assume that the sets C_i have no common direction of recession. If every subcollection consisting of $n+1$ or fewer sets has a nonempty intersection, then the entire collection has a nonempty intersection.*

Proposition 2.2.1. *Let $X(\epsilon) = \{x \in R^n | f_i(x) = x^T Q_i x + 2q_i^T x + c_i \leq \epsilon_i, i = 1, \cdots, m\}$ with Q_i being positive semi-definite. Assume that $X(\epsilon^k)$ is nonempty for some positive sequence $\epsilon^k = (\epsilon_1^k, \cdots, \epsilon_m^k)^T, k = 1, 2, \cdots$ approaching nonincreasingly to zero. Then, the set $X(0)$ is also nonempty.*

With this proposition, as Luo and Zhang [44] shows, the attainability of the convex program with only convex quadratic data involved easily follows. Below we provide a new proof of this proposition. As to Luo and Zhang's proof in [44], from the linear system (4) on Page 90, they obtained the existence of a solution $\bar{x}^k$ and a constant ρ, independent of k such that they satisfy some inequality. It seems that the constant ρ is dependent of k if the argument follows from the Hoffman bound. Our proof uses some basic tools, linear transformation and decomposition of set, to avoid the argument used in [44].

Proof of Proposition 2.2.1. First, all the sets $X(\epsilon^k)$ are nonempty closed convex sets and $\bigcap_{k \in I} X(\epsilon^k) \neq \emptyset$ for any finite index set I by the monotonicity of the sets $X(\epsilon)$. We know from Lemma 2.2.3 that $X(\epsilon^k)$ have the same recession cone, say X^∞. And we may assume that for any k, $X(\epsilon^k)$ contains no lines.

Otherwise, let d be the direction of the line. Then d, $-d$ are both in X^∞. Let $L = X^\infty \cap -X^\infty$ be the linearity of $X(\epsilon^k)$. Hence, from lemma 2.2.1, there exists a decomposition $X(\epsilon^k) = L + (X(\epsilon^k) \cap L^\perp)$ with $X(\epsilon^k) \cap L^\perp$ having no lines. And we also have $X(\epsilon^k) \cap L^\perp \subset (X(\epsilon^{k-1}) \cap L^\perp)$. Hence, we may consider $X(\epsilon^k) \cap L^\perp$ instead of $X(\epsilon^k)$.

$m = 1$. After a nonsingular linear transformation, we may assume that $f_1(x) = \sum_{1\le i\le r} \lambda_i (x_i - x_i^0)^2 + \sum_{r+1\le i\le k} q_i x_i + c, r \le k \le n$. If $X(\epsilon^k) \neq \emptyset$, then we can easily choose a point $\bar{x}$ such that $f_1(\bar{x}) \le 0$. Just consider two cases: if $r = n$, then $c \le 0$ and hence take $\bar{x}_i = x_i^0$; if $r < n$, it follows similarly. This implies that $X(0) \neq \emptyset$.

By induction, assume that for $m = k$, the conclusion is true. Now consider the case $m = k+1$. Let $x_k \in X(\epsilon^k)$ be a sequence with the smallest norm. We claim that $\{x_k\}$ is bounded. Otherwise, assume that $\|x_k\| \to \infty$ and $\frac{x_k}{\|x_k\|} \to d$. Then $Q_i d = 0, q_i^T d \le 0$.

Case 1. There is some i_0, such that $q_{i_0}^T d < 0$. Let $i_0 = k+1$. Then by induction assumption, there is an $\bar{x}$ such that $f_i(\bar{x}) \le 0, 1 \le i \le k$. Then $f_i(\bar{x}+td) \le 0, 1 \le i \le k$ and $f_{k+1}(\bar{x}+td) = f_{k+1}(\bar{x}) + t q_{k+1}^T d \le 0$ for t large enough. This is to say, $\bar{x}+td \in X(0)$ for some t, contradicting the unboundedness of the sequence $\{x_k\}$.

Case 2. $q_i^T d = 0$ for all $i = 1, \cdots, k+1$. Then for any $x \in X(\epsilon^k), x+td \in X(\epsilon^k), \forall t \in R$. This contradicts the assumption that $X(\epsilon^k)$ contains no lines.

This completes the proof. $\square$

Next, we prove the existence of a solution of the convex quadratically constrained quadratic problem by adopting the analytical method of Blum and Oettli [13] proving the Frank-Wolfe theorem.

Proposition 2.2.2. *[44] Consider the convex quadratically constrained quadratic problem*

$$\begin{array}{rcl} \min\ f_0(x) & = & x^T Q_0 x + 2q_0^T x + c_0 \\ \text{subject to } f_i(x) & = & x^T Q_i x + 2q_i^T x + c_i \le 0, i = 1, \cdots, m, \end{array} \tag{QCQP}$$

where $f_i(x) = x^T Q_i x + 2q_i^T x + c_i, i = 0, 1, \cdots, m,$ are convex.
Assume the feasible set $F = \{x \in R^n | f_i(x) \le 0, i = 1, \cdots, m\}$ is nonempty and the objective value is bounded from below on F. Then there is a global solution for the problem.

Proof. Let $I = \{1, \cdots, m\}$. Assume that $x = 0$ is in F. Then F_ρ defined by $F \cap B(0, \rho)$ is nonempty compact. Let $\hat{V} = \inf_{x \in F} f_0(x)$ and $V(\rho) = \inf_{x \in F_\rho} f_0(x)$. So, $V(\rho)$ is monotonely nonincreasing and $\lim_{\rho \to \infty} V(\rho) = \hat{V}$. Since F_ρ is compact, there exists at least one $x_\rho \in F_\rho$ such that $f_0(x_\rho) = V(\rho)$. We claim that $\{x_\rho\}$ is bounded. In this case the result holds obviously. In the latter part, we prove by contradiction that the converse case couldn't occur. Without loss of generality, we can choose x_ρ to be the one with this property of minimal modulus.

Assume on the contrary we can find a sequence ρ_k with $\rho_k \to \infty, \|x_{\rho_k}\| = \rho_k$. Put $x_k = x_{\rho_k}$ for brevity. Since $x_k \in F$, $f_i(x_k) \leq 0, \forall i \in I$. Let $I_0 = \{i \in I | \limsup_{k \to \infty} f_i(x_k) = 0\}$ and $I_1 = I \setminus I_0$.
Choose an $\epsilon > 0$ and a subsequence $\{\rho_j\}$ of $\{\rho_k\}$ such that

$$f_i(x_j) \to 0, \forall i \in I_0, \text{ and } f_i(x_j) \leq -\epsilon, \forall i \in I_1. \tag{2.2.5}$$

Taking a subsequence if necessary, let $t_j = \frac{x_j}{\rho_j} \to t$. Then t is a unit vector.
Since $f_i(x_j) = \rho_j^2 t_j^T Q_i t_j + 2\rho_j q_i^T t_j + c_i$ and, since $\rho_j \to \infty$, it follows in the limit and (2.2.5) that

$$t^T Q_i t = 0, q_i^T t \leq 0, \forall i \in I_0; \tag{2.2.6}$$

$$t^T Q_i t = 0, q_i^T t \leq 0, \forall i \in I_1. \tag{2.2.7}$$

Similarly, by $f_0(x_j) \to \hat{V}$, we have

$$t^T Q_0 t = 0. \tag{2.2.8}$$

We consider two different subcases: (i) $q_i^T t = 0, \forall i \in I_0, (ii)$ $q_{i_0}^T t < 0$, for some $i_0 \in I_0$.

Case (i): Since $f_i(x_j + st) = f_i(x_j) + 2s(x_j^T Q_i t + q_i^T t) + s^2 t^T Q_i t$,
it follows from (2.2.6) and (2.2.7) that for $i \in I_0$, $f_i(x_j + st) = f_i(x_j) \leq 0, \forall s \geq 0$,
for $i \in I_1$, $f_i(x_j + st) \leq -\epsilon + 2s q_i^T t \leq -\epsilon, \forall s \geq 0$,
So,

$$x_j + st \in F, \forall s \geq 0. \tag{2.2.9}$$

By $f_0(x_j + st) \geq \hat{V}$, we similarly get

$$q_0^T t \geq 0. \tag{2.2.10}$$

Now, choose ρ_j so large that $t_j^T t > 0$. Then, because of $x_j^T t \to \infty$, it follows that

$$\|x_j - \lambda t\| < \|x_j\| = \rho_j \text{ for } \lambda \text{ small.} \tag{2.2.11}$$

We also have, for all $i \in I_0$,

$$f_i(x_j - \lambda t) = f_i(x_j) \le 0, \forall \lambda, \tag{2.2.12}$$

and, for all $i \in I_1$,
$f_i(x_j - \lambda t) = f_i(x_j) - 2\lambda q_i^T t \le -\epsilon - 2\lambda q_i^T t \le 0$, for all λ small enough.
Therefore,

$$x_j - \lambda t \in F \text{ for } \lambda \text{ small enough.} \tag{2.2.13}$$

On the other hand, it follows from (2.2.8) and (2.2.10) that $f_0(x_j - \lambda t) = f_0(x_j) - 2\lambda(x_j^T Q_0 t + q_i^T t) \le f_0(x_j)$, $\forall \lambda > 0$.

Hence, we obtained that for $x^* = x_j - \lambda t$,for some λ, $x^* \in F, \|x^*\| < \|x_j\|$, and $f_0(x^*) < f_0(x_j)$. This contradicts the definition of x_j as being a solution of $\min\{f_0(x) | x \in F_{\rho_j}\}$ having minimal modulus. Therefore, the optimal solution exists.

Case (ii): Let $I_0' = \{i \in I_0 | q_i^T t = 0\}$ and $I_0 \backslash I_0' = \{1, \cdots, r\}, r \ge 1$. Claim that there is an optimal solution $\bar{x}$ for problem (QCQP), which contradicts the unboundedness of the sequence $\{x_\rho\}$ as the solution the approximate problems. Thus the sequence $\{x_\rho\}$ is bounded and the Proposition follows.

We make induction on r. For $r = 1$. Consider the problem

$$(P') \quad \min f_0(x) \text{ s.t. } f_i(x) \le 0, i \in I_0' \cup I_1.$$

There are three cases for problem (P'): $\inf(P') = -\infty$; $\inf(P') < \hat{V}; \inf(P') = \hat{V}$. In the third case, the infimum is attainable from the proof of case (i). In all, by the continuity of f_0, there is an x' feasible for (P') such that $f_0(x') = \hat{V}$.

If $f_1(x') \le 0$, then x' is a solution of (QCQP). Assume $f_1(x') > 0$. Since $f_0(x_j) \to f_0(x')$ and f_0 is convex, $(x_j - x')^T \nabla f_0(x') = f_0(x_j) - f_0(x') - \frac{1}{2}(x_j - x')^T Q_0 (x_j - x') \le f_0(x_j) - f_0(x')$. Dividing $\|x_j - x'\|$ on both side of the above inequality and taking the limit, $t^T \nabla f_0(x') \le 0$. Let $s^* = -\frac{f_1(x')}{q_1^T t}$ and $\bar{x} = x' + s^* t$. We have $s^* > 0$. We may easily check that $f_1(\bar{x}) = 0, f_i(\bar{x}) \le 0, i \in I_0' \cup I_1$ and $f_0(\bar{x}) \le f_0(x') = \hat{V}$. This is to say $\bar{x}$ is an optimal solution for (QCQP).

Assume the claim is true for $k < r$. Let $k = r$ and consider the problem

$$(P'') \quad \min f_0(x) \ \text{ s.t. } \ f_i(x) \le 0, i \in I \backslash \{r\}.$$

Similarly, there are three cases for the problem (P'') and by the induction assumption, there is an x' feasible for (P'') such that $f_0(x') = \hat{V}$. Also,two subcases are to be considered. If $f_r(x') \le 0$, it is trivial. If $f_r(x') > 0$, we have $t^t \nabla f_r(x') \le 0$, as is shown in the case $m = 1$. Similarly, define $s^* = -\frac{f_r(x')}{q_r^T t}$ and $\bar{x} = x' + s^* t$. We obtain an optimal solution $\bar{x}$ for the original problem (QCQP).

This completes the proof. □

Remark. Belousov and Klatte pointed out in [9] that this result is even true if $f_0, f_1, \cdots, f_m$ are convex polynomials of arbitrary order, as proved by Belousov in his 1977 book [8]. Since this book is written in Russian, it is not readily available (at least in the West) and hence a wide part of the optimization community is not aware of this fact until recently.

Chapter 3

Relations between QGC and GSO

3.1 Introduction

Second order optimality conditions of an optimization problem are closely linked with the sensitivity analysis and with the study of the convergence properties of numerical algorithms for solving optimization problems, see [18, 24, 25]. In fact, various results concerning optimality conditions were obtained as byproducts of researches on sensitivity analysis. Another condition closed linked with the sensitivity analysis is the quadratic growth condition [14, 29, 39, 66, 67]. For standard nonlinear programming problems, the weak second order sufficient condition is equivalent to the quadratic growth condition as far as the set of minima consists of isolated points and some constraint qualification hypothesis holds [10, 14, 38]. In case of non-isolated minima set, Bonnans and Ioffe [16] studied the relations between the general second order sufficient condition and the quadratic growth condition. In [15], they devoted to a special case, the convex problems, and gave a complete characterization.

Yang [75, 76] proposed a second order sufficient condition of a global solution by introducing a generalized representation condition. Quadratic functions and linear fractional functions satisfy this representation condition. For the quadratic problems, second order optimality conditions for global solutions are also obtained for some problems of specific structures, such as the trust region subproblem, quadratic problem with one or two quadratic constraints and quadratic problem with a two-sided quadratic constraint, see Sorenson [68], Moré [49], Peng and Yuan [55] and Stern and Wolkowicz [69].

In this chapter, we will give some characterizations of the global quadratic growth condition and the global second order sufficient condition. Due to the hardness of the global characterization, we will only study some specially structured problems, such as the problem of min-max form and the quadratically constrained quadratic problems. We can't expect more for the global condition of the general problems. We will generally follow the scheme proposed by Bonnans and Ioffe [16]. We will first study the equivalence of the global quadratic growth condition and the global second-order sufficient condition for the problem of min-max form. At last, we will apply these results to the standard quadratically constrained quadratic problem via a reformulation.

3.2 The global quadratic condition and the global second order condition

In this section, we study the problem of the following min-max form

$$\min_{x \in R^n} f(x) := \max_{1 \le i \le m} f_i(x), \tag{3.2.1}$$

where $f_i : R^n \to R, i = 1, \cdots, m$, are real-valued functions.

To start with, we introduce some notations and terminologies used in this section. The index set $I(x) := \{i : 1 \le i \le m, f_i(x) = f(x)\}$ denotes the set of active indices of $f(x)$ at x. The function

$$\mathcal{L}(\lambda, x) := \sum_{i=1}^{m} \lambda_i f_i(x), \tag{3.2.2}$$

is defined as the Lagrangian of $f(x)$. The set

$$\mathscr{S}^m := \{\lambda \in R^m : \lambda \ge 0, \sum_{i=1}^{m} \lambda_i = 1\}, \tag{3.2.3}$$

denotes the standard simplex of R^m. The set

$$\Omega(x) := \{\lambda \in \mathscr{S}^m : \lambda_i \ge 0, \lambda_i = 0 \text{ if } i \notin I(x); \sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0\} \tag{3.2.4}$$

is the set of Lagrange multipliers of f at x and

$$\Omega_\delta(x) := \{\lambda \in \mathscr{S}^m : \lambda_i \ge 0, \lambda_i = 0 \text{ if } i \notin I(x); \| \sum_{i=1}^{m} \lambda_i \nabla f_i(x)\| \le \delta\} \tag{3.2.5}$$

the set of Lagrange $\delta-$multipliers.

In the following, suppose that $f(x)$ is a constant c_0 on the set S.

Definition 3.2.1. *A mapping π from a neighborhood U of S onto S will be called a* regular projection *onto S if there exists $\varepsilon > 0$ such that*

$$\varepsilon\|x - \pi(x)\| \le \mathrm{dist}(S, x), \quad x \in U. \tag{3.2.6}$$

Definition 3.2.2. *We say f satisfies the* quadratic growth condition *(QGC) on S if :*

there exists $\beta > 0$ and a neighborhood U of S such that

$$f(x) \ge c_0 + \beta \mathrm{dist}^2(S, x), \quad \forall x \in U. \tag{3.2.7}$$

We say that the (QGC) holds globally if the inequality (3.2.7) holds for all $x \in R^n$.

Definition 3.2.3. *We say f satisfies the* general second-order sufficient condition *(GSO) on S if :*

for any $\delta > 0$ there exists a neighborhood U of S,

a regular projection $\pi : U \to S$ and $\alpha > 0$ such that, for all $x \in U\backslash S$,

$$\max_{\lambda\in\Omega_\delta(\pi(x))} [\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h)] \ge \alpha\|h\|^2, \tag{3.2.8}$$

where $h = x - \pi(x)$.

We say the (GSO) holds globally if the inequality (3.2.8) holds for all $x \in R^n$.

Definition 3.2.4. *Let C, D be sets and $x \in C\bigcap D$. We say that C and D are nontangent at x if*

$$T_C(x)\bigcap T_D(x) = \{0\}.$$

Definition 3.2.5. *We say that f satisfies the tangency condition (TC) on $D \subset R^n$ if for any x in D and for any $i \in I(x)$, either $i \in I(y)$ for all $y \in D$ sufficiently close to x, or D and $\{y : f_i(y) = f_i(x) = c_0\}$ are nontangent at x.*

For $m = 1$, we have the following results.

Proposition 3.2.1. *If $f(x) = x^tQx + 2q^tx$, then global (QGC) is equivalent to global (GSO).*

Proof. We have $\Omega(x) = \Omega_\delta(x) = \{1\}$ and $\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h) = 2h^t(Qx + q) + h^tQh$.

Suppose that (QGC) holds locally on S. Then S is the local solution set of f , $\nabla f(x) = 2(Qx + q) = 0$ on S and $Q \succeq 0$. And $S = \{x : Qx + q = 0\}$, an affine set. Let π be the projection mapping on S. Then for any $x \in R^n \backslash S$,

$$\max_{\lambda \in \Omega_\delta(\pi(x))} [\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h)] = (x - \pi(x))^tQ(x - \pi(x)) \geq \lambda_{\min}\|x - \pi(x)\|^2,$$

where $\lambda_{\min}$ is the smallest nonzero eigenvalue of Q. The last inequality follows from the spectral theorem for the symmetric matrix Q and the fact that $x - \pi(x)$ is orthogonal to the null space of Q.

Now suppose that global (GSO) holds and then the local (GSO) holds, i.e. $\forall \delta > 0, \exists U \in \mathcal{N}(S)$, a regular projection $\pi : U \to S$ and $\alpha > 0$, s.t. $\forall x \in U \backslash S$,

$$2h^t(Q\pi(x) + q) + h^tQh \geq \alpha\|h\|^2,$$

where $h = x - \pi(x)$. Then $\forall x \in U$ and $x_0 = \pi(x)$,

$$f(x) = f(x_0) + 2(x - x_0)^t(Qx_0 + q) + (x - x_0)^tQ(x - x_0) \geq c_0 + \alpha\|x - x_0\|^2 \geq c_0 + \alpha \mathrm{dist}(S, x)^2.$$

Hence (QGC) holds locally and therefore $\nabla f(x) = 0$ on S and $Q \succeq 0$. From the necessity proof, we can let π be the projection mapping on S.

Then for any $x \in R^n \backslash U$ and $x_0 = \pi(x)$,

$$f(x) = f(x_0) + (x - x_0)^tQ(x - x_0) \geq c_0 + \lambda_{\min}\|x - x_0\|^2 \geq c_0 + \lambda_{\min}\mathrm{dist}(S, x)^2,$$

where $\lambda_{\min}$ is the smallest nonzero eigenvalue of Q.

Let $\beta = \min\{\alpha, \lambda_{\min}\}$.

Then

$$f(x) \geq c_0 + \beta \mathrm{dist}(S, x)^2, \ \forall x \in R^n.$$

This completes the proof. $\square$

We want to know whether the (QGC) holds globally in the presence of the (GSO) hypothesis. Unfortunately, this is not in general true. In case of $m \geq 2$, when the (GSO) holds on a set S and S is the global optimal set of the quadratic program, the global (QGC) may not hold. The following is a simple counter example.

Example 3.2.1. *For $f(x) = \max\{x, -x\}$, $x = 0$ is the global solution and the (GSO) holds for any $|x| \leq 1$, and therefore, from Theorem 1 of [16], (QGC) holds. If $|x| \geq Cx^2$ for some positive constant C, we have $|x| \leq C$. It follows that the global (QGC) does not hold.*

However, under the global (GSO) condition assumption we have

Proposition 3.2.2. *Let $f(x) = \max\{f_i(x)\}$ and f_i be quadratic functions. Then the global (GSO) condition implies the global (QGC) condition.*

Proof. If the global (GSO) holds, for any $\delta > 0$, there exists a regular projection $\pi : R^n \to S$, and $\alpha > 0$ such that, for all $x \in R^n \backslash S$,

$$\max_{\lambda \in \Omega_\delta(\pi(x))} [\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h)] \geq \alpha \|h\|^2,$$

where $h = x - \pi(x)$.

Then, for any $x \in R^n$,

$$\begin{aligned} f(x) - c_0 &= f(\pi(x) + h) - f(\pi(x)) \\ &\geq \max_{\lambda \in \Omega_\delta(\pi(x))} [\mathcal{L}(\lambda, \pi(x) + h) - \mathcal{L}(\lambda, \pi(x))] \\ &= \max_{\lambda \in \Omega_\delta(\pi(x))} [\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h)] \\ &\geq \alpha \|h\|^2 \geq \alpha \text{dist}^2(x, S). \end{aligned}$$

Then the global (QGC) holds. □

Bonnans and Ioffe [16] showed that (QGC) and (TC) imply the (GSO). We want to know whether the global (QGC) and (TC) imply the global (GSO) or not.

The following is a counter example when $m = 4$.

Example 3.2.2. *Let* $f_1(x) = x^2 - \frac{1}{5}x - \frac{6}{5}$, $f_2(x) = -(x+1)^2 + \frac{1}{4}$, $f_3(x) = -x^2 + 2x$, *and* $f_4(x) = 0$.

Then $f(x) = \max_{1\le i\le 4} f_i(x) = \begin{cases} x^2 - \frac{1}{5}x - \frac{6}{5} & \text{if } x \in (-\infty, \frac{-9-\sqrt{139}}{20}] \cup [\frac{3}{2}, \infty) \\ -(x+1)^2 + \frac{1}{4} & \text{if } x \in (\frac{-9-\sqrt{139}}{20}, -\frac{1}{2}) \\ 0 & \text{if } x \in [-\frac{1}{2}, 0] \\ -x^2 + 2x & \text{if } x \in [0, \frac{3}{2}]. \end{cases}$

The global solution set $S = [-\frac{1}{2}, 0]$. Then it can be easily checked that $f(x) \geq \frac{1}{8}\text{dist}^2(x, S)$.The global (QGC) holds.

Next, we verify the tangency condition. For any $x \in \text{int}S$, $I(x) = \{3\}$. It follows that for any $y \in S$ close to x, $I(x) = I(y)$. If $x = -\frac{1}{2}$, then $I(x) = \{2, 3\}$. It is easy to see that $3 \in I(y)$ for any $y \in S$ close to x and $\{y : f_2(y) = f_2(x)\} = \{-\frac{1}{2}\}$, nontangent to S. The case for $x = 0$ is similar. Hence the (TC) condition also holds.

It remains to check that, for any $\delta > 0$,

$$\max_{\lambda\in\Omega_\delta(\pi(x))} [\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h)] \geq \alpha\|h\|^2,$$

for all $x \in R$ and some regular projection $\pi : R \to S$, where $h = x - \pi(x)$.

If $\pi(x) \in (-\frac{1}{2}, 0)$, then $\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h) = 0$. Then inequality would never hold. So, $\pi(x)$ takes the value either $-\frac{1}{2}$ or 0. Simple calculations indicate that $\mathcal{L}_x(\lambda, \pi(x)) \equiv 0$ and $\mathcal{L}_{xx}(\lambda, \pi(x)) = -2\lambda_3$ if $\pi(x) = 0, -2\lambda_2$ if $\pi(x) = -\frac{1}{2}$. Hence the global (GSO) does not hold since the left hand side is linear in h, which will less than $\alpha\|h\|^2$ for x large enough.

In case of $m = 3$, the answer is also negative.

Example 3.2.3. *Let* $f = \max\{x, -x, x^2\}$. *Then the global (QGC) and (TC) conditions hold but the global (GSO) condition does not.*

The above example involves affine functions. What about the case involving only non-affine quadratic functions? Consider the following case.

Example 3.2.4. *Let* $f(x) = \max\{f_1(x), f_2(x), f_3(x)\}$ *with* $f_1(x) = -x^2 + 2x, f_2(x) = -x^2 - 2x$ and $f_3(x) = 2x^2 - 1$. *We have that* $x = 0$ *is the global solution and* $f(x) \geq x^2$. *This is to say the global (QGC) condition holds. Since the optimal solution set is a singleton, the tangency condition (TC) holds trivially. However, the global (GSO)* *doesn't hold since the related Hessian is negative definite. In fact, since for any $x \in R^n$, $\pi(x) = 0$, $\max_{\lambda\in\Omega_\delta(\pi(x))}[\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h)] = \max_{\lambda\in\Omega_\delta(0)}\{(4\lambda - 2)x - x^2\}$. Thus, the global (GSO) couldn't hold.*

Now we consider the case of $m = 2$ and the global solution set is a singleton. First, we review a result by J. J. Moré [49].

Lemma 3.2.1. *Consider the following problem*

$$\min\{f_0(x) : f_1(x) \le 0\},$$

where $f_0, f_1 : R^n \to R$ are quadratic functions. Assume that $\min\{f_1(x) : x \in R^n\} < 0$ and $\nabla^2 f_1 \ne 0$. A vector x^* is a global solution of the problem if and only if there is a $\lambda \ge 0$ such that

$$f_1(x^*) \le 0,$$

$$\nabla f_0(x^*) + \lambda\nabla f_1(x^*) = 0,$$

$$\nabla^2 f_0(x^*) + \lambda\nabla^2 f_1(x^*) \ge 0.$$

Applying this Lemma, we may obtain the following result.

Proposition 3.2.3. *Assume that $\max\{f_0(x), f_1(x)\} \ge c + \alpha\|x - x_0\|^2$ where $f_i(x) = x^T Q_i x + 2q_i^t x + c_i, i = 0, 1,$ and α is a positive constant. Then the global (GSO) condition holds.*

Proof. It suffices to prove that there are multipliers $\lambda_0 \ge 0, \lambda_1 \ge 0$ such that $\lambda_0 + \lambda_1 = 1$, $\lambda_0\nabla f_0(x_0) + \lambda_1\nabla f_1(x_0) = 0$ and $\lambda_0 Q_0 + \lambda_1 Q_1 > 0$.

To avoid the trivial case, we may assume that $f_0(x_0) = f_1(x_0) = c$. Otherwise, there is at least one function being strictly convex. Note that $\max\{f_0(x), f_1(x)\} \ge c+\alpha\|x-x_0\|^2 \Leftrightarrow f_0(x)+\max\{0, f_1(x)-f_0(x)\} \ge c+\alpha\|x-x_0\|^2$. Let $y^2 = \max\{0, f_1-f_0\}$. Then $(x_0^T, 0)^T$ is the global solution of the problem (P') :

$$\min\{f_0(x) + y^2 - \alpha\|x - x_0\|^2 : g(x, y) := f_1(x) - f_0(x) - y^2 \le 0\}.$$

It is obvious that $g(x, y)$ satisfies the assumptions in the Lemma 3.2.1. Hence, there is a $\lambda \ge 0$ such that

$$\nabla f_0(x_0) + \lambda(\nabla f_1(x_0) - \nabla f_0(x_0)) = 0,$$

$$\begin{pmatrix} Q_0 - 2\alpha I & 0 \\ 0 & 2 \end{pmatrix} + \lambda \begin{pmatrix} Q_1 - Q_0 & 0 \\ 0 & 2 \end{pmatrix} \geq 0.$$

Similarly, $\max\{f_0(x), f_1(x)\} \geq c + \alpha\|x - x_0\|^2 \Leftrightarrow f_1(x) + \max\{0, f_0(x) - f_1(x)\} \geq c + \alpha\|x - x_0\|^2$.

Then there is a $\lambda' \geq 0$ such that

$$\nabla f_1(x_0) + \lambda'(\nabla f_0(x_0) - \nabla f_1(x_0)) = 0,$$

$$\begin{pmatrix} Q_1 - 2\alpha I & 0 \\ 0 & 2 \end{pmatrix} + \lambda' \begin{pmatrix} Q_0 - Q_1 & 0 \\ 0 & 2 \end{pmatrix} \geq 0.$$

When $\lambda > 0, \lambda' > 0$, we have $\lambda'\nabla f_0(x_0) + \lambda\nabla f_1(x_0) = 0$, and $\lambda'\nabla^2 f_0(x_0) + \lambda\nabla^2 f_1(x_0) \geq 2\alpha\lambda\lambda' I$.

Therefore, there are λ_0, λ_1 such that $0 \leq \lambda_0, \lambda_1 \leq 1$, $\lambda_0 + \lambda_1 = 1$, $\lambda_0\nabla f_0(x_0) + \lambda_1\nabla f_1(x_0) = 0$ and $\lambda_0 Q_0 + \lambda_1 Q_1 \geq \alpha' I > 0$. This completes the proof. □

However, for the case the solution set is not a singleton, we have not obtained the implication yet. In general case, under what circumstances or with what strengthened conditions will the global quadratic growth condition and the global general second-order sufficient condition be equivalent?

As a special case, we consider the following problem of homogeneous quadratic forms:

$$\min_{x \in R^n} \; f(x) = \max_{1 \leq i \leq m} x^T Q_i x \tag{P1}$$

Proposition 3.2.4. *Assume that the problem (P1) is bounded from below and the solution set S is bounded. Then $x = 0$ is the only global solution of the problem (P1) and both the global (QGC) and global (GSO) hold.*

Proof. The optimal value is 0; otherwise, if there exists an $x \in R^n$ such that $f(x) < 0$, by the homogeneity of order 2 of f, f is unbounded from below.

Next, we have that the solution set of the problem is either an unbounded set or the origin point.

When the solution set is bounded, we claim that $f(x)$ satisfies the quadratic growth condition

$$f(x) \geq \alpha\|x\|^2, \text{ for some positive constant } \alpha.$$

Otherwise, assume there exists a sequence $(x_n)_{n=1}^{\infty}$ such that $f(x_n) < \frac{1}{n}\|x_n\|^2$. Letting $y_n = \frac{x_n}{\|x_n\|}$, and taking a subsequence, if necessary, we have $f(y_n) < \frac{1}{n}$ and $y_n \to y_0$. Then, we have $f(y_0) \le 0$ and $\|y_0\| = 1$, contradicting with the uniqueness of the solution set of the problem.

Also, under the assumption of boundedness of the solution set, the global (GSO) for f still holds. Since $S = \{0\}$, $\pi(x) = 0$ for any $x \in R^n$, $\Omega_\delta(\pi(x)) = \Omega_\delta(0) = \mathscr{S}^m$, and $\mathcal{L}_x(\lambda, \pi(x)) = 0$, $\mathcal{L}_{xx}(\lambda, \pi(x))(h,h) = \max_{\lambda \in \mathcal{T}^m} \sum \lambda_i x^T Q_i x = \max_{1 \le i \le m} x^T Q_i x$,

$$\max_{\lambda \in \Omega_\delta(\pi(x_n))} [\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h,h)] = \max_{1 \le i \le m} x^T Q_i x.$$

From this, we see that the global (QGC) condition is the same as the global (GSO) condition. □

Remark. An alternative result from Yuan [79] is as follows:

$$\max\{x^T Q_1 x, x^t Q_2 x\} \ge 0, \forall x \in R^n \Leftrightarrow \exists \lambda \in [0,1], \text{ s.t. } \lambda Q_1 + (1-\lambda) Q_2 \ge 0.$$

It is not true if there is more than two quadratic forms to be considered. This can be seen from the following example from Martinez and Seeger [46].

Let $f(x) = \max\{x_1^2 + 4x_1x_2 - 3x_2^2, x_1^2 - 8x_1x_2 - 3x_2^2, -5x_1^2 + 4x_1x_2 + 3x_2^2\}$. Then, $f(x) \ge 0$ and for any $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, the minimal eigenvalue of $\lambda_1 Q_1 + \lambda_2 Q_2 + \lambda_3 Q_3$ is no larger than -1. Compared with Proposition 3.2.4, we may say that the global (GSO) is a weaker form of second order condition.

3.3 Reformulation for the constrained programs

Consider the following standard form of nonlinear programming problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \ i = 1, \cdots, k; \ f_j(x) = 0, \ j = k+1, \cdots, m; \end{array} \tag{P}$$

where $f_i : R^n \to R, i = 0, 1, \cdots, m$, are real-valued functions.
First, we introduce the following reduction for problem (P).

Proposition 3.3.1. *Let S be a closed subset of feasible points of the problem (P) such that $f_0(x) = c$ on S. Set*

$$f(x) = \max\{f_0(x) - c, f_1(x), \cdots, f_k(x), |f_{k+1}(x)|, \cdots, |f_m(x)|\}. \tag{3.3.9}$$

Then the following properties are equivalent:
(a) $f_0(x) > c$ *for any feasible points outside of* S*;*
(b) $f(x) > 0$ *for any* $x \in R^n \backslash S$*.*

Proof. The implication of (b)$\Rightarrow$(a) is obvious. Conversely, if (a) holds, then $f(x) \geq f_0(x) - c > 0$ for any feasible x outside of S. On the other hand if x is not feasible, then either $f_i(x) > 0$ for some $i = 1, \cdots, k$ or $|f_i(x)| > 0$ for some $i = k+1, \cdots, m$; in either case $f(x) > 0$. $\square$

Thanks to this proposition we can reformulate the quadratic growth condition and second order sufficient condition for the constrained program (P), using the specific form of the function f given by (3.3.9).

The results of reformulation can be summarized as follows. Consider the set $\Lambda(x)$ of Lagrange multipliers of (P) at x:

$$\Lambda(x) = \{\lambda = (\lambda_0, \cdots, \lambda_m) : \lambda_i \geq 0, i = 1, \cdots, k; \\ \lambda_i f_i(x) = 0, i = 1, \cdots, k; \sum \lambda_i \nabla f_i(x) = 0\}.$$

the set of $\delta-$multipliers:

$$\Lambda_\delta(x) = \{\lambda = (\lambda_0, \cdots, \lambda_m) : \lambda_i \geq 0, i = 1, \cdots, k; \\ \lambda_i f_i(x) = 0, i = 1, \cdots, k; \|\sum \lambda_i \nabla f_i(x)\| \leq \delta\},$$

the subset of normalized multipliers and $\delta-$multipliers:

$$\Lambda^N(x) = \{\lambda \in \Lambda(x); \sum |\lambda_i| \leq 1\},$$

$$\Lambda^N_\delta(x) = \{\lambda \in \Lambda_\delta(x); \sum |\lambda_i| \leq 1\},$$

and the critical cone for (P) at x:

$$K(x) = \{h : \nabla f_i(x) h \leq 0, i = 1, \cdots, k, \nabla f_i(x) h = 0, i = k+1, \cdots, m\}.$$

Now we say that

Definition 3.3.1 (Global (QGC_P))**.** *Problem (P) satisfies the global quadratic growth condition on* S *if* $f(x)$ *defined by (3.3.9) satisfies global (QGC) on* S*.*

Definition 3.3.2 (Global (GSO_P)). *Problem (P) satisfies the general global second order sufficient condition on S if there are regular projection $\pi : R^n \to S$ and an $\alpha > 0$, such that (3.2.8) is valid with $\Omega_\delta(\pi(x))$ replaced by $\Lambda_\delta^N(x)$.*

Similarly, we have the following relation for the global (QGC_P) and global (GSO_P).

Proposition 3.3.2. *Let $f_i(x),\ i = 1, \cdots, m$, be quadratic functions. Then the following implication holds:*

$$\textit{Global } (GSO_P) \Longrightarrow \textit{Global } (QGC_P).$$

Next, we study a special case of quadratic problem, the two-sided constrained quadratic problem.

Proposition 3.3.3. *Consider the following two-sided constrained quadratic problem*

$$\begin{aligned} &\text{minimize } f_0(x) = x^T Ax + 2b^T x + c \\ &\text{subject to} \quad \alpha \le x^T Bx \le \beta, x \in R^n. \end{aligned}$$

Let x^* be a feasible point of the problem and $f_0(x^*) \ge 0$. Assume that $Bx^* = 0$ implies $\alpha < 0 < \beta$. Then the following two assertions are equivalent.
(1) There exists a $\lambda_0 \in R$ such that it satisfies:

$$(A - \lambda_0 B)x^* = b, \tag{3.3.10}$$

$$A - \lambda_0 B > 0, \tag{3.3.11}$$

and

$$\lambda_0(\alpha - (x^*)^T Bx^*) \ge 0 \ge \lambda_0((x^*)^T Bx^* - \beta). \tag{3.3.12}$$

(2) There is some $\gamma > 0$ such that $f(x) = \max\{f_0(x), \alpha - x^T Bx, x^T Bx - \beta\} \ge f_0(x^*) + \gamma\|x - x^*\|^2$, for all $x \in R^n$.
Consequently, we have that x^* is a unique global solution of the two-sided constrained quadratic problem.

Proof. (1) $\Rightarrow$ (2): From (1), we can derive that the global (GSO) holds for $f(x)$. In this case, $\mathcal{L}(\lambda, x) = \lambda_1 f_0(x) + \lambda_2(\alpha - x^T Bx) + \lambda_3(x^T Bx - \beta)$. Under the assumption (1), we may take $\lambda_1 = \frac{1}{1+\lambda_0}, \lambda_2 = \frac{\lambda_0}{1+\lambda_0}, \lambda_3 = 0$ if $\alpha = (x^*)^T Bx^*$, $\lambda_1 = \frac{1}{1-\lambda_0}, \lambda_2 = 0, \lambda_3 = \frac{-\lambda_0}{1-\lambda_0}$ if $\beta = (x^*)^T Bx^*$, or $\lambda_1 = 1, \lambda_2 = 0, \lambda_3 = 0$ if $\alpha < (x^*)^T Bx^* < \beta$. Then

$\mathcal{L}_x(\lambda, \pi(x))h + \frac{1}{2}\mathcal{L}_{xx}(\lambda, \pi(x))(h, h) = \frac{1}{k}[h^T(A - \lambda_0 B)h] \geq \hat{\alpha}\|h\|^2$, for some $\hat{\alpha} > 0$, where $h = x - x^*, k = 1 - \lambda_0$ or $1 + \lambda_0$. And it follows that the global (GSO) holds for the function $f(x)$. Then, by Proposition 3.3.2, the global (QGC) holds for $f(x)$.

(2) $\Rightarrow$ (1): If condition (2) holds, and since $f_0(x^*) \geq 0$, then $f(x) = f_0(x) \geq f_0(x^*) + \gamma\|x - x^*\|^2$, $\forall x \in \{x | \alpha \leq x^T B x \leq \beta\}$. It follows that x^* is the unique global minimizer of the problem $\min\{f_0(x) - \gamma\|x - x^*\|^2 : \alpha \leq x^T B x \leq \beta\}$ and also is the global minimizer of the original problem. Applying the result of Stern and Wolkowicz, we have that there is a $\lambda_0 \in R$ such that $(A - \lambda_0 B)x^* = b, A - \lambda_0 B \geq \gamma I > 0$, and $\lambda_0(\alpha - (x^*)^T B x^*) \geq 0 \geq \lambda_0((x^*)^T B x^* - \beta)$. $\square$

Chapter 4

Fractional program

4.1 Introduction

One key technique frequently used in optimization community is S-lemma or S-procedure which originally arose from the stability analysis of nonlinear systems, see, for example, [74] and the references therein. The S-procedure concerns the problem when a quadratic (in)equality is a consequence of other quadratic (in)equalities, such as

$$f_0(x) \geq 0 \text{ for } f_1(x) \geq 0, \cdots, f_m(x) \geq 0, \forall x \in X.$$

To ease the complexity of the the problem, the S-procedure considers the solvability of the following auxiliary function

$$S(x) = f_0(x) - \lambda_1 f_1(x) - \cdots - \lambda_m f_m(x) \geq 0, \lambda_1, \cdots, \lambda_m \geq 0.$$

When the functions $f_i(x), i = 1, \cdots, m$ are linear or convex, the results of this kind refer to the Farkas lemma or Farkas theorem. For more details for S-lemma, we may refer to the article by Polyak [59]and the review by Pólik and Terlaky [58], and the references therein.

A known variant of S-lemma states that, for quadratic functions $f_i : R^n \to R, i = 1, 2$, the following two assertions are equivalent:

$$\begin{aligned} &(1)\ \max\{f_1(x), f_2(x)\} \geq 0, \forall x \in R^n; \\ &(2)\ \exists \lambda_1, \lambda_2 \geq 0 \text{ s.t. } \lambda_1 f_1(x) + \lambda_2 f_2(x) \geq 0, \forall x \in R^n. \end{aligned}$$

In this chapter, we will mainly concentrate on the quadratic fractional programming problems. First, we will generalize the S-lemma to the quadratic fractional function case. Then by applying S-lemma and previously presented Frank-wolfe theorem we will obtain existence results for various kinds of the quadratic fractional programs.

4.2 Quadratic Fractional Programs

Consider the following quadratical fractional program

$$f(x) = \max\{f_1(x), f_2(x)\},\ \forall x \in H, \tag{P}$$

where $f_i(x) = \frac{x^T A_i x}{2(a^T x+s)},\ x \in R^n,\ \ i = 1, 2,\ H = \{x \in R^n | a^T x + s > 0\}, A_i, i = 1, 2,$ are $n \times n$ matrices, $a \in R^n$ and s is a real.

We will characterize the conditions such that $f(x) \geq 0$. Note that the problem $f(x) \geq 0,\ \forall x \in H$ is equivalent to that $\max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in H$. We have the following result similar to the Yuan's alternative theorem.

Proposition 4.2.1. *The following two assertions are equivalent:*
(i) $\max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in H$;
(ii) $\exists t \in [0, 1],$ s.t. $tA_1 + (1 - t)A_2 \geq 0$.

Proof. First, we have $\max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in H \Leftrightarrow \max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in R^n$.

This follows from the two cases below:
(1) If $0 \in H$, then there exists a ball $B(0, r) \subset H$. For any $y \in R^n$, $\exists k > 0$ s.t. $ky \in B(0, r) \subset H$.
(2) Assume $0 \notin H$. We have that $\max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in H \Leftrightarrow \max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in \{x|a^T x > 0\} \Leftrightarrow \max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in \{x|a^T x \geq 0\} \Leftrightarrow \max\{x^T A_1 x, x^T A_2 x\} \geq 0,\ \forall x \in R^n$. For the equivalences, we used, respectively, the facts that $\{x|a^T x > 0\}$ is a cone generated by H, the continuity of the max-function, and the symmetry of A_i.

Second, by the Yuan's alternative theorem, the equivalence is obtained. □

Proposition 4.2.2. *Assume that $s \neq 0$. Then the following three assertions are equivalent:*
(i) $f(x) = \max\{f_1(x), f_2(x)\} \geq 0,\ \forall x \in H$;

(ii) For any $x \in H$, $\exists t \in [0,1]$, s.t. $t \bigtriangledown^2 f_1(x) + (1-t) \bigtriangledown^2 f_2(x) \geq 0$;

(iii) $\exists t \in [0,1]$, s.t. $tA_1 + (1-t)A_2 \geq 0$.

Proof. That (i)$\Leftrightarrow$ (iii) follows from Proposition 4.2.1.

(ii)$\Leftrightarrow$ (iii): First, we do some calculus. The gradients and Hessians of f_i, respectively, are

$$\bigtriangledown f_i(x) = \frac{2(a^T x + s)A_i x - (x^T A_i x)a}{2(a^T x + s)^2}, \tag{4.2.1}$$

and

$$\bigtriangledown^2 f_i(x) = \frac{(a^T x + s)^2 A_i - (a^T x + s)((A_i x)a^T + a(A_i x)^T) + (x^T A_i x)aa^T}{(a^T x + s)^3}, i = 1, 2. \tag{4.2.2}$$

So,

$$y^T \bigtriangledown^2 f_i(x) y = \frac{((a^T x + s)y - (a^T y)x)^T A_i((a^T x + s)y - (a^T y)x)}{(a^T x + s)^3}, i = 1, 2, \tag{4.2.3}$$

and

$$\begin{aligned} y^T(t \bigtriangledown^2 f_1(x) &+ (1-t) \bigtriangledown^2 f_2(x))y \\ &= \frac{((a^T x + s)y - (a^T y)x)^T (tA_1 + (1-t)A_2)((a^T x + s)y - (a^T y)x)}{(a^T x + s)^3}. \end{aligned}$$

Hence, (iii) implies (ii).

For the converse, the semi-positiveness of $tA_1 + (1-t)A_2$ for some $t \in [0,1]$ follows from the fact that there exists an $x \in H$ such that the transformation $y \in R^n \mapsto (a^T x + s)y - (a^T y)x \in R^n$ is surjective, or equivalently, there is an $x \in H$ such that the matrix $(a^T x + s)I - xa^T$ is nonsingular.

If $0 \in H$, taking $x = 0$, then the transformation is surjective. Now assume that $0 \notin H$. Obviously that $a \neq 0$. Without loss of generality, assume that, the first component of a, $a_1 \neq 0$. Let $x = (ta_1, 0, \cdots, 0)$, for some $\lambda > 0$ large enough such that $a^T x + s > 0$. And therefore, $\det((a^T x + s)I - xa^T) = s(\lambda a_1^2 + s)^{n-1} \neq 0$. $\square$

Remark. For the case $f_i = \frac{x^T A_i x + 2b_i^T x + c_i}{a^T x + s}$ with $s \neq 0$, we will have a similar result. Let $y^T = (t, x^T)$, and $B_i = \begin{pmatrix} c_i & b_i^T \\ b_i & A_i \end{pmatrix}$, $i = 1, 2$. So, the program can be reformulated as

$$\min_{y \in H'} g(y) = \max\{g_1(y), g_2(y)\},$$

where $g_i(y) = \frac{y^T B_i y}{(0,a^T)y+s}$, $y \in R^{n+1}$, $i = 1, 2$, and $H' = \{y \in R^{n+1} | (0, a^T)y + s > 0\} = \{(t, x) | t \in R, x \in H\}$.

So, we have $g(y) \geq 0, \forall x \in H'$ if and only if $\exists t \in [0, 1]$, s.t. $tB_1 + (1-t)B_2 \geq 0$ if and only if, for any $x \in H'$, $\exists t \in [0, 1]$, s.t. $t \bigtriangledown^2 g_1(x) + (1-t) \bigtriangledown^2 g_2(x) \geq 0$.

Another form of the above conclusion is given as follows.

Proposition 4.2.3. *Assume that there is an $\bar{x} \in H$ such that $\bar{x}^T A_2 \bar{x} < 0$. Then the two assertions below are equivalent.*
(i) There is no $x \in H$ such that the following system holds:

$$\frac{x^T A_1 x}{a^T x+s} < 0, \frac{x^T A_2 x}{a^T x+s} \leq 0,$$

(ii) There exists a $t \geq 0$ such that $A_1 + tA_2$ is positive semi-definite;
If $s \neq 0$, the above two assertions also equivalent to
(iii) For any $x \in H$, $\exists t \geq 0$, s.t. $\bigtriangledown^2 f_1(x) + t \bigtriangledown^2 f_2(x) \geq 0$.

As a direct application of Proposition 4.2.2, we have

Proposition 4.2.4. *Consider the following quadratic fractional program*

$$\begin{array}{ll} \min_{x \in H} & \frac{x^T A_1 x}{a^T x+s} \\ \text{subject to} & \frac{x^T A_2 x}{a^T x+s} \leq 0, \end{array} \qquad \text{(QFP)}$$

where $H = \{x \in R^n | a^T x + s > 0\}$. Assume that $s > 0$. Then $x = 0$ is the global solution of the problem (QFP) if and only if there exists a $t \in [0, 1]$ such that $tA_1 + (1-t)A_2$ is positive semi-definite.

Next, we give a sufficient condition for a specific problem of the quadratic fractional programs.

Proposition 4.2.5. *Consider the following quadratic fractional program with a two-sided quadratic fractional constraint:*

$$\begin{array}{ll} \min_{x \in H} & f_1(x) = \frac{x^T A_1 x}{a^T x+s} \\ \text{subject to} & \gamma \leq f_2(x) = \frac{x^T A_2 x}{a^T x+s} \leq \beta, \end{array}$$

where $H = \{x \in R^n | a^T x + s > 0\}$ and $s \neq 0$. If there are a $y \in H$ and a $\lambda \in R$ such that the following conditions hold:

$$2(a^T y + s)((A_1 - \lambda A_2)y) - (y(A_1 - \lambda A_2)y)a = 0, \tag{4.2.4}$$

$$A_1 - \lambda A_2 \geq 0, \tag{4.2.5}$$

and

$$\lambda(\gamma - f_2(y)) \geq 0 \geq \lambda(f_2(y) - \beta). \tag{4.2.6}$$

Then y is the global solution of the problem.

Proof. There are three cases to consider.
Case (i). Suppose that $\gamma < f_2(y) < \beta$. Then $\lambda = 0$ from the condition (4.2.6) and $A_1 \geq 0$. Since $s \neq 0$, by (4.2.3), $\bigtriangledown^2 f_1(x) \geq 0$ on H, i.e. f_1 is convex on the feasible region. Then, y solves the problem.
Case (ii). Suppose that $\gamma = f_2(y)$. By (4.2.4), (4.2.5), y minimizes the Lagrangian function

$$L(x, \lambda) = f_1(x) + \lambda(\gamma - f_2(x))$$

over R^n. That is

$$f_1(y) = L(y, \lambda) \leq L(x, \lambda), \forall x \in R^n.$$

Since condition (4.2.6) implies $\lambda \geq 0$, it follows that $\lambda(\gamma - f_2(x)) \leq 0$, for all feasible x. Then, $f_1(y) \leq f_1(x)$, for all feasible x.
Case (iii). Suppose that $f_2(y) = \beta$. This is similar to case (ii). □

We now consider the existence of a global solution of a quadratic fractional problem over a polyhedra:

$$\min_{x \in P} \frac{x^T A_1 x + b_1^T x + c_1}{x^T A_2 x + b_2^T x + c_2}$$

where P is a polyhedron.

Proposition 4.2.6. *Assume that the above problem is bounded from below and $x^T A_2 x + b_2^T x + c_2$ is positive and bounded over P. Then there exists a global solution x_0 in P.*

Proof. Let $\inf_{x \in P} \frac{x^T A_1 x + b_1^T x + c_1}{x^T A_2 x + b_2^T x + c_2} = \alpha$. Then, for any $x \in P$,

$$x^T A_1 x + b_1^T x + c_1 \geq \alpha(x^T A_2 x + b_2^T x + c_2).$$

That is,

$$x^T A_1 x + b_1^T x + c_1 - \alpha(x^T A_2 x + b_2^T x + c_2) \geq 0.$$

Note that

$$\inf_{x \in P} x^T A_1 x + b_1^T x + c_1 - \alpha(x^T A_2 x + b_2^T x + c_2) = 0.$$

By Frank-Wolfe Theorem, we know that there is an $x_0 \in P$ such that

$$x_0^T A_1 x_0 + b_1^T x_0 + c_1 - \alpha(x_0^T A_2 x_0 + b_2^T x_0 + c_2) = 0$$

and

$$x_0^T A_2 x_0 + b_2^T x_0 + c_2 > 0.$$

Therefore, x_0 is a global solution of the original problem. □

Remark. If the denominator is not bounded over the polyhedron, the assertion may not hold. For example, the problem $\min\{\frac{1}{x} : x \geq 1\}$ has no solutions.

Next, we extend the Proposition 4.2.2 to the multiple quadratic fractional case. Let $f_i = \frac{x^T A_i x}{2(a^T x + s)}, i = 1, \cdots, m$, with $m \geq 3$. Consider the following assertions:
(i) $\max_{1 \leq i \leq m}\{\frac{x^T A_i x}{2(a^T x + s)}\} \geq 0$, for all $x \in H = \{x \in R^n | a^T x + s > 0\}$;
(ii) There exist $t_i \geq 0$, $i = 1 \cdots, m$, with $t_1 + \cdots + t_m = 1$, such that $t_1 A_1 + \cdots + t_m A_m$ is positive semi-definite;
(iii) There exist $t_i \geq 0$, $i = 1 \cdots, m$, with $t_1 + \cdots + t_m = 1$, such that $t_1 \bigtriangledown^2 f_1(x) + \cdots + t_m \bigtriangledown^2 f_m(x)$ is positive semi-definite for all $x \in H$.

Lemma 4.2.1. *(Martinez-Legaz and Seeger) Let the symmetric matrices $A_1, \cdots, A_m$, have non-positive extradiagonal terms. Then the following two statements are equivalent:*
(1) $\max_{1 \leq i \leq m}\{x^T A_i x\} \geq 0, \forall x \in R^n$;
(2) There exist $t_i \geq 0$, with $t_1 + \cdots + t_m = 1$, such that $t_1 A_1 + \cdots + t_m A_m$ is positive semi-definite.

Applying this Lemma, we may obtain:

Proposition 4.2.7. *Let $s \neq 0$. Suppose that there exists a nonsingular $n \times n$ matrix Q such that $Q^T A_1 Q, \cdots, Q^T A_m Q$ have non-positive extradiagonal terms. Then the above statements (i),(ii), and (iii) are equivalent.*

Proof. First, similar to the case of $m = 2$, we have that $\max_{1\le i\le m}\{x^T A_i x\} \ge 0, \forall x \in H \Leftrightarrow \max_{1\le i\le m}\{x^T A_i x\} \ge 0, \forall x \in R^n$. Then, the equivalence between (i) and (ii) follows from Lemma 4.2.1.
We can easily have

$$
\begin{aligned}
& y^T(\sum_{i=1}^{m} t_i \bigtriangledown^2 f_i(x))y \\
& \qquad\qquad = \frac{((a^T x + s)y - (a^T y)x)^T(\sum_{i=1}^{m} t_i \bigtriangledown^2 f_i(x))((a^T x + s)y - (a^T y)x)}{(a^T x + s)^3}.
\end{aligned}
$$

The equivalence between (ii) and (iii) follows from the proof of Proposition 4.2.2. $\square$

Remark. If the matrices $A_1, \cdots, A_m$ commute pairwise, then they can be simultaneously diagonalized under an orthogonal transformation. Hence the pairwise commuting property is sufficient to the non-positive extradiagonal property.

Chapter 5

Conclusions

In this thesis, we mainly concentrated on some global analysis for the quadratically constrained quadratic programming problems. We studied the existence of global solutions, the general global second-order sufficient condition, and the global quadratic growth condition for the quadratic problems.

In chapter 2, as a complement to the existing theory of the existence results for the quadratic problems, we only provided two new proofs for two assertions: the continuity of a sequence of nested sets defined by a number of convex quadratic inequalities; the attainability of the quadratic problem with all functions being convex.

In chapter 3, we studied the global second-order conditions for quadratic problems with non-isolated global solutions. First, we defined the global quadratic growth condition and the general global second-order sufficient condition for the function of the form of maximizing a finite number quadratic functions. Then, we investigated the relations between the two conditions. It was shown that the global quadratic growth condition is implied by the global second-order sufficient condition and that, when the solution set is a singleton and the number of quadratic term is 2, the reverse implication also holds. Finally, after a reformulation, we applied these results to the quadratically constrained quadratic problems.

In chapter 4, we established some alternative results for quadratic fractional functions. Then, we studied some corresponding quadratic fractional problems by applying these alternative results.

Bibliography

[1] W. Alt, *Lipschitzian Perturbation of infinite optimization problems*, Mathematical programming with Data Perturbations (A. V. Fiacco, ed.), Marcel Dekker, New York, 1983, pp. 7–21.

[2] ______, *Stability of Solutions for a Class of Nonlinear Cone Constrained Optimization Problems Part I: Basic Theory*, Numer. Funct. Anal. Optim. **10** (1989), 1053–1064.

[3] A. Auslender, *Stability in mathematical programming with non-differentiable data*, SIAM J. Control Optim. **22** (1984), 239–254.

[4] ______, *Noncoercive Optimization Problems*, Mathematics of Operations Research **21** (1996), 769–782.

[5] ______, *How to Deal With the Unbounded in Optimization: Theory and Algorithms*, Mathematical Programming **79** (1997), 3–18.

[6] ______, *Existence of Optimal Solutions and Duality Results Under Weak Conditions*, Mathematical Programming **88** (2000), 45–59.

[7] A. Auslender and M.Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer Monographs in Mathematics, Springer, 2003.

[8] E.G. Belousov, *Introduction to Convex Analysis and Integer Optimization*, University Publisher, Moscow, 1977 (Russian).

[9] Evgeny G. Belousov and D. Klatte, *A Frank-Wolfe Type Theorem for Convex Polynomial Programs*, Computational Optimization and Applications **22** (2002), 37–48.

[10] A. Ben-Tal and J. Zowe, *An Unified Theory of First and Second Order Conditions for Extremum Problems in Topological Vector Spaces*, Math. Progr. Study **19** (1982), 39–77.

[11] D. Bertsekas, A. Nedic, and Asuman E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.

[12] Dimitri P. Bertsekas and P. Tseng, *Set Intersection Theorems and Existence of Optimal Solutions*, Mathematical Programming **110** (2007), 287–314.

[13] E. Blum and W. Oettli, *Direct Proof of the Existence Theorem for Quadratic Programming*, Operations Research **20** (1972), 165–167.

[14] J. F. Bonnans, *Directional Derivatives of Optimal Solutions in Smooth Nonlinear Programming*, J. Optim. Theory Appl. **73** (1992), 27–45.

[15] J. F. Bonnans and A. D. Ioffe, *Quadratic Qrowth and Stability in Convex Programming Problems with Multiple Solutions*, J. Convex Analysis **2** (1995), 41–57.

[16] ———, *Second-order Suffciency and Quadratic Growth for Non-isolated Minima*, Mathematics of Operations Research **20** (1995), 801–817.

[17] J. F. Bonnans and A. Shapiro, *Optimization Problems With Perturbations: A Guided Tour*, SIAM Review **40** (1998), 228–264.

[18] ———, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[19] J. V. Burke and M. C. Ferris, *Weak Sharp Minima In Mathematical Programming*, SIAM Journal on Control and Optimization **31** (1993), 1340–1359.

[20] C. Baiocchi and G. Buttazzo and F. Gastaldi and F. Tomarelli, *General Existence Theorems for Unilateral Problems in Continuum Mechanics*, Archive for Rational Mechanics and Analysis **100** (1987), 149–189.

[21] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Trust-Region Methods*, MPS/SIAM Ser. Optim. 1, SIAM, Philadelphia.

[22] R. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, 1992.

[23] B. C. Eaves, *On Quadratic Programming*, Management Science **17** (1971), 698–711.

[24] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Progmmming*, Academic Press, New York, 1983.

[25] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.

[26] C. A. Floudas and V. Visweswaran, *Quadratic Optimization*, Handbook of global optimization (R. Horn and P.M. Pardalos, eds.), Kluwer Academic Publishers: The Netherlands, 1995, pp. 217–269.

[27] M. Frank and P. Wolfe, *An Algorithm for Guadratic Programming*, Naval Research Logistics Quarterly **3** (1956), 95–110.

[28] W. Gander, G.H. Golub, and U. von Matt, *A Constrained Eigenvalue Problem*, Linear Algebra Appl. **114-115** (1989), 815–839.

[29] J. Gauvin and R. Janin, *Directional Behaviour of Optimal Solutions in Nonlinear Mathematical Programming*, Mathematics of Operations Research **13** (1988), 629–649.

[30] David M. Gay, *Computing Optimal Locally Constrained Steps*, SIAM J. Sci. and Stat. Comput. **2** (1981), 186–197.

[31] B. Gollan, *Perturbation Theory for Abstract Optimization Problems*, J. Optimization Theory Appl. **35** (1981), 417–441.

[32] ———, *On the Marginal Function in Nonlinear Progmmming*, Math. Oper. Res. **9** (1984), 208–221.

[33] G. H. Golub and U. von Matt, *Quadratically Constrained Least Squares and Quadratic Problems*, Numerische Mathematik **59** (1991), 561–580.

[34] J.-B. Hiriart-Urruty, *Conditions for Global Optimality in: Handbook for Global Optimization*, pp. 1–26, Kluwer Academic Publishers, Dordrecht, 1995.

[35] ———, *Conditions for Global Optimality 2*, J. Global Optim. **13** (1998), 349–367.

[36] ———, *Global optimality conditions in maximizing a convex quadratic function under convex quadratic constraints*, J. Global Optim. **21** (2001), 445–455.

[37] R. Horst and H. Tuy, *Global Optimization*, Springer, Berlin, Springer-Verlag, 1990.

[38] A. Ioffe, *Necessary and Sufficient Conditions for a Local Minimum*, SIAM J. Control. Optim. **17** (1979), 245–288.

[39] A.D. Ioffe, *On Sensitivity Analysis of Nonlinear Programs in Banach Spaces : the Approach via Composite Unconstrained Optimization*, SIAM J. Optimization **4** (1994), 1–43.

[40] K. Jittorntrum, *Solution Point Differentiability Without Strict Complementarity in Nonlinear Progmmming*, Math. Programming **21** (1984), 127–138.

[41] N. Karmarkar, *An Interior-Point Approach to 'NP-Complete Problems*, Comtemporary Mathematics **114** (1990), 351–366.

[42] H. W. Kuhn and A. W. Tucker, *Nonlinear Programming*, Proceedings of the Second Berkeley Symposium on mathematical Statistics and Probability (1951), 481–492.

[43] G. M. Lee, N. N. Tam, and N. D. Yen, *Quadratic Programming and Affine Variational Inequalities: A Qualitative Study*, Nonconvex optimization and its applications, Springer, 2005.

[44] Z. Q. Luo and S. Zhang, *On the Extension of Frank-Wolfe Theorem*, Computational Optimization and Applications **13** (1999), 87–110.

[45] J.M. Martinez, *Local Minimizers of Quadratic Functions on Euclidean Balls and Spheres*, SIAM J. Optimization **4** (1994), 159–176.

[46] J. E. Martinez-Legaz and A. Seeger, *Yuan's Alternative Theorem and the Maximization of the Minimum Eigenvalue Function*, Journal of Optimization Theory and Applications **82** (1994), 159–167.

[47] T. Matsui, *NP-Hardness of Linear Multiplicative Programming and Related Problems*, Journal of Global Optimization **9** (1996), 113–119.

[48] H. Maurer and J. Zowe, *First and Second-Order Necessary and Suficient Optimality Conditions for Infinite-dimensional Programming Problems*, Mathathematical Programming **16** (1979), 98–110.

[49] J. J. Moré, *Generalizations of the Trust Region Problem*, Optimization Methods and Software **2** (1993), 189–209.

[50] J. J. Moré and D. C. Sorensen, *Computing a Trust Region Step*, SIAM J. Sci. and Stat. Comput. **4** (1983), 553–572.

[51] K. G. Murty and S. N. Kabadi, *Some NP-Complete Problems in Quadratic and Nonlinear Programming*, Mathematical Programming **39** (1987), 117–129.

[52] W. T. Obuchowska, *On Generalizations of the Frank-Wolfe Theorem to Convex and Quasi-Convex Programmes*, Computational Optimization and Applications **33** (2006), 349–364.

[53] A. Ozdaglar and P. Tseng, *Existence of Global Minima for Constrained Optimization*, Journal of Optimization Theory and Applications **128** (2006), 523–546.

[54] Panos M. Pardalos and Stephen A. Vavasis, *Quadratic Programming with One Negative Eigenvalue is NP-Hard*, Journal of Global Optimization **1** (1991), 15–22.

[55] 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.

[56] Andre F. Perold, *A Generalization of the Frank-Wolfe Theorem*, Mathematical Programming **18** (1980), 215–227.

[57] Elijah Polak, *Optimization : Algorithms and Consistent Approximations*, Springer-Verlag, New York, 1997.

[58] I. Pólik and T. Terlaky, *A Survey of the S-Lemma*, Siam Review **49** (2007), 371–418.

[59] B. T. Polyak, *Convexity of Quadratic Transformations and its Use in Control and Optimization*, J. Optim. Theory Appl. **99** (1998), 553–583.

[60] F. Rendl and H. Wolkowicz, *A Semidefinite Framework for Trust Region Subproblems with Applications to Large Scale Minimization*, Mathematical Programming **77** (1997), 273–299.

[61] S. M. Robinson, *Perturbed kuhn-tucker points and rates of convergence for a class of nonlinear-programming algorithms*, Mathematical Programming **7** (1974), 1–16.

[62] ———, *Generalized Equations and Their Solutions, Part II: Applications to Nonlinear Programming*, Math. Programming Stud. **19** (1982), 200–221.

[63] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[64] ———, *Second order optimality conditions in nonlinear programming obtained by way of epi-derivatives*, Mathematics of Operations Research **14** (1989), 462–484.

[65] S. Sahni, *Computationally Related Problems*, SIAM J. Comput. **3** (1974), 262–279.

[66] A. Shapiro, *Perturbation Theory of Nonlinear Programs When the Set of Optimal Solutions is Not a Singleton*, Appl. Math. Optim. **18** (1988), 215–229.

[67] ———, *Perturbation Analysis of Optimization Problems in Banach Spaces*, Numer. Funct. Anal. Optim. **13** (1992), 97–116.

[68] D. C. Sorensen, *Newton's Method with a Model Trust Region Modification*, SIAM J. Numer. Anal. **19** (1982), 409–426.

[69] R. J. Stern and H. Wolkowicz, *Indefinite Trust Region Subproblems and Nonsymmetric Eigenvalue Perturbations*, SIAM J. Optim. **5** (1995), 286–313.

[70] M. Studniarski and Doug E. Ward, *Weak Sharp Minima: Characterizations and Sufficient Conditions*, SIAM Journal on Control and Optimization **38** (1999), 219–236.

[71] S. A. Vavasis, *Quadratic Programming is in NP*, Information Processing Letters **36** (1990), 73–77.

[72] D. E. Ward, *Characterizations of Strict Local Minima and Necessary Conditions for Weak Sharp Minima*, Journal of Optimization Theory and Applications **80** (1994), 551–571.

[73] ———, *Sufficient Conditions for Weak Sharp Minima of Order Two and Directional Derivatives of the Value Function*, Mathematical Programming with Data Perturbations, Lect. Notes Pure Appl. Math.,195 (New York) (A.V. Fiacco, ed.), Dekker, 1998, pp. 419–436.

[74] V. A. Yakubovich, *S-Procedure in Nonlinear Control Theory*, Vestnik Leningrad. Univ. **1** (1971), 62–77, in Russian.

[75] X. Q. Yang, *Second-Order Global Optimality Conditions for Convex Composite Optimization*, Mathematical Programming **81** (1998), 327–347.

[76] ———, *Second-Order Global Optimality Conditions for Optimization Problems*, Journal of Global Optimization **30** (2004), 271–284.

[77] Y. Y. Ye, *Interior-Point Algorithms for Global Optimization*, Annals of Operations Research **25** (1990), 59–73.

[78] ———, *On Affine Scaling Algorithms for Nonconvex Quadratic Programming*, Mathematical Programming **56** (1992), 285–300.

[79] Y. Yuan, *On a Subproblem of Trust Region Algorithms for Constrained Optimization*, Mathematical Programming **47** (1990), 53–63.