October 1997

Econometric Institute Report No. 9748/A

ON THE EXTENSIONS OF FRANK-WOLFE THEOREM[1]

by

Zhi-Quan Luo[2] and Shuzhong Zhang[3]

ABSTRACT

In this paper we consider optimization problems defined by a quadratic objective function and a finite number of quadratic inequality constraints. Given that the objective function is bounded over the feasible set, we present a comprehensive study of the conditions under which the optimal solution set is nonempty, thus extending the so-called Frank-Wolfe theorem. In particular, we first prove a general continuity result for the solution set defined by a system of convex quadratic inequalities. This result implies immediately that the optimal solution set of the aforementioned problem is nonempty when all the quadratic functions involved are convex. In the absence of the convexity of the objective function, we give examples showing that the optimal solution set may be empty either when there are two or more convex quadratic constraints, or when the Hessian of the objective function has two or more negative eigenvalues. In the case when there exists only one convex quadratic inequality constraint (together with other linear constraints), or when the constraint functions are all convex quadratic and the objective function is quasi-convex (thus allowing one negative eigenvalue in its Hessian matrix), we prove that the optimal solution set is nonempty.

KEY WORDS: Convex quadratic system, existence of optimal solutions, quadratically constrained quadratic programming.

---

[1] The research of the first author is supported by the Natural Sciences and Engineering Research Council of Canada, Grant No. OPG0090391.

[2] Room 225/CRL, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, L8S 4L7, Canada.

[3] Econometric Institute, Erasmus University Rotterdam, The Netherlands.

1 Introduction

In this paper we are concerned with the question whether an optimization problem has an optimal solution or not. The answer to this question can be trivial if, for instance, all the feasible solutions form a compact set, or if the objective function is unbounded over the feasible region. In the former case, the well-known Weierstrass theorem applies and so the answer is positive. In the latter case, clearly no optimal solution can exist, and the problem is called *unbounded* by convention. The situation becomes more subtle if the set of feasible solutions is not bounded in norm, and yet the objective function value is bounded for all the feasible solutions. In the setting of primal-dual convex conic programs, a detailed discussion on issues related to this and others can be found in Luo, Sturm and Zhang [9].

In mathematical terms we consider the following constrained optimization problem:

$$\text{(P)} \quad \begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \;\; i = 1, 2, ..., m \end{array}$$

where all f_i's are continuous differentiable functions, $i = 0, 1, ..., m$. In case these functions are all affine linear, the problem is called *linear programming*. It is well-known that for linear programming, a bounded feasible problem always has an optimal solution. This property is remarkable, and fails to hold for general nonlinear programs. That is why it is also known as the *fundamental theorem of linear programming* (see Chvátal [4]). Frank and Wolfe [6] showed that if f_i's remain affine linear for $i = 1, 2, ..., m$ and f_0 is an arbitrary quadratic function, then (P) being feasible and bounded from below over the feasible region implies that an optimal solution exists. This result is known as the Frank-Wolfe theorem and can be considered as a generalization of the fundamental theorem of linear programming. Several alternative proofs for the Frank-Wolfe theorem were proposed; see [3, 5]. Perold [13] further generalized the Frank-Wolfe theorem to a class of non-quadratic objective functions (but the constraints are still affine linear).

In this paper we consider generalizations of the Frank-Wolfe theorem as well. However, we will restrict ourselves to the case where all the functions involved are either affine linear or quadratic. We also assume the problem under consideration always has a non-empty feasible set.

As a first and important step we consider the solution set defined by convex quadratic inequalities; that is, we consider the feasible set of (P) where

$$\begin{aligned} f_1(x) &:= \tfrac{1}{2} x^T Q_1 x + q_1^T x + c_1 \\ f_2(x) &:= \tfrac{1}{2} x^T Q_2 x + q_2^T x + c_2 \\ &\vdots \\ f_m(x) &:= \tfrac{1}{2} x^T Q_m x + q_m^T x + c_m \end{aligned}$$

with each $Q_i \in \Re^{n \times n}$ being positive semidefinite, $q_i \in \Re^n$, and $c_i \in \Re$. More specifically, for each $\epsilon = (\epsilon_1, \epsilon_2, ..., \epsilon_m)^T \ge 0$, we let $X(\epsilon)$ denote the feasible set of the following perturbed quadratic system:

$$X(\epsilon) := \{x \in \Re^n \, : \, f_i(x) \le \epsilon_i, \;\; i = 1, 2, ..., m\}. \tag{1.1}$$

In that case we prove that the following implication holds:

$$X(\epsilon^k) \neq \emptyset \quad \text{for some sequence } \epsilon^k \downarrow 0 \quad \Rightarrow \quad X(0) \neq \emptyset. \tag{1.2}$$

This result was first established in an unpublished report by Luo [8] and is documented here for ease of future reference. As a corollary of this result, we readily establish the attainability of the infimum (assumed to be finite) of (P) when f_0, f_1, ..., f_m are all convex. Since convex quadratically constrained quadratic programming can be viewed as a special case of the so-called l_p programming, this attainability result also follows from the strong duality relation for l_p programming established in Terlaky [15].

Starting from the above attainability result for Convex Quadratically Constrained Quadratic Programming (CQCQP), which can be called the *fundamental theorem for CQCQP* as parallel to the LP case, we seek possibilities of removing some of the convexity assumptions either in the objective function or in the constraints. Our findings are summarized below:

(1) Even if f_0 is convex, and at least one of the constraint functions f_i $(i = 1, 2, ..., m)$ is nonlinear $(Q_i \neq 0)$ and non-convex, the infimum of (P) (assumed to be finite) is in general <u>not</u> attainable.

(2) If f_0 is non-convex and at least two or more functions f_i $(i = 1, 2, ..., m)$ are nonlinear (but convex), then the infimum of (P) is in general <u>not</u> attainable.

(3) If f_0 is nonconvex and at most one of the constraint functions f_i $(i = 1, 2, ..., m)$ is nonlinear (but convex), then the infimum of (P) (assumed to be finite) is always attained.

(4) If f_0 is quasi-convex over the feasible region and all of the constraint functions f_i $(i = 1, 2, ..., m)$ are convex, then the infimum of (P) (assumed to be finite) is always attained.

To put the above results in perspective, our results (3)–(4) can be viewed as natural extensions of the Frank-Wolfe theorem. Secondly, the continuity of the feasible set defined by convex quadratic inequalities is an extension of a similar result for the polyhedral set. The latter is a direct consequence of Hoffman's well known error bound [7] for linear inequality systems and also of Robinson's theorem [14] on the upper Lipschitzian continuity of polyhedral multi-functions.

The organization of the paper is as follows. Next section is devoted to the discussion on the feasible set defined by convex quadratic inequalities. In Section 3 we extend the Frank-Wolfe theorem to the case where at most one constraint function is nonlinear and convex (while the other constraints are all linear). In Section 4 the fundamental theorem of CQCQP is extended to the case where the objective function is quasi-convex. Finally we conclude the paper in Section 5.

Our notations are standard. For example, the vector notation $x \geq 0$ means that each component of x is nonnegative. The superscript "T" indicates either vector or matrix transpose. Also, for any square matrix Q, the notation $Q \geq 0$ indicates Q positive semidefinite. In addition, for any optimization problem (OP), we use inf((OP)) to denote the infimum of (OP). The notation $\|\cdot\|$ denotes the usual Euclidean norm. Finally, we write $\epsilon^k \downarrow 0$ when the sequence ϵ^k approaches to zero monotonically from above.

2 Convex quadratic inequality systems

We start our discussion by first investigating the continuity of the solution set defined by convex quadratic inequalities. The proof of this result first appeared in the unpublished report [8] by the first author, and is included here for completeness.

Theorem 1 *Suppose that the perturbed solution set* $X(\epsilon)$ *(as defined by (1.1)) is nonempty for some positive sequence* $\{\epsilon^k\}$ *approaching to zero. Then the unperturbed feasible set* $X(0)$ *is also nonempty.*

Proof: We prove by induction on m, the number of quadratic inequalities. Let $m = 1$. Suppose $f_1(x) \le \epsilon_1^k$ has a solution x^k, where $\epsilon_1^k \downarrow 0$. In other words,

$$\frac{1}{2}(x^k)^T Q_1 x^k + q_1^T x^k + c_1 \le \epsilon_1^k, \quad \forall k. \tag{2.1}$$

If $\{x^k\}$ has a bounded subsequence, then the cluster point of this subsequence must be a solution of $f_1(x) \le 0$. Otherwise, we have $\|x^k\| \to \infty$. In this case, we divide (2.1) by $\|x^k\|^2$, let $k \to \infty$ and use $Q_1 \ge 0$ to obtain

$$0 \le \limsup_{k\to\infty} \frac{(x^k)^T Q_1 x^k}{\|x^k\|^2} \le 0, \quad \text{and} \quad \limsup_{k\to\infty} \frac{q_1^T x^k}{\|x^k\|} \le 0.$$

Since $Q_1 \ge 0$, we further deduce

$$\lim_{k\to\infty} \frac{Q_1 x^k}{\|x^k\|} = 0.$$

By passing to a subsequence if necessary, we assume

$$u := \lim_{k\to\infty} \frac{x^k}{\|x^k\|}.$$

We consider two cases.

Case 1.1. $q_1^T u < 0$. Since $Q_1 u = 0$, we see u is a recession direction for $f_1(x) \le 0$. For $t > 0$, we consider

$$\begin{aligned} f_1(tu) &= \frac{t^2}{2} u^T Q_1 u + t q_1^T u + c_1 \\ &= c_1 + t q_1^T u, \end{aligned}$$

which is nonpositive with

$$t = \left| \frac{c_1}{q_1^T u} \right|.$$

Case 1.2. $q_1^T u = 0$ and $Q_1 u = 0$. Let us assume, with out loss of generality, that $\{x^k\}$ is the smallest norm solution to $f_1(x) \le \epsilon_1^k$. Consider the linear system

$$Q_1 x = Q_1 x^k, \quad q_1^T x = q_1^T x^k. \tag{2.2}$$

Clearly, there exits a solution $\bar{x}^k$ to (2.2) such that

$$\|\bar{x}^k\| \leq \rho \left(\|Q_1 x^k\| + |q_1^T x^k| \right)$$

where $\rho > 0$ is a constant independent of n. Since $f_1(\bar{x}^k) = f_1(x^k) \leq \epsilon_1^k$ and x^k is the smallest norm solution, we have

$$\|x^k\| \leq \|\bar{x}^k\| \leq \rho \left(\|Q_1 x^k\| + |q_1^T x^k| \right).$$

Dividing both sides by $\|x^k\|$ and letting $k \to \infty$, we get

$$1 \leq \rho \left(\|Q_1 u\| + |q_1^T u| \right),$$

contradicting $Q_1 u = 0$ and $q_1^T u = 0$. This completes the proof for the case $m = 1$.

Now we assume, by induction, that the theorem holds with $m \leq l$. Consider the case $m = l+1$. Let $\{x^k\}$ be the smallest norm solution in $X(\epsilon)$. If the sequence $\{x^k\}$ has a bounded subsequence, then any cluster point of this bounded subsequence lies in $X(0)$ and the theorem holds trivially. It remains to consider the case $\|x^k\| \to \infty$. As before, we let $u := \lim_{k\to\infty} x^k/\|x^k\|$. By an argument similar to that used in Case 1.1, we can obtain

$$u^T Q_i u = 0, \quad q_i^T u \leq 0, \quad i = 1, 2, ..., l+1.$$

Since $Q_i \geq 0$, this further implies $Q_i u = 0$ for all i. We once again consider two cases.

Case 2.1. There exists an j such that $q_j^T u < 0$. Without loss of generality, let $j = l+1$. Since the system

$$f_1(x) \leq \epsilon_1^k,\ f_2(x) \leq \epsilon_2^k,\ ...,\ f_l(x) \leq \epsilon_l^k$$

has a solution for each k, the induction hypothesis implies there exists some $\bar{x}$ satisfying

$$f_1(\bar{x}) \leq 0,\ f_2(\bar{x}) \leq 0,\ ...,\ f_l(\bar{x}) \leq 0.$$

Consider the vector $x(t) = \bar{x} + tu$, for $t > 0$. Then, we have

$$\begin{aligned} f_i(x(t)) &= f_i(\bar{x}) + t\nabla f_i(\bar{x})^T u + \frac{t^2}{2} u^T Q_i u \\ &= f_i(\bar{x}) + t\left(Q_i \bar{x} + q_i\right)^T u \\ &\leq f_i(\bar{x}) \leq 0, \qquad \text{for } t > 0,\ i = 1, 2, ..., l. \end{aligned}$$

Furthermore, we have

$$\begin{aligned} f_{l+1}(x(t)) &= f_{l+1}(\bar{x}) + t\left(Q_{l+1}\bar{x} + q_{l+1}\right)^T u + \frac{t^2}{2} u^T Q_{l+1} u \\ &= f_{l+1}(\bar{x}) + t q_{l+1}^T u \\ &\leq 0, \qquad \text{for } t \geq \left| \frac{f_{l+1}(\bar{x})}{q_{l+1}^T u} \right|. \end{aligned}$$

Let

$$t^* = \left| \frac{f_{l+1}(\bar{x})}{q_{l+1}^T u} \right|.$$

Then $x(t^*)$ is a solution in $X(0)$.

Case 2.2. $q_i^T u = 0$, $Q_i u = 0$, for $i = 1, 2, ..., l+1$. Consider the linear system

$$q_i^T x = q_i^T x^k, \quad Q_i x = Q_i x^k, \quad i = 1, 2, ..., l+1. \tag{2.3}$$

Then there exists some $\bar{x}^k$ satisfying (2.3) such that

$$\|\bar{x}^k\| \le \rho \left(\sum_{i=1}^{l+1} \left(\|Q_i x^k\| + |q_i^T x^k| \right) \right),$$

where $\rho > 0$ is independent of k. By equation (2.3), we have

$$f_i(\bar{x}^k) = f_i(x^k) \le \epsilon_i^k, \qquad i = 1, 2, ..., l+1.$$

Since x^k is the smallest norm solution in $X(\epsilon^k)$ (i.e., the least norm solution satisfying the above inequalities), we have

$$\|x^k\| \le \|\bar{x}^k\| \le \rho \left(\sum_{i=1}^{l+1} \left(\|Q_i x^k\| + |q_i^T x^k| \right) \right), \quad \forall k.$$

Dividing the both sides by $\|x^k\|$ and letting $k \to \infty$ yields

$$1 \le \rho \left(\sum_{i=1}^{l+1} \left(\|Q_i u\| + |q_i^T u| \right) \right).$$

This contradicts the conditions $q_i^T u = 0$, $Q_i u = 0$, for $i = 1, 2, ..., l+1$. The proof is complete. **Q.E.D.**

The following example shows that the convexity of the functions f_i, $i = 1, 2, ..., m$, is necessary for Theorem 1 to hold.

Example 1. Consider the following quadratic inequality system:

$$1 \le xy \le 1, \quad 0 \le x \le 0.$$

Clearly the above system does not have a solution, i.e., $X = \emptyset$. On the other hand, the perturbed system

$$1 \le xy \le 1, \quad 0 \le x \le \epsilon$$

has a solution for each $\epsilon > 0$. Thus $X(\epsilon) \ne \emptyset$. This shows that Theorem 1 cannot hold if the convexity of the functions f_i, $i = 1, 2, ..., m$, is removed.

The following is a corollary of Theorem 1. It is kindly pointed out to us by Professor Olvi Mangasarian who, in an early paper [11, Thereom 3.2], derived a similar result for the polyhedral case.

Corollary 1 *A linear (or affine) map of a convex region in $\Re^n$ defined by convex quadratic constraints is closed.*

Proof: Let the convex region $X := X(0)$, and let $y = Ax + a$ be an affine mapping where A and a are some matrix and vector of matching dimension. To prove the claim, we must show that the set

$$Y := \{y \ : \ y = Ax + a, \ \text{for some } x \in X\}$$

is closed. Consider a convergent sequence $\{y^k\} \subset Y$ with $\lim_{k\to\infty} y^k = y^*$. Then there exists some $\{x^k\} \subset X$ such that

$$\|y^* - Ax^k - a\|^2 \le \|y^* - y^k\|^2, \quad f_i(x^k) \le 0, \quad i = 1, 2, ..., m; \ \forall k. \tag{2.4}$$

Consider the following convex quadratic inequality system (in the variable x):

$$\|y^* - Ax - a\|^2 \le 0, \quad f_i(x) \le 0, \quad i = 1, 2, ..., m.$$

From (2.4), this system has a solution when the right hand side of the first inequality is perturbed to $\|y^* - y^k\|^2$, for all k. Letting $k \to \infty$ and using Theorem 1, we conclude that the above system has a solution x^*. In other words, there exists some $x^* \in X$ such that $y^* = Ax^* + a$, showing that $y^* \in Y$. Thus, Y is closed. **Q.E.D.**

We can use Theorem 1 to establish the attainability of the infimum when all the quadratic functions (objective and constraints) are convex.

Corollary 2 *Consider the following convex quadratically constrained quadratic program:*

$$\begin{aligned} \textit{minimize} \quad & f_0(x) := \tfrac{1}{2}x^TQ_0x + q_0^Tx \\ \textit{subject to} \quad & \tfrac{1}{2}x^TQ_1x + q_1^Tx + c_1 \le 0 \\ & \tfrac{1}{2}x^TQ_2x + q_2^Tx + c_2 \le 0 \\ & \qquad \vdots \\ & \tfrac{1}{2}x^TQ_mx + q_m^Tx + c_m \le 0 \end{aligned} \tag{2.5}$$

where each Q_i is symmetric positive semidefinite, $i = 0, 1, ..., m$. Suppose the objective function is bounded from below over the feasible region. Then (2.5) has an optimal solution.

Proof: Let $f^* > -\infty$ denote the infimum of $f_0(x)$ over the feasible region of (2.5). Then there exists a sequence $\{x^k\}$ in the feasible region such that

$$\begin{cases} f_0(x^k) \le f^* + \frac{1}{k} \\ f_i(x) \le 0, \quad i = 1, 2, ..., m. \end{cases}$$

By Theorem 1, this implies that

$$\begin{cases} f_0(x) \le f^* \\ f_i(x) \le 0, \quad i = 1, 2, ..., m. \end{cases}$$

has a solution. This shows that (2.5) has an optimal solution. **Q.E.D.**

Corollary 2 was first established by Terlaky [15] in his study of duality for l_p programming. The approach in this reference is by regularization so that constraint qualification holds and no duality gap exists. In contrast, our proof of this result is more direct and relies on the continuity property of the solution set defined by convex quadratic inequalities. It is natural to ask if it is possible to remove the convexity of the objective function f_0. The following example shows that this is not possible.

Example 2. Consider the following minimization problem in $\Re^4$:

$$\begin{aligned} \text{minimize} \quad & f_0(x_1, x_2, x_3, x_4) := -2x_1x_2 + x_3x_4 + x_1^2 \\ \text{subject to} \quad & f_1(x_1, x_2, x_3, x_4) := x_1^2 - x_3 \le 0, \\ & f_2(x_1, x_2, x_3, x_4) := x_2^2 - x_4 \le 0. \end{aligned}$$

Clearly, f_0 is nonconvex, but f_1 and f_2 are both convex. Moreover, since $x_3 \ge x_1^2$ and $x_4 \ge x_2^2$, we have $x_3x_4 \ge x_1^2x_2^2$. Thus, for any feasible vector $(x_1, x_2, x_3, x_4)^T$, there holds

$$f_0(x_1, x_2, x_3, x_4) \ge -2x_1x_2 + x_1^2x_2^2 + x_1^2 = (x_1x_2 - 1)^2 + x_1^2 - 1 > -1. \tag{2.6}$$

On the other hand, consider the sequence

$$(x_1^k, x_2^k, x_3^k, x_4^k) = (1/k, k, 1/k^2, k^2), \qquad k = 1, 2, ...$$

It can be easily checked that $(x_1^k, x_2^k, x_3^k, x_4^k)^T$ is feasible and

$$f(x_1^k, x_2^k, x_3^k, x_4^k) = \left(\frac{1}{k^2} - 1\right) \to -1.$$

This, together with (2.6), shows that the infimum of f_0 over the feasible set is -1. However, the inequality (2.6) shows that this infimum cannot be attained by any feasible vector.

The above example indicates that if the feasible set involves two nonlinear convex quadratic inequalities and the Hessian matrix of the objective has two negative eigenvalues, then the infimum may not always be attained. Thus, the most we can hope for, as far as the extension of Frank-Wolfe Theorem is concerned, is that it holds when there is at most one nonlinear convex quadratic inequality constraint, or alternatively, if the constraints are all convex and the objective function allows only one negative eigenvalue in its Hessian matrix. In the next two sections we will show respectively that in these two cases a Frank-Wolfe type theorem indeed holds true.

3 One convex quadratic constraint case

In this section we will consider the quadratically constrained quadratic programming problem (QCQP) in which only one quadratic constraint is nonlinear and convex, and all the other constraints are simply affine linear. By applying an orthonormal transformation if necessary, we assume the QCQP is given in the following form:

$$\begin{array}{ll} \text{minimize} & f_0(x, y_0, y) := \frac{1}{2}x^T Qx + q^T x + (g^T + x^T G)y + \frac{1}{2}y^T Hy \\ \text{subject to} & \|x\|^2 \leq h^T y + d, \quad Ax + By \leq c \end{array} \tag{3.1}$$

where y is a vector variable, g, h, c, q are vectors, d is a scalar, and A, B, G, H, Q are matrices of appropriate size.

Throughout this section we wish to establish that if (3.1) is bounded from below over the feasible region, then the infimum of (3.1) is attained. The proof is rather involved and is broken into several intermediate steps. The main idea of the proof is to first perform several variable transformations and problem reformulations so that the essential features of the problem are explicitly revealed, and then apply Hoffman's error bound and Frank-Wolfe Theorem to the resulting problem. The main tool to be used in such process is the following lemma which shows that, under certain conditions, the process of variable restriction and transformation does not change the infimum of the problem nor its attainability.

Lemma 1 *Consider the following minimization problem*

$$\begin{array}{ll} \textit{minimize} & f(x, y) \\ \textit{subject to} & (x, y) \in \Omega \end{array} \tag{3.2}$$

where f is a given function bounded from below over the feasible region Ω. Let $y = g(x)$ be a given function and consider a restricted version of (3.2)*:*

$$\begin{array}{ll} \textit{minimize} & h(x) := f(x, g(x)) \\ \textit{subject to} & x \in \bar{\Omega}, \quad (x, g(x)) \in \Omega \end{array} \tag{3.3}$$

where $\bar{\Omega}$ is some subset of $\Re^n$. Suppose $\inf((3.3)) \leq \inf((3.2))$ *and problem* (3.3) *attains its infimum. Then* $\inf((3.3)) = \inf((3.2))$ *and the infimum of* (3.2) *is also attained.*

Proof: Let $\{x^k\} \subset \bar{\Omega}$ be a sequence such that

$$(x^k, g(x^k)) \in \Omega \quad \text{and} \quad \lim_{k \to \infty} h(x^k) = \inf((3.3)).$$

Since $(x^k, g(x^k))$ is feasible for (3.2), we obtain

$$h(x^k) = f(x^k, g(x^k)) \geq \inf((3.2)).$$

This implies $\inf((3.3)) = \lim_{k \to \infty} h(x^k) \geq \inf((3.2))$. By assumption, we have the inequality in the reverse direction. Thus, $\inf((3.3)) = \inf((3.2))$.

If the infimum of (3.3) is attained, say at a feasible vector x^*, then

$$f(x^*, g(x^*)) = h(x^*) = \inf((3.3)), \quad \text{and} \quad (x^*, g(x^*)) \in \Omega.$$

Thus, $(x^*, g(x^*))$ is feasible for (3.2). Since $\inf((3.3)) = \inf((3.2))$, it follows $(x^*, g(x^*))$ attains the infimum of (3.2). The proof is complete. **Q.E.D.**

We first consider a special case of (3.1). Namely we consider the case where y consists of only one scalar variable y_0. Later we will show that this in fact can be extended to the general setting. Specifically, we first consider the following version of problem (3.1):

$$\begin{array}{ll} \text{minimize} & f(x, y_0) := \frac{1}{2}x^T Q x + q^T x + (f_0 + f^T x)y_0 + \lambda y_0^2 \\ \text{subject to} & \|x\|^2 \le y_0, \quad Ax + ay_0 \le c \end{array} \tag{3.4}$$

where y_0 is a scalar variable, a, c are vectors and A is a matrix of appropriate size. Suppose that (3.4) is bounded from below over the feasible region. We like to show that the infimum of (3.4) is attained.

Let $\{(x^k, y_0^k)\}$ be a sequence of feasible vectors such that

$$f(x^k, y_0^k) \to \inf((3.4)).$$

If $\{y_0^k\}$ is bounded, then so is $\{x^k\}$. In that case any limit point of $\{(x^k, y_0^k)\}$ attains the infimum of (3.4). Now we assume $y_0^k \to \infty$. We consider several separate cases.

Case 1. $\lambda \neq 0$. In this case, the term $\lambda(y_0^k)^2$ will dominate in the objective function since $\|x^k\|^2 \le y_0^k$. Thus, if $\lambda > 0$, then we should have $f(x^k, y_0^k) \to \infty$, a contradiction. Similarly, if $\lambda < 0$, then $f(x^k, y_0^k) \to -\infty$, contradicting the fact that f is bounded from below over the feasible region.

Case 2. $\lambda = 0$ and $a_i > 0$ for some i. In this case, we claim that the feasible region is bounded. In fact, consider the ith linear constraint: $A_i x + a_i y_0 \le c_i$. Since $y_0 \ge \|x\|^2$ and $a_i > 0$. It follows

$$A_i x + a_i \|x\|^2 \le c_i,$$

implying x is bounded. (Here A_i denotes the ith row of A.) Furthermore, we have $0 \le y_0 \le (c_i - A_i x)/a_i$, showing that y_0 is also bounded. Therefore, the feasible region is bounded, in which case the desired property holds trivially.

Case 3. $\lambda = 0$ and $a_i \le 0$ for all i. In this case, $(x, y_0) = (0, 1)$ is a recession direction for the feasible set. Thus, for the objective to be bounded, we must have $(f_0 + f^T x) \ge 0$ for all (x, y_0) feasible. We first consider a subcase in which $\|x^k\|$ is bounded. Then we can select a finite $\bar{y}_0$ such that $(x^k, \bar{y}_0)$ is feasible for all k. Since $y_0^k \to \infty$, it follows that $y_0^k \ge \bar{y}_0$ for all large k. Moreover, since $(f_0 + f^T x^k) \ge 0$, we have $f(x^k, \bar{y}_0) \le f(x^k, y_0^k)$. Thus,

$$\liminf_{k \to \infty} f(x^k, \bar{y}_0) \to \inf((3.4)).$$

As a result, $\bar{y}_0$ together with any limit point of $\{x^k\}$ will attain the infimum of (3.4).

It remains to consider the case $\|x^k\| \to \infty$. Let $I = \{i : a_i = 0\}$ and let $\tilde{I} = \{i : a_i < 0\}$. Since $y_0^k \ge \|x^k\|^2$ and $a_{\tilde{I}} < 0$, the constraints $A_{\tilde{I}}x + a_{\tilde{I}}y_0 \le c_{\tilde{I}}$ are automatically satisfied once $y_0 \ge d$ for some large enough $d > 0$. Consequently, the vector $(x^k, \|x^k\|^2)$ is feasible for large k. Moreover, $f(x^k, \|x^k\|^2) \le f(x^k, y_0^k)$ since $(f_0 + f^T x^k) \ge 0$ and $y_0^k \ge \|x^k\|^2$. This shows $\liminf_{k\to\infty} f(x^k, \|x^k\|^2) \to \inf((3.4))$. Thus, without loss of generality we may assume $y^k = \|x^k\|^2$.

Since $\|x^k\| \to \infty$, we may assume without loss of generality that $x_1^k \to \infty$ (or $x_1^k \to -\infty$; but this case can be treated symmetrically and thus omitted). Now consider the following minimization problem obtained by restricting $y_0 = \|x\|^2$ and $x_1 \ge \sqrt{d}$ in (3.4):

$$\begin{array}{ll} \text{minimize} & f(x, \|x\|^2) := \frac{1}{2}x^T Q x + q^T x + (f_0 + f^T x)\|x\|^2 \\ \text{subject to} & x_1 \ge \sqrt{d}, \quad A_I x \le c_I. \end{array} \tag{3.5}$$

Notice that in our case we have $\lambda = 0$ which explains the absence of the term $\lambda\|y_0\|^2$. Also, the constraints $A_{\tilde{I}}x + a_{\tilde{I}}\|x\|^2 \le c_{\tilde{I}}$ are absent since they are satisfied automatically due to the fact $\|x\|^2 \ge x_1^2 \ge d$. Clearly, if x is a feasible solution of (3.5) then $(x, \|x\|^2)$ is a feasible solution of (3.4). Thus, $\inf((3.5)) \ge \inf((3.4))$, and so the problem (3.5) is bounded from below. On the other hand, since x^k is feasible for (3.5) and

$$\lim_{k\to\infty} f(x^k, \|x^k\|^2) = \inf((3.4)),$$

it follows $\inf((3.5)) \le \inf((3.4))$. However, (3.5) is restricted from (3.4). Therefore, by Lemma 1 if problem (3.5) attains its infimum, then so does problem (3.4). Thus, the key of our analysis will be to show that the infimum of (3.5) is attainable. To this end, we concentrate on the problem of the form:

$$\begin{array}{ll} \text{minimize} & g(x) := \frac{1}{2}x^T Q x + q^T x + (f_0 + f^T x)\|x\|^2 \\ \text{subject to} & Ax \le c \end{array} \tag{3.6}$$

where c is a vector and A is a matrix of appropriate size.

We now prove the following result.

Lemma 2 *Suppose that* (3.6) *is bounded below over the feasible region, then the infimum of* (3.6) *is attained.*

Proof: Let $\{x^k : k = 1, 2, ...\}$ be a sequence of feasible vectors such that $g(x^k) \downarrow \inf((3.6))$ as $k \to \infty$. Clearly, if this sequence is bounded then the lemma holds trivially. In what follows, let us assume $\|x^k\| \to \infty$ as $k \to \infty$.

First we note that the sequence $\{f_0 + f^T x^k : k = 1, 2, ...\}$ is bounded. This is because

$$|g(x^k)| \ge \left(|f_0 + f^T x^k| - \tfrac{1}{2}\|Q\| - \|q\|/\|x^k\|\right)\|x^k\|^2$$

and so if $\{f_0 + f^T x^k : k = 1, 2, ...\}$ were not bounded then the subsequence $\{g(x^k)\}$ would tend to infinity, contradicting the selection of $\{x^k : k = 1, 2, ...\}$.

Without loss of generality, we assume $(f_0 + f^T x^k) \to f^*$ for some scalar f^*, as $k \to \infty$. Furthermore, by a rescaling of objective function if necessary, we assume $\|f\| = 1$. Define an affine orthonormal coordinate transformation $L : \Re^n \mapsto \Re^n$:

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} := L(x) = \begin{bmatrix} f^T \\ H \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} f_0 - f^* \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where H is an $(n-1) \times n$ matrix whose rows are mutually orthogonal, normalized (to 1) and orthogonal to f. Let us denote $\bar{z} = (z_2, z_3, ..., z_n)^T$. Then, under the transformation L, we have $z_1 = (f_0 + f^T x) - f^*$. Furthermore, using the orthonormality, we have

$$\|x\|^2 = (z_1 - f_0 + f^*)^2 + \sum_{i=2}^{n} \bar{z}_i^2 = (z_1 - f_0 + f^*)^2 + \|\bar{z}\|^2.$$

Substituting this and the expression of L into the problem (3.6), simplifying the expressions and renaming the matrices and vectors, we can rewrite the objective function in the form

$$g(x) = h(z) = \tfrac{1}{2} z^T Q z + q^T z + z_1 \|z\|^2,$$

and rewrite the constraints as

$$A\bar{z} + a z_1 \le c$$

for some new matrices Q, A and new vectors a and q. In this way, the problem (3.6) is transformed into

$$\begin{array}{ll} \text{minimize} & h(z) := \tfrac{1}{2} z^T Q z + q^T z + z_1 \|z\|^2 \\ \text{subject to} & A\bar{z} + a z_1 \le c \end{array} \tag{3.7}$$

with $\inf((3.6)) = \inf((3.7))$. Since $g(x^k) \to \inf((3.6))$ and $(f_0 + f^T x^k) \to f^*$, we obtain a sequence $\{z^k\}$ which is feasible for (3.7) and

$$h(z^k) \to \inf((3.6)), \quad z_1^k \to 0 \quad \text{and} \quad \|z^k\| \to \infty. \tag{3.8}$$

Clearly, (3.7) has a finite infimum. Moreover, if this infimum is attained, then under the inverse mapping of L, the infimum of (3.6) is also attained.

To show that the infimum of (3.7) is attainable, we consider the following linearly constrained QP (in the variable $\bar{z}$) obtained by restricting $z_1 = 0$ in (3.7):

$$\begin{array}{ll} \text{minimize} & h(0, \bar{z}) := \tfrac{1}{2}(0, \bar{z})^T Q \begin{pmatrix} 0 \\ \bar{z} \end{pmatrix} + q^T \begin{pmatrix} 0 \\ \bar{z} \end{pmatrix} \\ \text{subject to} & A\bar{z} \le c \end{array} \tag{3.9}$$

Since (3.9) is bounded from below, we have

$$\tau := \min \left\{ \frac{1}{2}(0, \bar{u})^T Q \begin{pmatrix} 0 \\ \bar{u} \end{pmatrix} \ : \ A\bar{u} \le 0, \ \|\bar{u}\| = 1 \right\} \ge 0.$$

We consider two cases. First, if $\tau > 0$, then we consider any vector $z = (z_1, \bar{z})$ satisfying $A\bar{z} + z_1 a \leq c$. By Hoffman's bound, there exists some y with $Ay \leq 0$ such that

$$\|y - \bar{z}\| \leq \mu(\|c\| + |z_1|\|a\|),$$

where $\mu > 0$ is a constant independent of z; i.e.

$$\|y - \bar{z}\| = O(1 + |z_1|). \tag{3.10}$$

Therefore, we obtain

$$\begin{aligned}
\frac{1}{2}(z_1, \bar{z})^T Q \begin{pmatrix} z_1 \\ \bar{z} \end{pmatrix} &\geq \frac{1}{2}(0, \bar{z})^T Q \begin{pmatrix} 0 \\ \bar{z} \end{pmatrix} - O(|z_1|\|\bar{z}\| + z_1^2) \\
&\geq \frac{1}{2}(0, y)^T Q \begin{pmatrix} 0 \\ y \end{pmatrix} - O(\|\bar{z}\|\|y - \bar{z}\| + \|y - \bar{z}\|^2)) - O(|z_1|\|\bar{z}\| + z_1^2) \\
&\geq \tau\|y\|^2 - O(|z_1| + 1 + |z_1|\|\bar{z}\| + z_1^2)
\end{aligned}$$

for all $z = (z_1, \bar{z})$ satisfying $A\bar{z} + z_1 a \leq c$. By (3.10) we further obtain

$$\frac{1}{2}(z_1, \bar{z})^T Q \begin{pmatrix} z_1 \\ \bar{z} \end{pmatrix} \geq \tau\|\bar{z}\|^2 - O(|z_1| + 1 + |z_1|\|\bar{z}\| + z_1^2).$$

This implies

$$\begin{aligned}
h(z^k) &= \frac{1}{2}{z^k}^T Q z^k + q^T z^k + z_1^k\|z^k\|^2 \\
&\geq (\tau + z_1^k)\|\bar{z}^k\|^2 - O(|z_1^k| + 1 + (|z_1^k| + \|q\|)\|\bar{z}^k\| + (z_1^k)^2) \\
&\to \infty, \quad \text{as } k \to \infty,
\end{aligned}$$

where the last step is due to $\tau > 0$, $\|z^k\| \to \infty$ and $z_1^k \to 0$, This contradicts the assumption that $h(z^k) \to \inf((3.6))$.

Now consider the case $\tau = 0$. In this case, there exists some $u = (0, \bar{u})$ with $\|u\| = 1$ such that

$$\frac{1}{2}(0, \bar{u})^T Q \begin{pmatrix} 0 \\ \bar{u} \end{pmatrix} = 0, \quad \text{and} \quad A\bar{u} \leq 0. \tag{3.11}$$

We claim $z_1^k \geq 0$ for all k. Indeed, suppose $z_1^k < 0$ for some k. Then let $w^k(t) = z^k + tu$, for $t > 0$. For each $t > 0$, we note that $w^k(t)$ is a feasible vector of (3.9), and $h(w^k(t))$ is a quadratic function of t. Moreover, it can be checked using the condition (3.11) that the coefficient of the t^2 term in this quadratic polynomial is equal to z_1^k. Now if $z_1^k < 0$, then $h(w^{\bar{k}}(t))$ would tend to $-\infty$ for large t, contradicting the boundedness assumption for the problem (3.7). Thus, we have $z_1^k \geq 0$ for all k. Without loss of generality we assume $z_1^k \downarrow 0$.

As noted before, (3.9) must have a finite infimum as well. By the Frank-Wolfe theorem, this infimum is attained. Thus, by Lemma 1, if we can show $\inf((3.9)) = \inf((3.7))$, then any minimum solution of (3.9) also achieves minimum for (3.7) (Lemma 1). Clearly, $\inf((3.9)) \geq \inf((3.7))$. It remains to show $\inf((3.9)) \leq \inf((3.7))$. Consider the linear system in $\bar{z}$:

$$A\bar{z} \leq c.$$

By Hoffman's error bound [7], there exists $\bar{y}^k$ such that

$$A\bar{y}^k \leq c, \quad \|\bar{y}^k - \bar{z}^k\| = O(z_1^k), \tag{3.12}$$

where the constant in the big "O" notation is independent of k. Let $y^k = (0, (\bar{y}^k)^T)^T$. Then, by Taylor expansion, we obtain

$$\begin{aligned} h(y^k) - h(z^k) &= \tfrac{1}{2}(y^k - z^k)^T Q(y^k - z^k) + (Qz^k + q)^T(y^k - z^k) - z_1^k\|z^k\|^2 \\ &\leq O(\|y^k - z^k\|^2) + O(\|y^k - z^k\|\|z^k\|) - z_1^k\|z^k\|^2 \\ &\leq O((z_1^k)^2) + O(\|z^k\|z_1^k) - z_1^k\|z^k\|^2 \\ &< 0, \qquad \text{for large } k, \end{aligned}$$

where the second last step follows from (3.12) and the last step is due to $\|z^k\| \to \infty$ and $z_1^k \downarrow 0$. Taking limit as $k \to \infty$, we obtain

$$\liminf_{k\to\infty} h(y^k) - \inf((3.7)) \leq 0.$$

This implies $\inf((3.9)) \leq \inf((3.7))$. By Lemma 1 the desired result follows. **Q.E.D.**

Using Lemma 2 we immediately obtain the next result.

Lemma 3 *If* (3.4) *is bounded below over the feasible region, then the infimum of* (3.4) *is attained.*

Proof: We continue our discussion preceding Lemma 2. The only remaining case to be considered there is: $\lambda = 0$, $a_i \leq 0$ for all i, and that the norm of the sequence approaching inf((3.4)) is unbounded. Now we apply Lemma 2 to conclude that (3.5) has an optimal solution. Since $\inf((3.5)) = \inf((3.4))$, it follows that an optimal solution for (3.4) also exists. **Q.E.D.**

Now we are in a position to discuss the general problem (3.1). But before we start the discussion, we first note a property for quadratic programming with linear constraints.

Lemma 4 *Consider the following QP:*

$$\begin{array}{lll} (QP) & \textit{minimize} & \frac{1}{2}y^TQy + q^Ty \\ & \textit{subject to} & By \leq c. \end{array}$$

Suppose that there exist infinite sets U and V such that for any $q \in U$ and $c \in V$, the problem (QP) has an optimal solution. Then, there exist infinite subsets $U' \subseteq U$, $V' \subseteq V$, and an affine linear mapping L independent of U and V such that for all $q \in U'$, $c \in V'$, an optimal solution y of (QP) can be expressed as:

$$y = L(q, c).$$

Proof: First of all, by the Frank-Wolfe theorem, optimal solutions for (QP) are always attainable. Secondly, any optimal solution of (QP) will be a solution to the following LCP, which is the KKT condition, necessary for all optimal solutions to satisfy:

$$(LCP) \quad \left\{ \begin{array}{rcl} Qy + q + B^T x &=& 0 \\ By + s &=& c \\ s_i x_i &=& 0, \quad \text{for all } i \\ s \geq 0, \; x \geq 0 && \end{array} \right.$$

For convenience, assume that B has full-column rank. Let I be such an index set that B_I is invertible. Let J be the complement of I. Eliminating variables y we get an equivalent formulation of (LCP) as follows:

$$(LCP)' \quad \left\{ \begin{array}{rcl} x_I &=& -B_I^{-T}(QB_I^{-T}c_I + q) + B_I^{-T}QB_I^{-T}s_I - B_I^{-T}B_J^T x_J \\ s_J &=& c_J - B_J B_I^{-T}c_I + B_J B_I^{-T}s_I. \end{array} \right.$$

Interchanging variable(s) from right to left side is called *pivoting*. We call variables on the left side *basic* and variables on the right side *nonbasic*.

Observe that any complementary solution (x, s) can be obtained by simple linear transformation (pivots) from $(LCP)'$.

Since there are only finite number of possible ways to partition the whole index set, we conclude that there must be infinite subsets $U' \subseteq U$ and $V' \subseteq V$ such that for any $q \in U'$ and $c \in V'$ the basic-nonbasic partition on the vector variable (x, s) remains constant. As long as basic-nonbasic partition remains unchanged, the solution (x, s) can be expressed as an affine linear function in q and c. Finally, noting that the variable y is related to s affine linearly, i.e. $y = B_I^{-1}c_I - B_I^{-1}s_I$, the lemma thus follows. **Q.E.D.**

For ease of further discussion we introduce an auxiliary variable y_0 in (3.1) and consider

$$\begin{array}{ll} \text{minimize} & f(x, y_0, y) := \frac{1}{2}x^T Qx + q^T x + (g^T + x^T G)y + \frac{1}{2}y^T Hy \\ \text{subject to} & \|x\|^2 \leq y_0, \quad Ax + By \leq c \\ & y_0 = h^T y + d. \end{array} \tag{3.13}$$

Lemma 5 *Suppose* (3.13) *is bounded below over the feasible region, then the infimum of* (3.13) *is attained.*

Proof: Consider a sequence $\{(x^k, y_0^k, y^k)\}$ in the feasible region of (3.1) such that

$$f(x^k, y_0^k, y^k) \downarrow \inf((3.13)).$$

For each fixed k, consider the following linearly constrained QP:

$$\begin{array}{ll} \text{minimize} & (g^T + (x^k)^T G)y + \frac{1}{2}y^T Hy \\ \text{subject to} & By \leq c - Ax^k, \quad h^T y = y_0^k - d. \end{array} \tag{3.14}$$

Problem (3.14) is bounded from below since (3.13) is bounded from below. It follows from the Frank-Wolfe theorem that its infimum is attained. Now we apply Lemma 4 to conclude that there exists an infinite subset $\mathcal{K} \subset \{1, 2, 3, ...\}$ such that for all $k \in \mathcal{K}$ an optimal solution of (3.14) can be expressed as

$$\bar{y}^k = L_1 x^k + y_0^k l_2 + l_3$$

for some fixed matrix L_1 and fixed vectors l_2, l_3. Substituting the above relation

$$y = L_1 x + y_0 l_2 + l_3$$

into (3.13) we get a problem in the form of (3.4). That problem is bounded below and has the same infimum as (3.13). By Lemma 3 and Lemma 1 we conclude that the infimum of (3.13) is attained too. **Q.E.D.**

By Lemma 5, using some simple variable transformation we have proven the following main theorem of this section:

Theorem 2 *Suppose $Q_1 \geq 0$ and $Q_i = 0$ for $i = 2, 3, ..., m$. Then if the objective function $f_0(x)$ is bounded over the feasible set X, then the infimum of (P) is attained.*

Theorem 2 shows that if there is only one nonlinear convex quadratic inequality in the constraints, then the (finite) infimum of the quadratic objective function is always attained. The following example shows that the convexity of this quadratic inequality constraint cannot be relaxed.

Example 3. Consider the following quadratically constrained QP in $\Re^2$:

$$\begin{array}{ll} \text{minimize} & x^2 \\ \text{subject to} & xy \geq 1, \quad y \geq 0. \end{array}$$

Clearly, the infimum of the above problem is equal to 0, but it is never attained by any feasible solution. Notice in this case, the quadratic constraint $xy \geq 1$ is nonconvex, even though the feasible region is convex.

4 Quasi-convex objective function

Example 2 shows that Frank-Wolfe type theorem cannot hold for convex quadratically constrained problem if the Hessian of the quadratic objective function has more than one negative eigenvalues. A natural question arises: Can there still be such a theorem when the Hessian of the objective has no more than one negative eigenvalue? In this section we will show that under some conditions the answer to this question is positive.

A thorough treatment on concavity and various extensions can be found in [2]. In particular, conditions have been derived for the characterization of a quadratic quasi-convex or pseudo-convex function. Consider a general quadratic function $f(x) = \frac{1}{2}x^T Qx + q^T x$ with $x \in \Re^n$. Then, by performing a sequence of orthonormal, scaling and affine transformations it can be put into the following canonical form:

$$\bar{f}(y) = -\frac{1}{2}\sum_{i=1}^{p} y_i^2 + \frac{1}{2}\sum_{i=p+1}^{r} y_i^2 + \sum_{i=r+1}^{n} \lambda_i y_i.$$

If $p \geq 1$, this function is in general not convex. Furthermore, it is shown that this function can be quasi-convex or pseudo-convex on a solid domain only if $p \leq 1$. Thus, the interesting case left is $p = 1$. In that case consider two (convex) second-order cones (also known as ice-cream cones):

$$\mathcal{C}_1 = \{y : \, y_1^2 - \sum_{i=2}^{r} y_i^2 \geq 0, \, y_1 \geq 0\}$$

and

$$\mathcal{C}_2 = \{y : \, y_1^2 - \sum_{i=2}^{r} y_i^2 \geq 0, \, y_1 \leq 0\}.$$

Lemma 6 *The function $\bar{f}(y)$ is quasi-convex on a convex set $\mathcal{C}$ if and only if*

- *$p \leq 1$;*
- *$\mathcal{C} \subseteq \mathcal{C}_1$ or $\mathcal{C} \subseteq \mathcal{C}_2$;*
- *If $r < n$ then $\lambda_i = 0$ for $i = r+1, ..., n$.*

In simple words, the Hessian of a quasi-convex quadratic function has exactly one negative eigenvalue. Moreover, its gradient must always be contained in the range space of the Hessian matrix.

A similar characterization can be given for pseudo-convex quadratic functions. In fact, the only difference is that for a pseudo-convex function the domain $\mathcal{C}$ must be *strictly* contained in either $\mathcal{C}_1$ or $\mathcal{C}_2$.

Now we consider the following problem:

$$\begin{array}{lll} (P)' & \text{minimize} & f_0(x) = \frac{1}{2}x^T Q_0 x + q_0^T x \\ & \text{subject to} & f_i(x) = \frac{1}{2}x^T Q_i x + q_i^T x + c_i \leq 0, \; i = 1, 2, ..., m \\ & & Ax \leq b \end{array}$$

where $f_0(x)$ is a quasi-convex function over the polyhedral set $\{x \, : \, Ax \leq b\}$ and all the quadratic constraint functions $f_i(x)$'s are convex ($i = 1, ..., m$). Notice that compared with (P), problem $(P)'$ has its polyhedral constraints explicitly stated. Also, we have required that $f_0(x)$ is quasiconvex over a polyhedral set containing the feasible set of $(P)'$ rather than just over the feasible set itself. Such a requirement is not restrictive, since the recession directions of convex quadratic inequality constraints form a polyhedral set. Therefore, by adding cutting planes if necessary we can contain the feasible set of $(P)'$ by a polyhedron inside the second-order cone. In this way, the quasi-convexity of $f_0(x)$ over the feasible region of $(P)'$ is the same as the quasi-convexity of $f_0(x)$ over the containing polyhedral set (Lemma 6).

Theorem 3 *If $(P)'$ is bounded below and $f_0(x)$ is quasi-convex over the polyhedral set $\{x : Ax \le b\}$, then the optimal solution set of $(P)'$ is nonempty.*

Proof: The proof is quite similar, in spirit, to the analysis in Section 2. We will apply induction on m: the number of quadratic constraints. If $m = 1$, the theorem reduces to a special case of Theorem 2, and hence is true.

Assume that the theorem holds true for $m \le l$. Also, by the quasi-convexity of $f_0(x)$, we can assume without loss of generality that $f_0(x)$ (upon a linear transformation) is given by

$$f_0(x) = -x_1^2 + x_2^2 + \cdots + x_n^2 \tag{4.1}$$

and

$$Ax \le b \quad \Rightarrow \quad f_0(x) \le 0 \text{ and } x_1 \ge 0. \tag{4.2}$$

Now consider the case when $m = l + 1$. We construct a sequence of truncated problems as follows:

$$\begin{array}{lll} (P)'_k & \text{minimize} & f_0(x) = \frac{1}{2}x^TQ_0x + q_0^Tx \\ & \text{subject to} & f_i(x) = \frac{1}{2}x^TQ_ix + q_i^Tx + c_i \le 0, \;\; i = 1, 2, ..., l+1, \\ & & Ax \le b, \\ & & \|x\| \le k \end{array}$$

with $k = 1, 2, \ldots$. For each $(P)'_k$ an optimal solution exists due to the compactness of the feasible region. Let x^k denote the minimum norm solution of $(P)'_k$.

Certainly, if a subsequence of $\{x^k : k = 1, 2, ...\}$ is bounded, then the theorem follows immediately. Without loss of generality we assume that $\|x^k\| \to \infty$ and

$$\lim_{k\to\infty} \frac{x^k}{\|x^k\|} = u, \quad \text{for some } u \text{ with } \|u\| = 1.$$

Since $\{f_0(x^k) : k = 1, 2, ...\}$ is a monotonically decreasing sequence and that $f_i(x)$ is convex quadratic for $i = 1, 2, ..., l+1$ it follows that

$$\begin{aligned} u^TQ_0u &= 0 && (4.3) \\ u^TQ_iu &= 0, \quad \text{for } i = 1, 2, ..., l+1, && (4.4) \\ q_i^Tu &\le 0, \quad \text{for } i = 1, 2, ..., l+1, && (4.5) \\ Au &\le 0. && (4.6) \end{aligned}$$

Now we consider two separate cases:

Case 1. There exists $j \in \{1, 2, ..., l+1\}$ such that $q_j^T u < 0$. Without loss of generality, assume $j = l+1$. In this case we consider

$$\begin{array}{ll} \text{minimize} & f_0(x) = \frac{1}{2} x^T Q_0 x + q_0^T x \\ \text{subject to} & f_i(x) = \frac{1}{2} x^T Q_i x + q_i^T x + c_i \leq 0, \ i = 1, 2, ..., l \\ & Ax \leq b. \end{array}$$

There are two possibilities with the above minimization problem: Either it is unbounded, or it has an attainable minimum solution by the induction hypothesis. In both situations due to quasiconvexity there exists a solution x' such that $f_0(x') = \inf_{k \geq 1} f_0(x^k)$ and

$$\begin{array}{rcl} f_i(x') & \leq & 0, \quad \text{for } i = 1, 2, ..., l, \\ Ax' & \leq & b. \end{array} \tag{4.7}$$

If $f_{l+1}(x') \leq 0$, then x' is an optimal solution for $(P)'$, and so the theorem follows. Now consider the other possibility, i.e. $f_{l+1}(x') > 0$. Let us denote $x' = (x_1', \bar{x}')$, where $\bar{x}' = (x_2', x_3', ..., x_n')$. The notation of $\bar{x}^k$ is defined similarly: $\bar{x}^k = (x_2^k, x_3^k, ..., x_n^k)$. Recall that $f_0(x^k) \downarrow f_0(x')$ and $f_0(x)$ is quasi-convex over the domain. We claim that

$$u^T \nabla f_0(x') \leq 0. \tag{4.8}$$

To see this we recall from (4.2) that

$$x_1' \geq \|\bar{x}'\|, \quad x_1^k \geq \|\bar{x}^k\|, \quad \forall k.$$

By the definition of $f_0(x)$ (cf. (4.1)), we have

$$\begin{array}{rcl} \nabla f_0(x')^T (x^k - x') & = & 2(-x_1', \bar{x}')^T (x^k - x') \\ & = & 2(-x_1', \bar{x}')^T x^k - 2 f_0(x') \\ & = & -2x_1' x_1^k + 2 \sum_{i=2}^n x_i' x_i^k - 2 f_0(x') \\ & \leq & -2x_1' x_1^k + 2\|\bar{x}'\| \|\bar{x}^k\| - 2 f_0(x') \\ & \leq & -2\|\bar{x}'\|(x_1^k - \|\bar{x}^k\|) - 2 f_0(x') \\ & = & \dfrac{2\|\bar{x}'\|}{x_1^k + \|\bar{x}^k\|} f_0(x^k) - 2 f_0(x'), \end{array}$$

where the second last step follows from the preceding bound. Dividing both sides by $\|x^k - x'\|$ and letting $k \to \infty$, we immediately obtain

$$\limsup_{k \to \infty} \nabla f_0(x')^T \frac{(x^k - x')}{\|x^k - x'\|} \leq 0,$$

implying (4.8) holds.

Now we let

$$t^* := -\frac{f_{l+1}(\bar{x})}{q_{l+1}^T u} \ (> 0) \tag{4.9}$$

and

$$x'(t^*) := x' + t^* u.$$

Clearly,

$$\begin{aligned} f_{l+1}(x'(t^*)) &= f_{l+1}(x') + t^*(\nabla f_{l+1}(x'))^T u \\ &= f_{l+1}(x') + t^* q_{l+1}^T u = 0, \end{aligned} \tag{4.10}$$

where we used $Q_{l+1}u = 0$ and (4.9). By (4.4) and (4.5), we have

$$\begin{aligned} f_i(x'(t^*)) &= f_i(x') + t^*(\nabla f_i(x'))^T u \\ &= f_i(x') + t^* q_i^T u \\ &\le f_i(x') \le 0 \end{aligned} \tag{4.11}$$

for all $i = 1, 2, ..., l$. Finally, by (4.3) and (4.8) we get

$$f_0(x'(t^*)) = f_0(x') + t^* u^T \nabla f_0(x') \le f_0(x') = \inf((P)'). \tag{4.12}$$

Summarizing (4.10), (4.11) and (4.12) we conclude that $x'(t^*)$ is an optimal solution for $(P)'$.

Case 2. $q_i^T u = 0$ for all $i = 1, 2, ..., l+1$. In this case, we know that both u and $-u$ are recession directions for $(P)'$.

For any fixed k, since $f_0(x^p) < f_0(x^k)$ for all $p > k$, it follows from the quasi-convexity of f_0 that

$$(x^p - x^k)^T \nabla f_0(x^k) \le 0, \quad \text{for all } p > k.$$

Dividing $\|x^p\|$ on both sides and letting $p \to \infty$ we have

$$u^T \nabla f_0(x^k) \le 0.$$

Since u is a recession direction and $(P)'$ is bounded below, we conclude that

$$u^T \nabla f_0(x^k) = 0 \tag{4.13}$$

for all $k = 1, 2, ...$

Now, because $u = \lim_{k\to\infty} x^k/\|x^k\|$, it follows that for sufficiently large k we have

$$u^T x^k > 0,$$

and also we have the implication:

$$(Au)_i < 0 \quad \Rightarrow \quad (Ax^k - b)_i < 0 \tag{4.14}$$

for each index i. This means that there exists $\epsilon_0 > 0$ such that for all $0 < \epsilon \le \epsilon_0$,

$$x^k(\epsilon) := x^k - \epsilon u$$

is a feasible solution for $(P)'$ and $(P)'_k$. By (4.13),

$$f_0(x^k(\epsilon)) = f_0(x^k)$$

for $0 < \epsilon \le \epsilon_0$. However,

$$\|x^k(\epsilon)\|^2 = \|x^k\|^2 - 2\epsilon(u^T x^k) + \epsilon^2$$

and so we can choose $\epsilon > 0$ sufficiently small such that

$$\|x^k(\epsilon)\| < \|x^k\|.$$

This contradicts the fact that x^k is the minimum norm solution for $(P)'_k$, implying that **Case 2** can never occur. The proof is complete. **Q.E.D.**

5 Conclusions

Unboundedness and existence of solutions are important issues in optimization theory; see Auslender [1] for an extensive survey. In this paper we have provided a comprehensive study of the existence of optimal solutions for quadratically constrained quadratic programs in the absence of convexity. Our results extend the classical Frank-Wolfe theorem for linearly constrained QP to the case where the constraints are convex quadratic. As a basic step we first established a continuity property of the feasible set defined by a set of convex quadratic inequalities. By example we showed that if the Hessian of the objective function has more than one negative eigenvalue or if there are more than one quadratic inequality constraints, then optimal solution may not exist. On the other hand, we showed that an optimal solution always exists when either there is one convex quadratic constraint (plus linear constraints) or the objective function is quasi-convex.

Note that our results do not address the issue of how an optimal solution can be obtained. For a linearly constrained quadratic program this is a hard problem in general. For example Matsui [12] showed that minimizing the product of two linear functions over a polyhedron is NP-hard.

The problem of minimizing the quartic polynomial $(x_1x_2 - 1)^2 + x_1^2$ over the whole plane does not have an optimal solution. This means that Frank-Wolfe type theorem cannot be generalized to the situation where the objective function is a fourth (or higher) order multi-variate polynomial, even though the constraint set is polyhedral. However, as shown in Section 3, this generalization is possible for a special type of multi-variate polynomial function of degree three. Thus, an open question still remains: Can we prove or disprove the same result for a general multi-variate polynomial function of degree three?

Acknowledgement: We are deeply indebted to Wu Li who provided constructive comments on an early draft of this paper. In addition, we wish to express our appreciation to Olvi Mangasarian for bringing to our attention Corollary 1 and for his encouragement on this research. Finally, we wish to thank Jong-Shi Pang and Yinyu Ye for several stimulating discussions on the general subject of this paper.

References

[1] A. Auslender, "How to deal with the unbounded in optimization: Theory and algorithms", *Mathematical Programming* Series B, 79 (1997) 3-8.

[2] M. Avriel, W.E. Diewert, S. Schaible and I. Zang, *Generalized Concavity*, Plenum Press, New York, 1988.

[3] E. Blum and W. Oettli, "Direct proof of the existence theorem in quadratic programming", *Operations Research* 20 (1972) 165–167.

[4] V. Chvátal, *Linear Programming*, W.H. Freeman and Company, New York, 1983.

[5] B.C. Eaves, "On quadratic programming", *Management Science* 17 (1971) 698–711.

[6] M. Frank and P. Wolfe, "An algorithm for quadratic programming", *Naval Research Logistics Quarterly* 3 (1956) 95–110.

[7] A. J. Hoffman, "On approximate solutions of systems of linear inequalities", *Journal of Research of the National Bureau of Standards* 49 (1952) 263–265.

[8] Z.-Q. Luo, "On the solution set continuity of the convex quadratic feasibility problem", Unpublished manuscript, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada; September, 1995.

[9] Z.-Q. Luo, J.F. Sturm and S. Zhang, "Duality results for conic convex programming", Report 9719/A, Econometric Institute, Erasmus University Rotterdam, 1997.

[10] Z.-Q. Luo and P.Y. Tseng, "Error bound and convergence analysis of matrix splitting algorithms for the affine variational inequality problem", *SIAM J. Optimization* 2 (1992) 43–54.

[11] O. L. Mangasarian and L. L. Schumaker, "Discrete Splines via Mathematical Programming", *SIAM Journal on Control* 9 (1971) 174–183.

[12] T. Matsui, "NP-hardness of linear multiplicative programming and related problems", Report METR95-13, University of Tokyo, Japan, 1995.

[13] A.F. Perold, "A generalization of Frank-Wolfe theorem", *Mathematical Programming* 18 (1980) 215–227.

[14] S. M. Robinson, "Some continuity properties of polyhedral multifunctions", *Mathematical Programming Study* 14 (1981) 206–214.

[15] T. Terlaky, "On l_p programming", *European Journal of Operations Research* 22 (1985) 70–100.