A Parametric Solution for Local and Global Optimization

by

Baoyan Ding

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 1996

0-612-21341-2

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

The goal of this thesis is to present a method which when applied to certain nonconvex quadratic programming problems will locate the global minimum, all isolated local minima and some of the non-isolated local minima. The method proceeds by formulating a (multi) parametric QP or LP in terms of the data of the given nonconvex quadratic programming problem. Based on the solution of the parametric QP or LP, a minimization problem is formulated. This problem is unconstrained and piece-wise quadratic. A key result is that the isolated local minimizers (including the global minimizer) of the original non-convex problem are in one to one correspondence with those of the derived unconstrained problem. As an application, the method is applied to the problem of determining if a given symmetric matrix is copositive on a given polyhedral cone. We show that the copositivity problem in which the matrix has exactly one negative value can be solved in polynomial time. The results established for non-convex quadratic programming problems are generalized to the non-convex problems in which the objective function is nonquadratic and the constraints are nonlinear.

Acknowledgments

I would like to give my special thanks to Dr. M.J. Best, my supervisor. Without his great encouragement, patience and advise, I would not have finished this thesis at this time. I am very impressed by and appreciative of the teaching and the guidance he provided.

I would also like to thank the Department of Combinatorics and Optimization, the Faculty of Graduate Studies and my supervisor for arranging financial support during my studies at the University of Waterloo.

I am grateful to Dr. K.G. Murty, external examiner, and also Dr. D. Fuller, Dr. H. Wolkowicz and Dr. L. Tuncel for their review of my thesis and their constructive suggestions.

I would also like to thank Dr. H. Wolkowicz and Dr. L. Tuncel for providing me with useful references and helpful discussions during the writing of this thesis.

Last but not the least, my thanks go to Dongmei Liu, my wife, for her assistance, patience and encouragement in my doctoral studies.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 An Overview of the Thesis

In this thesis, we consider the following quadratic programming problem

$$\min\ \{c'x + \frac{1}{2}x'Cx \mid Ax \leq b\}, \tag{1.1}$$

where $c \in \mathrm{E}^n$, $b \in \mathrm{E}^m$, C is an (n, n) symmetric matrix and A is an (m, n) matrix. If C is positive semidefinite, (1.1) is a convex quadratic programming problem. This type of problem can be solved by any convex quadratic programming algorithm. For example, see Van de Panne and Whinston [30], Best and Ritter [10], and Gill and Murray [15]. This type of problem can also be solved in polynomial time. For example, see Monteiro and Adler [20], and Ye and Tse [34]. If C is not positive semidefinite, (1.1) is a non-convex quadratic programming problem. In this thesis, we will focus our attention on non-convex QP's. This type of problem was proved to be NP-hard even when C has exactly one negative eigenvalue (see Pardalos and Vavasis [26]). In the past several years, some algorithms have been developed to locate a global minimizer for a non-convex QP. For example, see Konno *et al.* [17], Bomze and Danninger [11], Murty [21], and Vavasis [32] and [33]. Unlike these algorithms, our goal is to develop an algorithm to locate both the global minimum and the local minima. The approach we use here is new. It is based on a one to one correspondence between local and global minimizers for the given problem and the unconstrained derived problem (Theorems 2.2.4 and 2.2.5). Unlike the convex case, the number of isolated local minimizers for the non-convex QP could be exponential in the dimensionality of the problem. For example, the problem

$$\min\{-\sum_{i=1}^{n}(x_i - \frac{1}{2})^2 \mid 0 \le x_i \le 1,\ i = 1, \cdots, n\}$$

has 2^n isolated local minimizers. In general, it will be difficult to develop an efficient algorithm to locate all isolated local minimizers and some local minimizers. However our motivation is a little bit different from this. By developing algorithms for finding all isolated local minimizers and some local minimizers, we hope that it can help us to further understand the structure of non-convex quadratic programming problems and it can provide intuition for developing further efficient algorithms.

In Chapter 2 of this thesis, we present a method which when applied to certain non-convex QP will locate the global minimum, *all* isolated local minima and some of the non-isolated local minima. The method proceeds by formulating a (multi) parametric convex QP in terms of the data of the given non-convex QP. Based on the solution of the parametric QP, an unconstrained minimization problem is formulated. This problem is piece-wise quadratic. A key result is that the isolated local minimizers (including the global minimizer) of the original non-convex problem are in one to one correspondence with those of the derived unconstrained problem. A numerical procedure is developed for a special class of non-convex QP's in which C has exactly one negative eigenvalue.

In Chapter 3, we develop a numerical decomposition method based on the theory established in Chapter 2 to locate the global minimum, all isolated local minima and some of the nonisolated local minima for general indefinite QP. The procedure proceeds by formulating a (multi) parametric LP in terms of the data of the given nonconvex QP. We use this and a decomposition procedure to transform the given n dimensional nonconvex QP having m linear inequality constraints into k subproblem QP's, each of which has $n-1$ variables and m constraints, where $1 \leq k \leq m$ and k depends on the problem data. Special techniques are developed to ensure that k is small. The decomposition procedure may then be applied to the subproblem QP's. A branch of the decomposition procedure terminates when either the subproblem is concave, the subproblem is convex, the subproblem has a Hessian matrix having exactly one negative eigenvalue, or, the subproblem has dimension 1.

In Chapter 4, we consider the application of the algorithms developed in Chapter 2 and 3 to a copositivity problem. Especially, for a given (n, n) symmetric matrix with exactly one negative eigenvalue and a polyhedral cone, it is proved that determining if the matrix is copositive on the cone can be solved in polynomial time and an algorithm is established for this class of problems. A slight modification of the algorithm can solve a class of copositivity problems in which the matrix has exactly two negative eigenvalues.

In Chapter 5, we generalize the techniques of Chapter 2 to a class of general nonconvex programming problems. Here the objective function need not be quadratic and the constraints need not be linear.

Finally, in Chapter 6, we summarize the contribution of the thesis and outline further research directions.

1.2 Parametric Quadratic Programming

In this section, we review some known results and solution methods for parametric programming problems.

We consider the following model problem

$$\left.\begin{array}{l} \min\ c(t)'x + \frac{1}{2}x'C(t)x \\ \text{subject to } a_i(t)'x \le b_i(t),\ i = 1, \cdots, m \end{array}\right\} \tag{1.2}$$

for $t \in T$, where $c(t), a_i(t) \in \mathrm{E}^n$, $b_i(t) \in \mathrm{E}^1$ for $i = 1, \cdots, m$, $C(t)$ is an (n, n) symmetric positive semidefinite matrix for each $t \in T$ and T is a subset of E^k. In the majority of cases, $C(t)$ and $a_i(t)$ for $i = 1, \cdots, m$ will be independent of t. We assume that $c(t)$, $a_i(t)$, $b_i(t)$ and $C(t)$ are continuous on T. Let $A(t) = (a_1(t), \cdots, a_m(t))'$, $b(t) = (b_1(t), \cdots, b_m(t))'$, $R(t) = \{x \in R^n \mid A(t)x \le b(t)$ for some $x \in \mathrm{E}^n\}$ and define a real valued function as follows:

$$f(t) = \begin{cases} \inf\{c(t)'x + \frac{1}{2}x'C(t)x \mid A(t)x \le b(t)\}, & \text{if } R(t) \ne \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

The following result gives a sufficient condition for f to be lower semicontinuous.

Theorem 1.2.1 *For any $t^* \in T$, if the following system*

$$A(t^*)s \le 0,$$
$$c(t^*)'s \le 0,$$
$$C(t^*)s = 0$$

has no nonzero solutions, then f is lower semicontinuous at t^* relative to T.

Proof. See Best and Ding [5].

There are some sufficient conditions for f to be continuous. However such results will not be needed in this thesis. So, we will not discuss them further.

Associated with (1.2) are the Karush-Kuhn-Tucker conditions

$$\left.\begin{array}{l} a_i(t)'x(t) \leq b_i(t),\ i = 1,\cdots,m, \\ -c(t) - C(t)x(t) = \sum_{i=1}^m u_i(t)a_i(t), \\ u_i(t) \geq 0,\ i = 1,\cdots,m, \\ u_i(t)(a_i(t)'x(t) - b_i(t)) = 0,\ i = 1,\cdots,m. \end{array}\right\} \tag{1.3}$$

The multiplier, or dual variable, associated with constraint i is $u_i(t)$. Theorems 1.2.2–1.2.5 and Corollary 1.2.1 are the well known Karush-Kuhn-Tucker conditions (see [19]). The importance of the Karush-Kuhn-Tucker conditions is demonstrated in the following theorem.

Theorem 1.2.2 *(Necessary and Sufficient Conditions). For each $t \in T$, n-vector $x(t)$ is an optimal solution for (1.2) if and only if there exist numbers $u_1(t),\cdots,u_m(t)$ which, together with $x(t)$, satisfy the Karush-Kuhn-Tucker conditions (1.3).*

Related to the primal problem (1.2) is the dual problem defined by

$$\left.\begin{array}{l} \max\ Q(x,t) - \sum_{i=1}^m u_i(b_i(t) - a_i(t)'x) \\ \text{subject to } -c(t) - C(t)x = \sum_{i=1}^m u_i a_i(t), \\ \qquad\qquad u_i \geq 0,\ i = 1,\cdots,m, \end{array}\right\} \tag{1.4}$$

where

$$Q(x,t) = c(t)'x + \frac{1}{2}x'C(t)x.$$

We will use u to denote the m-vector $(u_1,\cdots,u_m)'$ and $u(t)$ to denote m-vector $(u_1(t),\cdots,u_m(t))'$. The following theorems indicate the relationship between (1.2) and (1.4).

Theorem 1.2.3 *(Weak Duality). For each $t \in T$, if $x_1(t)$ is feasible for the primal and $(x_2(t),u(t))$ is feasible for the dual, then*

$$Q(x_1(t),t) \geq Q(x_2(t),t) - \sum_{i=1}^m u_i(t)(b_i(t) - a_i(t)'x_2(t)).$$

Corollary 1.2.1 *For each $t \in T$, (1.2) has an optimal solution if and only if both (1.2) and (1.4) are feasible.*

Let us define a set P as follows:

$$P = \left\{ t \in T \;\middle|\; \begin{array}{l} a_i(t)'x \leq b_i(t),\ i = 1, \cdots, m, \\ -c(t) - C(t)x = \sum_{i=1}^{m} u_i a_i(t), \\ u_i \geq 0 \text{ for some } x \text{ and } u_i,\ i = 1, \cdots, m \end{array} \right\}.$$

Corollary 1.2.1 implies that (1.2) has an optimal solution if and only if $t \in P$.

Theorem 1.2.4 *(Strong Duality). For each $t \in T$, if $x(t)$ is an optimal solution for (1.2), then there exists an m-vector $u(t)$ such that $(x(t), u(t))$ is optimal for (1.4) and*

$$Q(x(t), t) = Q(x(t), t) - \sum_{i=1}^{m} u_i(t)(b_i(t) - a_i(t)'x(t)).$$

Theorem 1.2.5 *(Complementary Slackness). For each $t \in T$, let $x(t)$ and $(x(t), u(t))$ be feasible for (1.2) and (1.4), respectively. Necessary and sufficient conditions for simultaneous optimality of $x(t)$ and $(x(t), u(t))$ are*

$$\begin{array}{l} u_i(t) > 0 \textit{ implies } a_i(t)'x(t) = b_i(t), \textit{ for } i = 1, \cdots, m, \\ a_i(t)'x(t) < b_i(t) \textit{ implies } u_i(t) = 0 \textit{ for } i = 1, \cdots, m. \end{array}$$

Let I be any subset of $\{1, \cdots, m\}$ and define a subset $P(I)$ as

$$P(I) = \left\{ t \in T \;\middle|\; \begin{array}{l} a_i(t)'x \leq b_i(t),\ i = 1, \cdots, m, \\ -c(t) - C(t)x = \sum_{i \in I} u_i a_i(t), \\ u_i \geq 0 \text{ for } i \in I, \\ a_i(t)'x = b_i(t) \text{ for } i \in I \text{ for some } x \text{ and } u_i \end{array} \right\}.$$

By defining $u_i = 0$ for $i \notin I$ and Theorem 1.2.5, we can see that (1.2) has an optimal solution for any $t \in P(I)$. The reason we define $P(I)$ is that it may be possible to write down the explicit expression for $x(t)$ when $t \in P(I)$. For example, if

$$\begin{bmatrix} C(t) & A_I(t)' \\ A_I(t) & 0 \end{bmatrix}$$

is nonsingular for any $t \in \mathrm{E}^k$, then

$$\begin{bmatrix} x(t) \\ u_i(t) \end{bmatrix} = \begin{bmatrix} C(t) & A_I(t)' \\ A_I(t) & 0 \end{bmatrix}^{-1} \begin{bmatrix} -c(t) \\ b(t) \end{bmatrix}$$

for any $t \in P(I)$, where $A_I'(t)$ is the matrix whose columns are the gradients of those constraints in I, and $u_I(t)$ is the vector whose components are those $u_i(t)$ associated with the columns of $A_I'(t)$. From the definition of $P(I)$, we have

$$P = \bigcup\{P(I) \mid I \subset \{1, \cdots, m\}\}. \tag{1.5}$$

Since $\{1, \cdots, m\}$ has a finite number of subsets, P is a union of the finite number of subsets. For any $I_1, I_2 \subset \{1, \cdots, m\}$, it is possible to have $P(I_1) \subset P(I_2)$ and $P(I_1) \neq P(I_2)$. So, the right hand side of (1.5) may not be a partition of P. A main idea of the numerical procedures for parametric programming problems is to efficiently partition P into finite number of regions $P(I)$ and get the explicit expression for the optimal solution on each of these regions.

We illustrate these ideas by taking $C(t) \equiv C$, $c(t) \equiv c + tq$, $a_i(t) \equiv a_i$ and $b_i(t) \equiv b_i + tp_i$ for $i = 1, \cdots, m$ and $T \equiv [\underline{t}, \bar{t}] \subset \mathrm{E}^1$. That is, we consider the following parametric quadratic programming problem

$$\left.\begin{array}{l} \min\ (c + tq)'x + \frac{1}{2}x'Cx \\ \text{subject to } a_i'x \leq b_i + tp_i,\ i = 1, \cdots, m \end{array}\right\} \tag{1.6}$$

for $T = [\underline{t}, \bar{t}] \subset \mathrm{E}^1$, where $c, q, a_i \in \mathrm{E}^n$, $b_i, p_i \in \mathrm{E}^1$ for $i = 1, \cdots, m$; C is an (n, n) symmetric positive semidefinite matrix, $\underline{t}$ is finite or $-\infty$ and $\bar{t}$ is finite or $+\infty$. Best [3] has developed a numerical method for (1.6). An application of Best's method to (1.6) will produce numbers $t_0, t_1, \cdots, t_\nu$ and n-vectors h_{0i}, h_{1i}, $i = 1, \cdots, \nu$ satisfying

$$x_i(t) = h_{0i} + th_{1i}$$

is optimal for (1.6) for all t with $t_{i-1} \leq t \leq t_i$ and for all $i = 1, \cdots, \nu$. It is possible to have $t_0 = \underline{t}$ and/or $t_\nu = \bar{t}$. If $t_0 > \underline{t}$, the method will conclude that (1.6) is either unbounded from below or infeasible for $t < t_0$, and, the relevant possibility will be given. Similarly, if $t_\nu > \bar{t}$, then the method will conclude that (1.6) is either unbounded from below or has no feasible solution for $t > t_\nu$ and the relevant possibility will be stated.

The results and ideas for parametric quadratic programming problems discussed in this section will be used in Chapters 2, 3 and 5.

Chapter 2

Global and Local Quadratic Minimization

2.1 Introduction

Here we consider the model non-convex quadratic programming problem

QP $$\min\{c'x + \frac{1}{2}x'Cx + x'DQ'x \mid Ax \le b\},$$

where $c \in \mathrm{E}^n$, $b \in \mathrm{E}^m$, A is an (m, n) matrix, D and Q are (n, k) matrices, C is a symmetric (n, n) positive semidefinite matrix, $k < n$ and $x \in \mathrm{E}^n$ is a variable. Corresponding to QP, we consider the parametric quadratic program:

QP(t) $$\min\{c'x + \frac{1}{2}x'Cx + t'Q'x \mid Ax \le b, D'x = t\},$$

where t is a parameter in E^k. Let R and $R(t)$ be the feasible regions for QP and QP(t), respectively. Let arg min{QP(t)} denote the set of all optimal solutions for QP(t). Finally, we formulate

MQP $$\min\{f(t) \mid t \in \mathrm{E}^k\},$$

where

$$f(t) = \begin{cases} \inf \{c'x + \frac{1}{2}x'Cx + t'Q'x \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

The non-convexity of the objective function of QP stems from the term $x'DQ'x$. One might question the generality of this model and suggest that the term be written as $x'Hx$, where H is a more general symmetric matrix, perhaps satisfying some properties. This situation has been analyzed in Chapter 3, where it is shown that for any symmetric matrix H having full rank, there exist (n, k) matrices Q and D satisfying $H = \frac{1}{2}[DQ' + QD']$ (and so $x'Hx = x'DQ'x$) if and only if H has at least two nonzero eigenvalues of opposite sign. In addition, when the required condition is satisfied a method to construct such D and Q is given. For the purposes of this chapter, we will assume that D and Q are already available.

We note that, in general, the problem of checking isolated local optimality is NP-hard, See Murty and Kabadi [22], and, Pardalos and Schnitger [24].

We will organize this chapter as follows. In Section 2.2, we will develop the relationships between QP, QP(t) and MQP. In particular, we will establish the one to one correspondence between isolated local minimizers of QP and MQP. In Section 2.3, we will specialize these results to the class of non-convex quadratic programs with a Hessian which has exactly one negative eigenvalue. We will give an algorithm that can not only find a global minimizer, but can also find all isolated minimizers and some non-isolated local minimizers.

2.2 The Relationships Between QP and MQP

We begin this section with a small example problem which will illustrate the critical relationship between QP and MQP.

Example 2.2.1

$$\text{QP} \qquad \begin{array}{rl} \text{minimize :} & x_1 x_2 \\ \text{subject to :} & x_1 \geq 0.5, \quad 22x_1 + 8x_2 \geq 27, \\ & 8x_1 + 22x_2 \geq 27, \quad x_2 \geq 0.5. \end{array}$$

Here, $C = 0$, $n = 2$ and we may take $D = (1,0)'$ and $Q = (0,1)'$. QP(t) can be written as

$$\text{QP(t)} \qquad \begin{array}{rl} \text{minimize :} & tx_2 \\ \text{subject to :} & x_1 \geq 0.5, \quad 22x_1 + 8x_2 \geq 27, \\ & 8x_1 + 22x_2 \geq 27, \quad x_2 \geq 0.5, \\ & x_1 = t. \end{array}$$

The solution of QP(t) is a piece-wise linear function of t and is summarized in Table 2.1. Examination of QP(t) with Table 2.1 gives $f(t)$:

$$f(t) = \begin{cases} t(27 - 22t)/8, & \text{if } 0.5 \leq t \leq 0.9, \\ t(27 - 8t)/22, & \text{if } 0.9 \leq t \leq 2, \\ t/2, & \text{if } t \geq 2. \end{cases}$$

Example 2.2.1 is illustrated in Figures 2.1(a) and (b). Figure 2.1(a) shows the given non-convex problem. The feasible region is shown as the shaded area. The level set $x_1 x_2 = 0.81$ is shown with a broken line. It is clear from the figure that there are local minima at $(0.5, 2)'$ and $(2, 0.5)'$ and the global minimum occurs at $(.9, .9)'$. Figure 2.1(b) shows $f(t)$, a piece-wise quadratic function which by inspection, has isolated local minima at $t = .5$, and 2 and a global minimum at $t = .9$. Using Table 2.1, we see that arg min$\{$QP(.5)$\}$ = $(0.5, 2)'$, arg min$\{$QP(2)$\}$ = $(2, .5)'$ and arg min$\{$QP(.9)$\}$ = $(.9, .9)'$. Thus, the local (global) minima of QP and $f(t)$ are in one to one correspondence for this example. Also note that $f(t)$ is a piece-wise

Table 2.1: Optimal Solution for QP(t) for Example 2.2.1

$t < 0.5$	$0.5 \le t \le 0.9$	$0.9 \le t \le 2.$	$t \ge 2.$
no feasible solution	$\begin{bmatrix} 0 \\ \frac{27}{8} \end{bmatrix} + t \begin{bmatrix} 1 \\ -\frac{22}{8} \end{bmatrix}$	$\begin{bmatrix} 0 \\ \frac{27}{22} \end{bmatrix} + t \begin{bmatrix} 1 \\ -\frac{8}{22} \end{bmatrix}$	$\begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

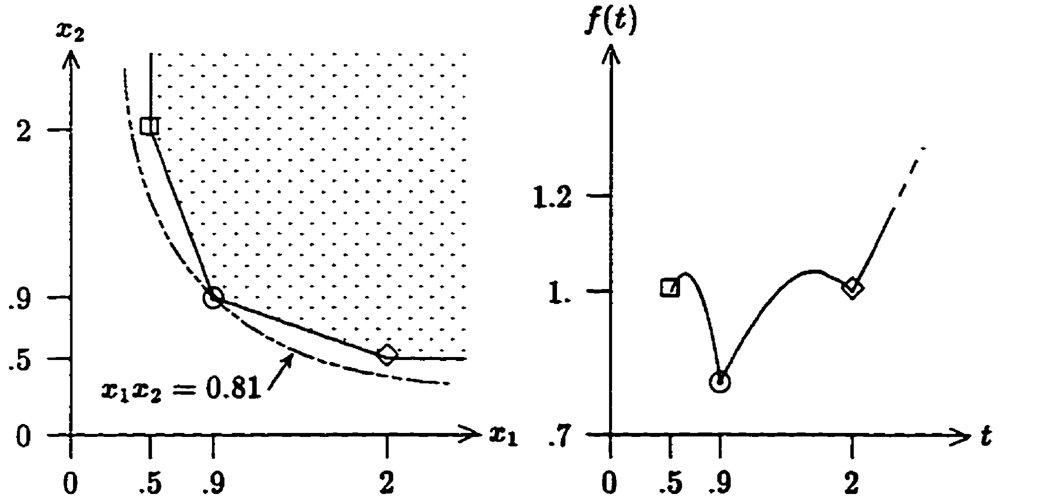

Figure 2.1: (a) Example 2.2.1 (b) $f(t)$ for Example 2.2.1

quadratic function of a single variable and so it is straightforward to obtain its local and global minimizers.

Notice that in Example 2.2.1, the local minimizers for both QP and MQP are isolated. The requirement that the local minimizers of QP be isolated is key in obtaining the one to one correspondence between such points of QP and MQP. The final result will be formulated in Theorem 2.2.4 and will be a consequence of Theorems 2.2.1–2.2.3, following.

Theorem 2.2.1 *Let t^* be a local minimizer for MQP with $f(t^*) > -\infty$. Then any $x^* \in \arg\min\{\mathrm{QP}(t^*)\}$ is a local minimizer for QP.*

Proof: Since t^* is a local minimizer for MQP, there exists a $\delta > 0$ such that

$$f(t) \geq f(t^*) \quad \text{for any } t \in B_\delta(t^*), \tag{2.1}$$

where $B_\delta(t^*) = \{t \in \mathrm{E}^k \mid \|t - t^*\| \leq \delta\}$. Now assume to the contrary, that there is an $x^* \in \arg\min\{\mathrm{QP}(t^*)\}$ which is not a local minimizer for QP. Then there exists a sequence $\{x^i\}$ such that

$$c'x^i + \frac{1}{2}(x^i)'Cx^i + (x^i)'DQ'x^i < c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^*,$$

where $Ax^i \leq b$, and $x^i \to x^*$. Since $x^* \in \arg\min\{\mathrm{QP}(t^*)\}$,

$$f(t^*) = c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^*.$$

Hence

$$c'x^i + \frac{1}{2}(x^i)'Cx^i + (x^i)'DQ'x^i < f(t^*). \tag{2.2}$$

Since $x^i \to x^*$, $D'x^i \to D'x^* = t^*$. Thus, there is an $M > 0$ such that

$$D'x^i \in B_\delta(t^*), \quad \text{whenever } i > M. \tag{2.3}$$

Let $t^i = D'x^i$ for $i > M$. Then

$$\begin{aligned} f(t^i) &= \inf\{c'x + \frac{1}{2}x'Cx + (t^i)'Q'x \mid Ax \le b, D'x = t^i\} \\ &\le c'x^i + \frac{1}{2}(x^i)'Cx^i + (t^i)'Q'x^i \qquad \text{(since } Ax^i \le b, D'x^i = t^i) \\ &= c'x^i + \frac{1}{2}(x^i)'Cx^i + (x^i)'DQ'x^i < f(t^*) \qquad \text{(from (2.2))}. \end{aligned}$$

But from (2.1) and (2.3), we have $f(t^i) \ge f(t^*)$, a contradiction. The assumption that there is an $x^* \in \arg\min\{\mathrm{QP}(t^*)\}$ which is not a local minimizer for QP leads to a contradiction and is therefore false. The proof of the theorem is thus complete. □

If t^* is an isolated local minimizer of f on E^k, we have the following further result.

Theorem 2.2.2 *If t^* is an isolated local minimizer for f on E^k with $f(t^*) > -\infty$ and* arg min$\{\mathrm{QP}(t^*)\}$ *is the singleton point $\{x^*\}$, then x^* is an isolated local minimizer for QP.*

Proof: From Theorem 2.2.1, x^* is a local minimizer for QP. If x^* is not an isolated local minimizer for QP, there exists a sequence $\{x^i\} \subset R$, $x^i \to x^*$ and $x^i \ne x^*$ for all i such that $c'x^i + \frac{1}{2}(x^i)'Cx^i + (x^i)'DQ'x^i = c'x^* + \frac{1}{2}(x^*)'Cx^* + (x^*)'DQ'x^*$. Let $t^i = D'x^i$. Then $t^i \to t^*$ and $f(t^i) = f(t^*)$. Since $\arg\min\{\mathrm{QP}(t^*)\} = \{x^*\}$ and $x^i \ne x^*$ for all i, $t^i \ne t^*$ for all i. This contradicts that t^* is an isolated local minimizer for MQP. The proof of the theorem is thus complete. □

Theorem 2.2.2 is illustrated in Example 2.2.1 where each of the three local minimizers for $f(t)$ are isolated, their corresponding sets, arg min$\{\mathrm{QP}(t)\}$, are singletons and each such point is an isolated local minimizer for QP. The following example illustrates Theorem 2.2.1 and in addition, shows that the condition arg min$\{\mathrm{QP}(t^*)\}$ be a single point cannot, in general, be removed from Theorem 2.2.2.

Example 2.2.2

QP $$\min\{-\frac{1}{4}x_1 - \frac{1}{2}x_2 + x_1x_2 \mid x_1 \geq \frac{1}{2}, \frac{3}{8} \leq x_2 \leq \frac{1}{2}\}.$$

Here we take $C = 0$, $n = 2$, $D = (1,0)'$ and $Q = (0,1)'$. Then $D'x = x_1$, $Q'x = x_2$ and QP(t) becomes

QP(t) $$\min\{-\frac{1}{4}t + (t - \frac{1}{2})x_2) \mid t \geq \frac{1}{2}, \frac{3}{8} \leq x_2 \leq \frac{1}{2}\},$$

from which $f(t)$ is derived as:

$$f(t) = \begin{cases} \frac{1}{8}t - \frac{3}{16}, & \text{if } t \geq \frac{1}{2}, \\ +\infty, & \text{otherwise.} \end{cases}$$

The situation is illustrated in Figures 2.2(a) and 2.2(b). The feasible region for QP is shown as the shaded area in Figure 2.2(a). It is clear that $t^* = \frac{1}{2}$ is an isolated local minimizer for f on E^1. Indeed, it is also the global minimizer. See Figure 2.2(b). However, arg min{QP(t^*)} $= \{(\frac{1}{2}, x_2)' \mid \frac{3}{8} \leq x_2 \leq \frac{1}{2}\}$ and by Theorem 2.2.1, each one of these points is a local (indeed, global) minimizer for QP. These are shown by the darkened line in Figure 2.2(a). Clearly, none of the local minimizers for QP is isolated. Thus the condition arg min{QP(t^*)} be a singleton is necessary in Theorem 2.2.2.

Example 2.2.3

QP $$\min\{-\frac{1}{4}x_1 - \frac{1}{2}x_2 + x_1x_2 \mid x_1 \geq \frac{1}{2}, 0 \leq x_2 \leq \frac{1}{2}\}.$$

Here we take $C = 0$, $n = 2$, $D = (1,0)'$ and $Q = (0,1)'$, Then $D'x = x_1$, $Q'x = x_2$ and QP(t) becomes

QP(t) $$\min\{-\frac{1}{4}t + (t - \frac{1}{2})x_2 \mid t \geq \frac{1}{2}, 0 \leq x_2 \leq \frac{1}{2}\},$$

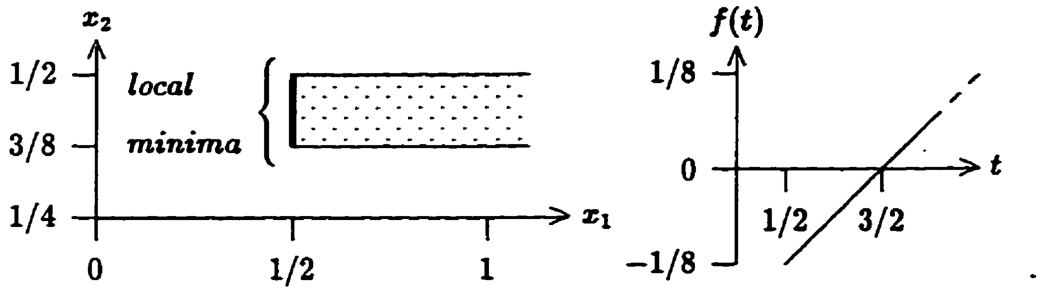

Figure 2.2: (a) Example 2.2.2 (b) $f(t)$ for Example 2.2.2

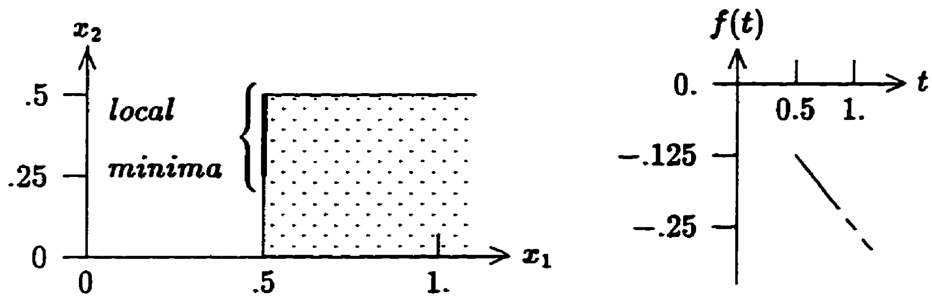

Figure 2.3: (a) Example 2.2.3 (b) $f(t)$ for Example 2.2.3

from which $f(t)$ is derived as:

$$f(t) = \begin{cases} -\frac{1}{4}t, & \text{if } t \geq \frac{1}{2}, \\ +\infty, & \text{otherwise.} \end{cases}$$

The situation is illustrated in Figures 2.3(a) and 2.3(b).

Observe that

$$\arg\min\{\mathrm{QP}(\frac{1}{2})\} \;=\; \{\,(\frac{1}{2}, x_2)' \mid 0 \leq x_2 \leq \frac{1}{2}\},$$

and for $t > \frac{1}{2}$, $\arg\min\{\mathrm{QP}(t)\} \;=\; (t, 0)'$. Also observe that $\{(\frac{1}{2}, x_2)' \mid \frac{1}{4} < x_2 \leq \frac{1}{2}\}$

are local optimal solutions for QP. In particular, $x^* = (\frac{1}{2}, \frac{1}{2})'$ is a local optimizer for QP but $t^* = D'x^* = \frac{1}{2}$ is *not* a local minimizer for f on E^1. Indeed f does not possess a local minimizer on E^1.

Example 2.2.3 shows that a one to one correspondence between local minima of QP and $f(t)$ will not hold without some restrictions. The key requirement in establishing the correspondence is that corresponding local minimizers for QP, QP(t) and $f(t)$ should each be isolated. This will be established subsequently. First we need the following lemma.

Lemma 2.2.1 *Let x^* be an isolated local minimizer for QP and let $t^* = D'x^*$. Let $\{t^i\}$ be any sequence with $t^i \to t^*$ and let $x^i \in \arg\min\{\mathrm{QP}(t^i)\}$. If there exists an $M > 0$ such that $f(t^i) \le M$ for all i, then $\{x^i\}$ is bounded.*

Proof: Since x^* is an isolated local minimizer for QP, there exists a $\delta > 0$ such that

$$c'x^* + \frac{1}{2}(x^*)'Cx^* + (x^*)'DQ'x^* < c'x + \frac{1}{2}x'Cx + x'DQ'x \tag{2.4}$$

for any $x \in (B_\delta(x^*) \bigcap R) \setminus \{x^*\}$. So,

$$c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^* < c'x + \frac{1}{2}x'Cx + (t^*)'Q'x$$

for any $x \in (B_\delta(x^*) \bigcap R(t^*)) \setminus \{x^*\}$. Since $c'x + \frac{1}{2}x'Cx + (t^*)'Q'x$ is convex, we have

$$c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^* < c'x + \frac{1}{2}x'Cx + (t^*)'Q'x \tag{2.5}$$

for any $x \in R(t^*) \setminus \{x^*\}$. Hence $\arg\min\{\mathrm{QP}(t^*)\} = \{x^*\}$. This implies that there does not exist a nonzero vector s satisfying the following conditions

$$As \le 0,\ D's = 0,\ (c + Qt^*)'s \le 0,\ Cs = 0, \tag{2.6}$$

otherwise $x^* + s$ will also be an optimal solution of QP(t^*) which contradicts arg min{QP(t^*)} $= \{x^*\}$. By Theorem 1.2.1, f is lower semi-continuous at t^*. Therefore for any $\gamma > 0$ there exists an $\epsilon > 0$ such that $f(t) \geq f(t^*) - \gamma$ for any $t \in B_\epsilon(t^*)$. Since $t^i \to t^*$, there exists an $N > 0$ such that $t^i \in B_\epsilon(t^*)$ for all $i \geq N$. So,

$$f(t^*) - \gamma \leq f(t^i) \leq M \tag{2.7}$$

for all $i \geq N$. Now assume that on the contrary, $\{x^i\}$ is unbounded, then $\{x^i/\|x^i\|\}$ has a convergent subsequence. Without loss of generality, let

$$\lim_{i\to\infty} \frac{x^i}{\|x^i\|} = s \text{ and } \lim_{i\to\infty} \|x^i\| = +\infty. \tag{2.8}$$

From $f(t^i) = c'x^i + \frac{1}{2}(x^i)'Cx^i + (t^i)'Q'x^i$, (2.7) and (2.8), we have

$$\frac{1}{2}s'Cs = \lim_{i\to\infty} \frac{f(t^i)}{\|x^i\|^2} = 0$$

and

$$(c + Qt^*)'s + \lim_{i\to\infty} \frac{(x^i)'Cx^i}{2\|x^i\|} = \lim_{i\to\infty} \frac{f(t^i)}{\|x^i\|} = 0,$$

i.e.;

$$Cs = 0 \text{ and } (c + Qt^*)'s = -\lim_{i\to\infty} \frac{(x^i)'Cx^i}{2\|x^i\|} \leq 0.$$

From $Ax^i \leq b$ and $D'x^i = t^i$, we have $As \leq 0$ and $D's = 0$. Thus we have exhibited a non-zero s satisfying (2.6). This is a contradiction and the proof of the lemma is complete. □

Theorem 2.2.3 *If x^* is an isolated local minimizer for QP, then $t^* = D'x^*$ is an isolated local minimizer for MQP, $f(t^*) = c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^*$ and arg min $\{QP(t^*)\} = \{x^*\}$.*

Proof: As the proof of Lemma 2.2.1, we have (2.4), (2.5) and $f(t) \geq f(t^*) - \gamma$ for any $t \in B_\epsilon(t^*)$. Thus,

$$f(t^*) = c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^* \text{ and } \arg\min\{\mathrm{QP}(t^*)\} = \{x^*\}.$$

Now assume to the contrary, that t^* is not an isolated local minimizer for MQP. Then there exist two sequences $\{x^i\}$ and $\{t^i\}$ with $t^i = D'x^i$, $t^i \to t^*$, $t^i \in B_\epsilon(t^*)$ and $x^i \in R \setminus B_\delta(x^*)$ such that

$$c'x^i + \frac{1}{2}(x^i)'Cx^i + (t^i)'Q'x^i = f(t^i) \leq f(t^*).$$

By Lemma 2.2.1, $\{x^i\}$ is bounded, so there exists a convergent subsequence. Without loss of generality, let $x^i \to x^0$. Then $x^0 \neq x^*$, $t^* = D'x^0$, $Ax^0 \leq b$ and

$$c'x^0 + \frac{1}{2}(x^0)'Cx^0 + (t^*)'Q'x^0 \leq f(t^*).$$

This contradicts (2.5). The proof of the theorem is thus complete. □

Combining Theorem 2.2.2 and Theorem 2.2.3, we have the following result.

Theorem 2.2.4 *A point x^* is an isolated local minimizer for QP if and only if $t^* = D'x^*$ is an isolated local minimizer of MQP, $f(t^*) = c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Q'x^*$ and* $\arg\min\{\mathrm{QP}(t^*)\} = \{x^*\}$.

Remark 2.2.1 From Theorem 2.2.1 and Theorem 2.2.4, we know that f will keep all of the critical information concerning isolated local minimizers of QP and some of the local minimizers of QP. Thus if we can locate all local minimizers of f we will obtain all isolated local minimizers and some local minimizers of QP.

Although the one to one correspondence between local minimizers of QP and MQP requires the condition of isolated local minima, this condition is not required for global minima as given in Theorem 2.2.5 below. The proof of this result can be obtained from definitions directly.

Theorem 2.2.5 *A point $t^* \in E^k$ with $f(t^*) > -\infty$ is a global minimizer of MQP if and only if QP has a global minimizer x^* such that $D'x^* = t^*$, $f(t^*) = c'x^* + \frac{1}{2}(x^*)'Cx^* + (t^*)'Qx^*$.*

We complete this section by showing how to recognize whether a local minimizer is an isolated local minimizer. Suppose that we know t^* is an isolated local minimizer and we want to know whether corresponding point x^* is also an isolated local minimizer. In doing so, we need only verify that $\arg\min\{\text{QP}(t^*)\} = \{x^*\}$. If C is positive definite, then QP(t) is strictly convex. In this case, x^* is necessarily uniquely determined and consequently $\arg\min\{\text{QP}(t^*)\} = \{x^*\}$. Otherwise, since QP$(t^*)$ is convex, we may assume that x^* is computed by some quadratic programming algorithm and $(u, v) = (u_1, \cdots, u_m, v_1, \cdots, v_k)$ is the associated vector of multipliers, where u and v correspond $Ax \leq b$ and $D'x = t^*$, respectively. Now by Theorem 4.14 of Best and Ritter [9], $\arg\min\{\text{QP}(t^*)\}$ can be represented by the set of x which satisfy

$$\left.\begin{array}{l} D'x = t^*,\ Cx = Cx^*, \\ a_i'x = b_i,\ \text{for all } i \text{ with } 1 \leq i \leq m,\ \text{and } u_i > 0, \\ a_i'x \leq b_i,\ \text{for all } i \text{ with } 1 \leq i \leq m,\ \text{and } u_i = 0. \end{array}\right\} \tag{2.9}$$

Let $I = \{i \mid 1 \leq i \leq m,\ a_i'x^* = b_i\}$ and A_I be a submatrix of A induced by i-th row of A for $i \in I$. If $\text{rank}([D, C, A_I']) < n$, then $\arg\min\{\text{QP}(t^*)\}$ is not a singleton, i.e.; x^* is not an isolated local minimizer. In this case, an alternative local minimizer can be computed easily from the null space of $[D, C, A_I']'$. In fact, for any $y \in E^n$, if $[D, C, A_I']'y = 0$ with $y \neq 0$, there is a nonzero number α such that $x^* + \alpha y$ satisfies (2.9), i.e.; $x^* + \alpha y$ is an alternate local minimizer. If $\text{rank}([D, C, A_I']) = n$, we need to consider the following linear programming problem:

$$\mu = \min\{\sum_{i \in I} a_i'x \mid (2.9)\}.$$

If $\mu = \sum_{i\in I} b_i$ then $\arg\min\{\mathrm{QP}(t^*)\}$ is a singleton, and x^* is an isolated local minimizer. Otherwise, x^* is not an isolated local minimizer and an optimal solution of the linear programming problem is an alternative local minimizer.

In next section, we are going to discuss some applications of the results established in this section.

2.3 The Case of a Single Negative Eigenvalue

In this section we consider our model problem with D and Q being n-dimensional *vectors*; i.e., $k = 1$. To emphasize this we replace D and Q with d and q, respectively. The model problem QP becomes

$$\mathrm{QP}_1 \qquad \min\{c'x + \frac{1}{2}x'Cx + (d'x)(q'x) \mid Ax \leq b\},$$

where $c \in \mathrm{E}^n$, $b \in \mathrm{E}^m$, A is an (m,n) matrix, d and $q \in \mathrm{E}^n$, C is a symmetric (n,n) positive semidefinite matrix, and $x \in \mathrm{E}^n$. $\mathrm{QP}(t)$ becomes

$$\mathrm{QP}_1(t) \qquad \min\{c'x + \frac{1}{2}x'Cx + tq'x \mid Ax \leq b, d'x = t\},$$

where t is a scalar parameter. Let R_1 and $R_1(t)$ be feasible regions for QP_1 and $\mathrm{QP}_1(t)$, respectively. Let $\arg\min\{\mathrm{QP}_1(t)\}$ denote the set of all optimal solutions for $\mathrm{QP}_1(t)$. Finally, we formulate

$$\mathrm{MQP}_1 \qquad \min\{f_1(t) \mid t \in \mathrm{E}^1\},$$

where

$$f_1(t) = \begin{cases} \inf\,\{c'x + \frac{1}{2}x'Cx + tq'x \mid x \in R_1(t)\}, & \text{if } R_1(t) \neq \phi, \\ +\infty, & \text{otherwise,} \end{cases}$$

and we have used the subscript "1" throughout the above to emphasize that $k = 1$.

If QP_1 were written with a more general Hessian matrix H, rather than $C + \frac{1}{2}(dq' + qd')$ then the resulting problem could be transformed into one having a Hessian matrix of the latter form provided H had exactly one negative eigenvalue (hence the title of this section). Details of this transformation are given in the next chapter.

The problem QP_1 has been shown to be NP-hard by Pardalos and Vavasis [26]. Konno *et al.* [17] proposed a solution method for a variation of QP_1 for which the constraints were equalities and non-negativity constraints. The method used a parametric form of the simplex algorithm and was designed solely to find a global minimizer. In this section, we will also develop a method to solve QP_1. However, in contrast to the method of Konno, our method will not only locate a global minimizer (if one exists), but also all isolated local minimizers and some non-isolated local minimizers. Indeed, even if QP_1 does not possess a global minimizer our method will locate all isolated local minimizers and some non-isolated local minimizers. Moreover, we will show that isolated local minimizers can be distinguished from non-isolated local minimizers by solving a linear programming problem.

Note that $\mathrm{QP}_1(t)$ is a convex parametric quadratic programming problem, with the parameter being a scalar. Note also that the parameter t occurs in both the linear part of the objective function as well as the right hand-side of a constraint. $\mathrm{QP}_1(t)$ must be solved for all possible t. An appropriate method to use is that of Best [3]. Best's method allows explicitly for a parameter in both the linear part of the objective function and the right hand-side of the constraints. Also, it allows for the possibility that the Hessian of the parametric QP is positive semidefinite, rather than just positive definite. In addition, it supplies critical information as to the status of $\mathrm{QP}_1(t)$ at the end points of the parametric interval.

Applying Best's method to $\mathrm{QP}_1(t)$ will produce numbers $t_0, t_1, \cdots, t_\nu$ and n–vectors h_{0i}, h_{1i}, $i = 1, \ldots, \nu$ satisfying

$$x_i(t) \;=\; h_{0i} \;+\; th_{1i} \tag{2.10}$$

is optimal for $\mathrm{QP}_1(t)$ for all t with $t_{i-1} \le t \le t_i$ and for all $i = 1, \ldots \nu$. It is possible to have $t_0 = -\infty$ and/or $t_\nu = +\infty$. If $t_0 > -\infty$, Best's method will conclude that $\mathrm{QP}_1(t)$ is either unbounded from below or infeasible for $t < t_0$, and, the relevant possibility will be given. Similarly, if $t_\nu < \infty$, then the method will conclude that $\mathrm{QP}_1(t)$ is either unbounded from below or has no feasible solution for $t > t_\nu$ and the relevant possibility will be stated. Table 2.1 gives the relevant information for Example 2.2.1.

Having solved $\mathrm{QP}_1(t)$, it remains to solve MQP_1. Using h_{0i} and h_{1i} from (2.10), define the constants

$$\left.\begin{aligned} \gamma_{1i} &= c'h_{0i} + \tfrac{1}{2}h_{0i}{}'Ch_{0i}, \\ \gamma_{2i} &= c'h_{1i} + h_{0i}{}'Ch_{1i} + q'h_{0i}, \\ \gamma_{3i} &= \tfrac{1}{2}h_{1i}{}'Ch_{1i} + q'h_{1i}, \end{aligned}\right\} \tag{2.11}$$

for $i = 1, \ldots, \nu$. From (2.10), (2.11) and the definition of $f(t)$, we now have

$$f_1(t) \;=\; \begin{cases} \gamma_{11} + \gamma_{21}t + \gamma_{31}t^2, & \text{if } t_0 \le t \le t_1, \\ \gamma_{12} + \gamma_{22}t + \gamma_{32}t^2, & \text{if } t_1 \le t \le t_2, \\ \quad\cdots & \quad\cdots \\ \gamma_{1\nu} + \gamma_{2\nu}t + \gamma_{3\nu}t^2, & \text{if } t_{\nu-1} \le t \le t_\nu. \end{cases} \tag{2.12}$$

This shows that $f_1(t)$ is piece-wise quadratic on ν adjacent intervals. This is illustrated in Figure 2.1(b) with $\nu = 3, t_0 = .5,\ t_1 = .9,\ t_2 = 2.$, and $t_3 = \infty$. The simple nature of $f(t)$ allows the determination of its local minima as summarized in Table 2.2.

Case 1 concerns points where the left derivative of $f_1(t)$ is negative and the right derivative is positive. This possibility is illustrated in Figure 2.1(b) with $t_1 = .9$ and $t_2 = 2$. Case 2 corresponds to $f_1(t)$ being strictly convex on $[t_{i-1}, t_i]$ and the unconstrained minimum of that quadratic piece lying within the interval. Case 3 requires that t_0 be finite, $f(t)$ be increasing at t_0 and that there be no feasible solutions below t_0. This is illustrated in Figure 2.1(b) for $t_0 = .5$. Note that the relevant possibility will be given by Best's parametric QP method. Also note that if the QP algorithm determines that $\text{QP}_1(t)$ is unbounded from below for $t < t_0$, then t_0 is *not* a local minimizer for $f_1(t)$. For the right-hand end of the interval, Case 4 is analogous to Case 3. Case 5 occurs when $f(t)$ is constant on the open interval (t_{i-1}, t_i), in which case any point in the interval is a local minimizer. The end points of the interval may or may not be local minimizers. See the discussion following Theorem 2.3.1.

The following result is an immediate consequence of Theorem 2.2.1

Theorem 2.3.1 *Let $t_1^*, t_2^*, \cdots, t_N^*$ be obtained from (2.10), (2.11) and Table 2.2. Let $x_i^* \in \arg\min\{\text{QP}(t_i^*)\}$ for $i = 1, \cdots, N$. Then $x_i^*, i = 1, \cdots, N$ are all local minimizers of QP_1. Moreover, if QP_1 possesses a global minimizer, then it is that x_k^* which gives the smallest objective function value for QP_1 among all the $\{x_i^* \mid i = 1, \cdots, N\}$.*

The formulation of Theorem 2.3.1 does not explicitly allow for Case 5 of Table 2.2. because it deals with particular points rather than points and intervals. If Case 5 does apply, then $\arg\min\{\text{QP}(t)\}$ are all local minimizers of QP_1. If the left derivative of $f(t)$ is negative at t_{i-1} then t_{i-1} is also a local minimizer of $f(t)$ and consequently all elements of $\arg\min\{\text{QP}(t_{i-1})\}$ are local minimizers of QP_1. The analogous result holds for the right-hand end of the interval. The information

Table 2.2: Determination of Local Minimizers for $f_1(t)$

Case	Range of i	Conditions	Local Min of $f(t)$
1	$1 \leq i \leq \nu - 1$	$\gamma_{2i} + 2\gamma_{3i}t_i \leq 0$, and, $\gamma_{2,i+1} + 2\gamma_{3,i+1}t_i \geq 0$	t_i
2	$1 \leq i \leq \nu$	$\gamma_{3i} > 0$, and, $t_{i-1} < -\gamma_{2i}/(2\gamma_{3i}) < t_i$	$-\gamma_{2i}/(2\gamma_{3i})$
3		$t_0 > -\infty$, $\gamma_{20} + 2\gamma_{30}t_0 > 0$ and $R_1(t) = \emptyset$ for $t < t_0$	t_0
4		$t_\nu < \infty$, $\gamma_{2\nu} + 2\gamma_{3\nu}t_\nu < 0$ and $R_1(t) = \emptyset$ for $t > t_\nu$	t_ν
5	$1 \leq i \leq \nu$	$\gamma_{2i} = \gamma_{3i} = 0$	$t_{i-1} < t < t_i$

concerning whether QP_1 possesses a global minimizer can be obtained from Best's algorithm, (2.10), (2.11) and Table 2.2. This can be summarized as follows. When Best's algorithm terminates with a finite t_0, it also specifies that either $\text{QP}_1(t)$ is unbounded from below for $t < t_0$, or, $R_1(t) = \phi$ for $t < t_0$. The analogous result holds when t_ν is finite. Thus if $R_1(t) \neq \phi$ for $t < t_0$ with t_0 being finite or $R_1(t) \neq \phi$ for $t > t_\nu$ with t_ν being finite, then QP_1 has no global minimizer. Otherwise both

$$\min \{\gamma_{11} + \gamma_{21}t + \gamma_{31}t^2 \mid t_0 \leq t \leq t_1\}$$

and

$$\min \{\gamma_{1\nu} + \gamma_{2\nu}t + \gamma_{3\nu}t^2 \mid t_{\nu-1} \leq t \leq t_\nu\}$$

have global minimizers if and only if QP_1 has global minimizer.

We illustrate this procedure by applying it to an example from Floudas and Pardalos [14].

Example 2.3.1

$$\text{QP}_1 \qquad \begin{array}{rl} \text{minimize}: & 6.5x - 0.5x^2 - y_1 - 2y_2 - 3y_3 - 2y_4 - y_5 \\ \text{subject to}: & AX \leq b,\ 0 \leq X = (x,y)',\ y_i \leq 1,\ i = 3,4, \\ & y_5 \leq 2, x \in \mathrm{E}^1,\ y \in \mathrm{E}^5, \end{array}$$

where

$$b = \begin{bmatrix} 16 \\ -1 \\ 24 \\ 12 \\ 3 \end{bmatrix} \text{ and } A = \begin{bmatrix} 1 & 2 & 8 & 1 & 3 & 5 \\ -8 & -4 & -2 & 2 & 4 & -1 \\ 2 & 0.5 & 0.2 & -3 & -1 & -4 \\ 0.2 & 2 & 0.1 & -4 & 2 & 2 \\ -0.1 & -0.5 & 2 & 5 & -5 & 3 \end{bmatrix}.$$

This problem has a known global minimizer $(x^*, y^*) = (0, 6, 0, 1, 1, 0)'$ with optimal objective function value of -11.

An application of our algorithm to this problem confirms that the above solution is indeed the global optimum. In addition, it also determines that the global minimizer is isolated and that $(\bar{x}, \bar{y}) = (13.83, 0, 0, 1, 0.19, 0.12)'$ is an isolated local minimizer with the objective function value -9.26. That is, the problem not only has an isolated global minimizer, but also a previously undiscovered isolated local minimizer.

In order to further test our algorithm, we formulated some variations of this problem. In all cases, the constraints remained the same and only the linear part of the objective function was changed. The modified objective functions, $g_i(x, y)$, $i = 1, \ldots 4$ are shown in Table 2.3 along with their corresponding vectors d and q. The original Floudas and Pardalos problem corresponds to $g_1(x, y)$. Each of the

Table 2.3: Objective Functions for Test Problems

$$\begin{aligned}
g_1(x,y) &= -0.5x^2 + 6.5x - y_1 - 2y_2 - 3y_3 - 2y_4 - y_5 \\
g_2(x,y) &= -0.5x^2 + 6.5x - 2y_2 - 3y_3 - y_5 + 2xy_1 + 3xy_2 \\
&\quad -5xy_3 - 4xy_4 + 6xy_5 \\
g_3(x,y) &= -0.5x^2 + 6.5x - 4y_2 - 4y_3 - y_5 + 2xy_1 + 3xy_2 \\
&\quad -5xy_3 - 4xy_4 + 6xy_5 \\
g_4(x,y) &= -0.5x^2 + 6.5x + 2xy_1 + 3xy_2 + 5xy_3 - 4xy_4 + 6xy_5
\end{aligned}$$

Table 2.4: Local and Global Minima for Four Test Problems

Objective Function	Objective Value	Solution Points	Type of Minimum
$g_1(x,y)$	-11	$(0,6,0,1,1,0)'$	global min, isolated
	-9.2567	$(13.83,0,0,1,0.19,0.12)'$	local min, isolated
$g_2(x,y)$	-105	$(12,0,0,1,1,0)'$	global min, isolated
	-5.6583	$(0,0.92,1.33,1,0.84,0)'$	local min, isolated
	-5.0718	$(0.52,0,1.44,1,1,0)'$	local min, isolated
$g_3(x,y)$	-106	$(12,0,0,1,1,0)'$	global min, isolated
	-9.3166	$(0,0.92,1.33,1,0.84,0)'$	local min, isolated
	-8.9409	$(0.5,0,1.45,1,0.97,0)'$	local min, isolated
	-8.9428	$(0.52,0,1.44,1,1,0)'$	local min, isolated
$g_4(x,y)$	-46.875	$(12.5,0,0,0,1,0)'$	global min, isolated
	1.3672	$(0.625,0,0,0,1,0)'$	local min, isolated
	0	$(0,1.25,0,0,1,0)'$	local min, non-isolated
	0	$(0,7.6,0,0.8,0,0)'$	local min, non-isolated

four examples was solved with $d = (1,0,0,0,0,0)'$. The first example used $q = (-0.5,0,0,0,0,0)'$ and the remaining three used $q = (-0.5,2,3,-5,-4,6)'$. The results of applying our method to these problems are summarized in Table 2.4. Note that the results summarized in Table 2.4 show that our method located two non-isolated local minimizers for the fourth test problem. This shows that although we cannot guarantee that our method will find all non-isolated local minimizers, it still may find some, or even all.

2.4 Conclusions

We have developed relationships between a given non-convex quadratic programming problem QP and a derived unconstrained (but non-differentiable) quadratic problem MQP. We have established that any local minimum of MQP gives a corresponding local minimum of QP. Furthermore, the isolated local minimizers of both QP and MQP are in a one to one correspondence.

For the case that the Hessian of QP has exactly one negative eigenvalue, we have developed an algorithm to compute all isolated local minimizers and some non-isolated local minimizers of QP. In addition, the algorithm will compute the global minimizer of QP, provided it exists, and will provide the information that QP is unbounded from below when that is the case. The algorithm is illustrated by applying it to a problem from the literature and some variations of it.

Chapter 3

A Decomposition Procedure For Non-Convex QP

3.1 Introduction

In Chapter 2, we have developed a theory to find all isolated local minimizers and some non-isolated local minimizers for the non-convex QP

$$\min\{c'x + x'DQ'x \mid Ax \le b\} \tag{3.1}$$

by parametric quadratic programming, where D and Q are (n, k) matrices. The model problem used in Chapter 2 includes a convex quadratic term in the objective function for (3.1). However, it is not useful to include it here and we omit it. As in Chapter 2, we proceed by formulating the parametric LP

$$\min\{c'x + t'Q'x \mid Ax \le b,\ D'x = t\}, \tag{3.2}$$

where t is a parameter vector in E^k. Letting $R(t)$ denote the set of feasible solutions for (3.2), the derived problem for (3.1) is

$$\min\{f(t) \mid t \in E^k\}, \tag{3.3}$$

where

$$f(t) = \begin{cases} \inf\{c'x + t'Q'x \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is shown in Chapter 2 that the isolated local minimizers of (3.1) and (3.3) are in one to one correspondence. In particular, if t^* is a local minimizer for (3.3) then any optimal solution for (3.2) with $t = t^*$ is a local minimizer for (3.1).

Although the theory was developed for arbitrary $k \leq n$, the numerical procedures developed in Chapter 2 were limited to the case of $k = 1$. In this chapter, we address the problem of arbitrary k by using a decomposition approach. The method proposed here will begin with the model problem

$$\min\{c'x \;+\; x'Cx \mid a_i'x \;\leq\; b_i,\; i = 1, \ldots, m\},$$

where C is (n, n) and symmetric. We then give a method which will either construct matrices D and Q satisfying $x'DQ'x = x'Cx$ (so that the model problem is rewritten in the model form (3.1)), or, determine that no such matrices D and Q exist. In the former case, the decomposition method then generates m subproblems each of dimension $n - 1$, where m is the number of constraints in (3.1). Solution of all of these m subproblems gives a solution to the given problem. Each of these smaller problems is in turn decomposed into m subproblems with their dimension reduced by 1. The process continues by generating smaller and smaller dimensional subproblems until the subproblem can be solved directly. One possibility is a 1 dimensional subproblem which can be solved directly. Other possibilities will be mentioned.

An obvious difficulty of this approach is that the number of subproblems can grow exponentially in m and n. However, in Section 3.3 we will show that performing the decomposition in a particular way will result in the number of subproblems being reduced. The matrices D and Q are not uniquely determined and by constructing them in a particular way, the subproblems may be reduced in number to between 1 and m.

We will give numerical examples to illustrate both the decomposition method and the subproblem reduction procedure.

3.2 A Decomposition Method

In this section we propose a method for solving

$$\min\{c'x + x'Cx \mid a_i'x \le b_i,\ i = 1, \ldots, m\}, \tag{3.4}$$

where C is (n, n) and indefinite. Although C is not the Hessian matrix for the objective function for (3.4) (the Hessian is $2C$), we shall refer to it as such in order to avoid introducing unnecessary terminology. The method decomposes the n dimensional problem into m subproblems each having dimension $(n-1)$. Each of these $(n-1)$ dimensional subproblems is in turn decomposed into m subproblems each of dimension $(n-2)$. The process is continued until 1 dimensional problems are reached. These can be solved directly and combined to provide all local minimizers for the previous 2 dimensional problem and so on back up to the $(n-1)$ dimensional problem and finally, the n dimensional problem is solved. The structure of the problem is that of a tree. The top node is the given problem from which emanate m branches to the $(n-1)$ level. From each of these m nodes emanate m branches leading to the $(n-2)$ level.

The tree structure of the proposed method is similar to a method proposed by Murty [21]. Murty's method produces the global minima, whereas ours will produce the global minima *plus all isolated local minima* and *some non-isolated local minima*. The subproblems generated by the two methods are quite distinct. Further, we will show that the number of subproblems at any level can be reduced to a number between 1 and m and the number of reduced subproblems depends on the problem data in a way which will be made explicit in Theorem 3.3.1. We assume the feasible region for (3.4) is non-null.

In order to apply the theory developed in Chapter 2, we require a method which will construct (n, k) matrices Q and D satisfying $C = \frac{1}{2}[QD' + DQ']$ or determine that no such Q and D exist. If such Q and D are found, then the objective function for (3.4) can be written as $x'Cx = \frac{1}{2}[x'QD'x + x'DQ'x] = x'DQ'x$ which is of the same form as the objective function for (3.1). We next formulate such a method, which we refer to as Procedure $\Psi(C)$.

Procedure Ψ(C)

Given an (n, n) symmetric matrix C, Procedure $\Psi(C)$ determines whether or not C is indefinite, positive semidefinite or negative semidefinite. If C is indefinite, Procedure $\Psi(C)$ constructs two $(n, n-1)$ matrices D and Q with $\text{rank}(D) = n-1$, and which satisfy $C = \frac{1}{2}[DQ' + QD']$. In this case, we write $\Psi(C) = (Q,\ D)$. The details of procedure $\Psi(C)$ are as follows.

We first require an (n, n) nonsingular matrix M and an (n, n) diagonal matrix Λ satisfying $M'CM = \Lambda$, where the diagonal elements of Λ are all either -1, 0 or $+1$. Such matrices may be found by either performing an eigenvalue decomposition for C or by using a modified conjugate direction method described in [7]. The latter method requires only $O(n^3)$ arithmetic operations. It is straightforward to show that the diagonal elements of Λ are all nonnegative if and only if C is positive semidefinite, and are all non-positive if and only if C is negative semidefinite. In either of these cases, Procedure $\Psi(C)$ terminates with the relevant information. The remaining possibility is that Λ has two nonzero diagonal elements of opposite sign and this is equivalent to C being indefinite. In this case, Procedure $\Psi(C)$ continues as follows.

Suppose k and l are such that λ_k and λ_l are both nonzero, have opposite signs, and assume $k < l$. Then $\lambda_k + \lambda_l = 0$. Let e_i denote the i-th unit vector of dimension $n-1$. If $l < n$, define $\hat{D}$ and $\hat{Q}$ according to

$$\hat{D}' = [\, e_1, \ldots, e_{l-1}, e_k, e_l, e_{l+1}, \ldots, e_{n-1} \,], \text{ and } \hat{Q}' = \hat{D}'\Lambda.$$

If $l = n$, define

$$\hat{D}' = [\, e_1, \ldots, e_{n-1}, e_k \,], \text{ and } \hat{Q}' = \hat{D}'\Lambda.$$

Note that $\hat{D}'$ differs from the $(n-1, n-1)$ identity matrix by the insertion of e_k after column $l-1$. It is straightforward to show that $\hat{D}\hat{Q}'$ differs from Λ only in the (k,l)-th and (l,k)-th elements which are

$$(\hat{D}\hat{Q}')_{kl} = \lambda_l \text{ and } (\hat{D}\hat{Q}')_{lk} = \lambda_k.$$

But then $\lambda_k + \lambda_l = 0$ implies

$$\Lambda = \frac{1}{2}[\, \hat{D}\hat{Q}' + \hat{Q}\hat{D}' \,].$$

Thus

$$C = (M')^{-1}\Lambda M^{-1} = \frac{1}{2}[(M')^{-1}\hat{D}\hat{Q}'M^{-1} + (M')^{-1}\hat{Q}\hat{D}'M^{-1}]$$

and thus

$$D = (M')^{-1}\hat{D} \text{ and } Q = (M')^{-1}\hat{Q}$$

satisfy the conditions of Procedure Ψ. The procedure is then complete with $\Psi(C) = (Q, D)$

The decomposition method for (3.4) proceeds as follows. Because C in (3.4) is indefinite, we can successfully invoke Procedure Ψ to obtain $\Psi(C) = (Q, D)$ with $\text{rank}(D) = n - 1$ and the parametric LP (3.2) for (3.4) becomes

$$\min\{c'x \;+\; t'Q'x \mid a_i{}'x \;\leq\; b_i,\; i = 1, \ldots, m,\; D'x \;=\; t\}. \tag{3.5}$$

For a fixed value of the parameter vector t, (3.5) is an LP. The two possibilities for its' solution are:

(a) an optimal solution which is an extreme point,

(b) an optimal solution which is not an extreme point, or, the problem is unbounded from below.

Each of these possibilities will be accounted for separately.

For the first possibility (a), an extreme point for (3.5) must have n active constraints having linearly independent gradients. These must include the $(n-1)$ linearly independent rows of D' plus at least one of $a_1, a_2, \ldots, a_m$. There are thus m possibilities. Let $1 \;\leq\; k \;\leq\; m$. The k-th subproblem to be considered is

$$\min\{c'x \;+\; t'Q'x \mid D'x \;=\; t,\; a_k{}'x \;=\; b_k\}. \tag{3.6}$$

Assume first that a_k and the columns of D are linearly independent. The associated extreme point is the solution of the simultaneous linear equations

$$B_k x \;=\; \begin{bmatrix} t \\ b_k \end{bmatrix},$$

where

$$B_k \;=\; \begin{bmatrix} D' \\ a_k{}' \end{bmatrix}.$$

Partition B_k^{-1} as

$$B_k^{-1} = [H_k\, ,\, d]\, ,$$

where H_k is the $(n, n-1)$ matrix of the first $(n-1)$ columns of B_k^{-1} and d is the last column.

Note that $B_k B_k^{-1} = I$, ie,

$$\begin{bmatrix} D' \\ a_k' \end{bmatrix} [H_k \;\; d] = \begin{bmatrix} D'H_k & D'd \\ a_k'H_k & a_k'd \end{bmatrix} = I,$$

so

$$D'H_k = I, \;\; D'd = 0, \;\; a_k'H_k = 0 \;\text{ and }\; a_k'd = 1. \tag{3.7}$$

The extreme point x is written as follows;

$$x = H_k t + b_k d. \tag{3.8}$$

The remaining $(m-1)$ constraints of (3.5) must be satisfied;

$$a_i'H_k t \le b_i - b_k a_i'd, \quad i = 1, \ldots, m, \;\; i \ne k. \tag{3.9}$$

For x in (3.8) to be optimal, it must satisfy dual feasibility for (3.5). However, $(n-1)$ of the constraints for (3.5) are equalities and their dual variables are not constrained in sign. The dual variable for the sole active inequality constraint is $-(c + Qt)'d$ and it must be nonnegative; ie,

$$d'Qt \le -c'd. \tag{3.10}$$

Note that (3.9) and (3.10) define exactly m inequality constraints on the $(n-1)$ vector t.

Using (3.8), the objective function for (3.5) is

$$F_k(t) = c'x + t'Q'x = b_k c'd + (H_k' + b_k Q'd)'t + t'Q'H_k t,$$

a quadratic function of the $(n-1)$ components of t. The Hessian term for $F_k(t)$ can be simplified as follows. Because $2C = QD' + DQ'$, it follows from (3.7) that

$$\begin{aligned} 2CH_k &= QD'H_k + DQ'H_k \\ &= Q + D(Q'H_k). \end{aligned}$$

Multiplying on the left by H_k' gives

$$H_k'CH_k = \frac{1}{2}[H_k'Q + Q'H_k],$$

which shows that the symmetrized Hessian for (3.11) can be written as $H_k'CH_k$. Thus the objective function for (3.5) is

$$F_k(t) = b_k c'd + (H_k' + b_k Q'd)'t + t'H_k'CH_k t. \tag{3.11}$$

It remains to consider the case that a_k and the columns of D are linearly dependent. If $\text{rank}(A, D') = n-1$, the (3.5) does not possess an extreme point. This is considered in possibility (b). Otherwise, $\text{rank}(A, D') = n$ and assuming the feasible region for (3.5) is non null, this implies that the feasible region for (3.5) possesses extreme points (Best and Ritter, 1985, page 69). Thus (3.5) possesses an optimal solution which is an extreme point. But because (3.5) contains $n-1$ equality constraints having linearly independent gradients, it follows that one of the constraints $a_i'x \leq b_i,\ i = 1, \ldots, m,\ i \neq k$ must be active. Suppose its index is j. But then this optimal solution can be obtained from subproblem j and moreover, a_j is not linearly dependent on the columns of D. Consequently, subproblem k, namely (3.6) need not be considered further. We summarize this as

Proposition 3.2.1 *Suppose a_k and the columns of D are linearly dependent. Then the k-th subproblem (3.6) may be omitted.*

The second possibility (b) is that either (3.5) has optimal solutions which are not extreme points, or, (3.5) is unbounded from below for certain values of t. We now account for this. Any solution for $D'x = t$ can be written as

$$x = x(t) = Ht + \sigma s, \tag{3.12}$$

where H is an $(n, n-1)$ matrix, s is a non zero n-vector with $D's = 0$ and σ is a scalar variable. Moreover, it is straightforward to compute such H and s. Using (3.12), the objective function for (3.5) can be written as

$$F_0(t) = \sigma(c's + t'Q's) + c'Ht + t'Q'Ht.$$

The analysis can be continued further by solving the LP

$$\min\{s'Qt \mid AHt + \sigma As \leq b\}, \tag{3.13}$$

where both t and σ are taken as variables. The possible conclusions are summarized in

Proposition 3.2.2 *Let H and s be as in (3.12).*
(a) If $As \leq 0$, then both (3.5) and (3.4) are unbounded from below for all t satisfying $(c's + t'Q's) < 0$, $AHt + \sigma As \leq b$ and $\sigma \geq 0$.
(b) If $As \geq 0$, then both (3.5) and (3.4) are unbounded from below for all t satisfying $(c's + t'Q's) > 0$, $AHt + \sigma As \leq b$ and $\sigma \leq 0$.
(c) If As has at least two nonzero entries of opposite sign, then for all t such that $AHt + \sigma As \leq b$ for some σ, $x(t)$ is an alternate optimal solution for one of the m extreme point subproblems and thus need not be considered further.

(d) If $As = 0$, then for all t with $(c's + t'Q's) = 0$ and $AHt \leq b$, $x(t)$ is optimal for (3.5) for all σ, (3.5) has no extreme points, each of the m subproblems (3.5) is vacuous and (3.4) reduces to the single $(n-1)$-dimensional problem

$$\min\{c'Ht + t'Q'Ht \mid AHt \leq b,\ s'Qt = -c's\}. \tag{3.14}$$

In summary, the decomposition method proceeds as follows. For $k = 1, \ldots, m$, we minimize the quadratic function $F_k(t)$ subject to the constraints (3.9) and (3.10), omitting those which satisfy the hypothesis of Proposition 3.2.1. We then account for the various possibilities of Proposition 3.2.2. If $As \leq 0$, (that is, Proposition 3.2.2(a) applies), it may or may not be true that there are t with $(c's + t'Q's) < 0$, $AHt + \sigma As \leq b$ and $\sigma \geq 0$. The relevant possibility may be determined by solving the LP

$$\min\{s'Qt \mid AHt + \sigma As \leq b,\ \sigma \geq 0\}.$$

A similar remark applies to Proposition 3.2.2(b). If Proposition 3.2.2(d) applies, then (3.4) is reduced to the $(n-1)$ dimensional problem (3.14) and none of the other m subproblems need be solved; i.e., the n dimensional problem (3.4) is simply reduced to the $n-1$ dimensional problem (3.14). Thus we have deco mposed (3.4), having n variables and m constraints, into at most m subproblems, each having $n-1$ variables and m constraints. The process can be continued by decomposing each of the $n-1$ variable problems into at most m $(n-2)$ variable problems. The decomposition process may be continued, generating subproblems of successively smaller dimension. There are several possibilities concerning the subproblems. If the Hessian of a subproblem has exactly one negative eigenvalue, then it may be solved directly by the method described in Chapter 2. Alternatively, the decomposition may be continued until 1 variable problems are generated and these may be solved by inspection. If a subproblem is convex, it may be solved by any convex

QP algorithm (e.g. [10]). However, if a subproblem is concave, it will be difficult to find all isolated local minima and this is a major drawback of this method.

Let R_k denote the feasible region for the k-th subproblem; i.e., R_k consists of those t which satisfy (3.9) and (3.10). Suppose $\hat{t}$ is a local minimum for the k-th subproblem. The corresponding point for (3.4) is given by(3.8); namely $\hat{x} = H_k\hat{t} + b_k d$. This may or may not be a local minimum for (3.4). It will be a local minimizer for (3.4) provided $\hat{t}$ is a local minimizer for each subproblem i for which $\hat{t} \in R_i$. In addition, if the conditions of Proposition 3.2.2 for unbounded from below are satisfied and $s'Q'\hat{t} = c's$, then $\hat{x}$ is not a local minimizer for (3.4). Note that these conditions are quite simple to check.

We illustrate these concepts with

Example 3.2.1

$$\begin{aligned} \text{minimize} : &\quad -x_1 - 2x_2 - x_3 + x'Cx \\ \text{subject to} : &\quad 0 \le x_i \le 1,\ i = 1, 2, 3, \end{aligned}$$

where

$$C = \begin{bmatrix} 2. & -.5 & 4.5 \\ -.5 & -1. & -1. \\ 4.5 & -1. & 5. \end{bmatrix} = [DQ' + QD']/2,$$

and for simplicity we take

$$D = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 2 & 1 \\ 0 & -1 \\ 1 & 1 \end{bmatrix}.$$

For this problem, the parametric LP (3.5) becomes

Table 3.1: Two Dimensional Subproblems for Example 3.2.1

Active 'x' Constraint	Objective Function	Constraints
$x_1 = 0$	$\frac{1}{5}(t_1^2 + 4t_1t_2 - 2t_2^2 + 3t_1 - 7t_2)$	$0 \leq t_1 + t_2 \leq 5$, $0 \leq -2t_1 + 3t_2 \leq 5$, $9t_1 + 2t_2 \geq 8$.
$x_1 = 1$	$\frac{1}{5}({t_1}^2 + 4t_1t_2 - 2{t_2}^2 + 12t_1 - 5t_2 - 8)$	$1 \leq t_1 + t_2 \leq 6$, $-2 \leq -2t_1 + 3t_2 \leq 3$, $9t_1 + 2t_2 \leq 8$.
$x_2 = 0$	$\frac{1}{2}(4{t_1}^2 - 3t_1t_2 - 2{t_2}^2 - 2t_1 + 2t_2)$	$0 \leq t_2 \leq 2$, $0 \leq 2t_1 - 3t_2 \leq 2$, $9t_1 + 2t_2 \geq 8$.
$x_2 = 1$	$\frac{1}{2}(4{t_1}^2 - 3t_1t_2 - 2{t_2}^2 + 7t_1 + 4t_2 - 8)$	$1 \leq t_2 \leq 3$, $-5 \leq 2t_1 - 3t_2 \leq -3$, $9t_1 + 2t_2 \leq 8$.
$x_3 = 0$	$2{t_1}^2 + 3t_1t_2 - t_1 - 3t_2$	$0 \leq t_1 + t_2 \leq 1$, $0 \leq t_2 \leq 1$, $9t_1 + 2t_2 \leq 8$.
$x_3 = 1$	$2{t_1}^2 + 3t_1t_2 - 10t_1 - 5t_2 + 8$	$2 \leq t_2 \leq 3$, $5 \leq t_1 + t_2 \leq 6$, $9t_1 + 2t_2 \geq 8$.

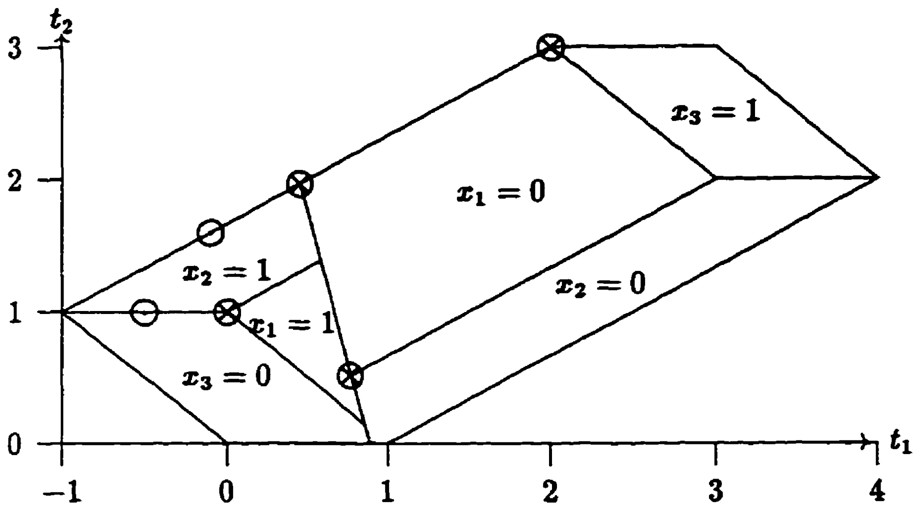

Figure 3.1: Two Dimensional Feasible Regions for Example 3.2.1

$$\begin{array}{lccccc}
\text{minimize:} & (2t_1 + t_2 - 1)x_1 & +(-t_2 - 2)x_2 & +(t_1 + t_2 - 1)x_3 & & \\
\text{subject to:} & x_1 & -x_2 & +3x_3 & = & t_1, \\
 & & x_2 & +2x_3 & = & t_2,
\end{array}$$

$$0 \le x_i \le 1, \quad i = 1, 2, 3.$$

There are 6 constraints in the problem. Each generates a subproblem with two variables. The objective function and constraints for each subproblem are summarized in Table 3.1. As well, the feasible regions for the subproblems are shown together in Figure 3.1.

In this example, each of the six subproblems has exactly five constraints rather than the six one might expect. This is because, for example, when $x_1 = 0$ is active, its gradient is linearly dependent on that of $x_1 \le 1$ and so produces a constraint of the form $0 \le 0$. Each of the six 2-dimensional problems is non-convex and can be decomposed into six 1-dimensional problems. Each set of six 1-dimensional problems consists of minimizing a piece wise quadratic function over at most 5 intervals. By Table 3.2, we know that local minima can be computed efficiently for the 1-dimensional case.

Each of the six 2 dimensional subproblems is non-convex and each such Hessian has exactly one negative eigenvalue. Each of these six subproblems could be solved using the method of Section 2.3 of Chapter 2. For problems with large numbers of constraints, this will be a computationally more attractive way to proceed. Local minima for the 2-dimensional problem are indicated in Figure 3.1 with circles. Those with a cross inside ($\otimes$) are local minima for at least one, but not all the regions in which they lie and are thus do not give local minima for the 3-dimensional problem. Those with a dot inside ($\odot$) are local minima for all the regions in which they lie and thus give local minima for the original problem. The local and global minimizers for Example 3.2.1 are shown in Table 3.2. Also shown are the local and global solutions for two variations of Example 3.2.1. These variations are obtained by changing the linear part of the objective function for Example 3.2.1.

We complete this section with an example which illustrates some of the possibilities for a problem having local minima but no global minima.

Example 3.2.2

$$\min\{x_1 x_2 \mid -x_2 \le 0\}$$

Here $A = [0, -1]$ and we take $Q = [0, 1]'$ and $D = [1, 0]'$. For this problem, the parametric LP (3.5) is $\min\{tx_2 \mid x_1 = t,\ -x_2 \le 0\}$. We first check to see if there is a region where the problem is unbounded from below. Take $H = [1, 0]'$ and $s = [0, 1]'$. Because $As \le 0$, $x(t) = Ht + \sigma s$ is feasible for all $\sigma \ge 0$. Furthermore, $c's + t'Q's = t < 0$ for all $t < 0$. Proposition 3.2.2(b) asserts that the example problem is unbounded from below for all $x_1 < 0$ and all $x_2 \ge 0$.

The remaining possibility for this example is that the single inequality constraint is active. In this case, the parametric LP (3.5) reduces to $\min\{0 \mid x_1 = t,\ -x_2 = 0\}$.

Table 3.2: Optimal Solutions for Example 3.2.1 and Some Variations

linear term	optimal type	objective value	point
$-3x_1 - 2x_2 - x_3$	global	-3.5	$(0.5, 1, 0)'$
	local	-3.45	$(0, 1, 0.3)'$
$-x_1 + 2x_2 - x_3$	global	-0.125	$(0.25, 0, 0)'$
	local	0.55	$(0, 1, 0.3)'$
	local	0.5	$(0.5, 1, 0)'$
$-0.5x_1 + 2x_2 - x_3$	global	-0.05	$(0, 0, 0.1)'$
	local	0.55	$(0, 1, 0.3)'$
	local	-0.03125	$(0.125, 0, 0)'$
	local	0.71875	$(0.375, 1, 0)'$

Since the original problem has just one constraint, the only restriction on t is (3.10), namely, $t \geq 0$. Thus $f(t) = 0$ for all $t \geq 0$ and all $t \geq 0$ are local minimizers of $f(t)$. However, $t = 0$ does not give a local minimizer for the 2 dimensional problem as it intersects the region $\{t \mid t < 0\}$ where the problem is unbounded from below.

In summary, the decomposition procedure has determined that for all (x_1, x_2) satisfying $x_1 < 0$ and $x_2 \geq 0$ the problem is unbounded from below, and, all points (x_1, x_2) satisfying $x_1 > 0$ and $x_2 = 0$ are local minimizers.

3.3 Reduction of the Number of Subproblems

In the previous section, we developed a decomposition method for locating all isolated local minimizers and some local minimizers. The method depends on the availability of two $(n, n-1)$ matrices D and Q satisfying

$$C = \frac{1}{2}(DQ' + QD'). \tag{3.15}$$

If D and Q satisfy (3.15) and H is any $(n-1, n-1)$ nonsingular matrix, then DH together with $Q(H^{-1})'$, respectively, also satisfy (3.15); i.e.,

$$C = \frac{1}{2}[(DH)(Q(H^{-1})')' + (Q(H^{-1})')(DH)'].$$

Thus, D and Q are not uniquely determined and it is reasonable to consider whether matrices D and Q can be constructed such that the number of subproblems will be reduced. If, for example, the columns of D include the gradients of one or more of the constraints of (3.4), i.e., $a_{\alpha_1}, \cdots, a_{\alpha_j}$ are columns of D, then from Proposition 3.2.1, the j subproblems $\alpha_1, \cdots, \alpha_j$ may be omitted. Before we establish the main result of this section, we need a lemma as follows.

Lemma 3.3.1 *Let $M = [d_1, \cdots, d_n]'$ be any (n, n) nonsingular matrix. If $(M^{-1})'$ C M^{-1} contains a (k, k) indefinite principal submatrix, then there exist two $(n, n-1)$ matrices D and Q with rank$(D) = n-1$ such that*

$$C = \frac{1}{2}(DQ' + QD') \tag{3.16}$$

and at least $n-k$ columns of D are identical to $n-k$ columns of M'.

Proof. Without loss of generality, assume that the (k, k) indefinite principal submatrix B is that induced by the last k rows and columns of $(M^{-1})'CM^{-1}$. Applying the decomposition procedure Ψ to B gives $\Psi(B) = (D_k, Q_k)$, where the dimensions of both D_k and Q_k are $(k, k-1)$, $\text{rank}(D_k) = k-1$ and

$$B = \frac{1}{2}(D_k Q_k' + Q_k D_k').$$

Let

$$(M^{-1})' C M^{-1} = \begin{bmatrix} F_1 & F_2 \\ F_2' & B \end{bmatrix},$$

and define

$$\hat{D} = \begin{bmatrix} I_{n-k} & 0 \\ 0 & D_k \end{bmatrix} \text{ and } \hat{Q} = \begin{bmatrix} F_1 & 0 \\ 2F_2' & Q_k \end{bmatrix},$$

where I_{n-k} denotes the identity matrix of dimension $n-k$. This implies

$$(M^{-1})' C M^{-1} = \frac{1}{2}(\hat{D}\hat{Q}' + \hat{Q}\hat{D}'),$$

and thus

$$C = \frac{1}{2}(M'\hat{D}\hat{Q}'M + M'\hat{Q}\hat{D}'M) = \frac{1}{2}(DQ' + QD'),$$

where $D = M'\hat{D}$ and $Q = M'\hat{Q}$. Since $\text{rank}(D_k) = k-1$, it follows that $\text{rank}(\hat{D}) = n-1$. Therefore $\text{rank}(D) = \text{rank}(\hat{D}) = n-1$ which completes the verification of (3.16). Since $D = M'\hat{D}$, the first $n-k$ columns of D are $d_1, \cdots, d_{n-k}$. This completes the proof of the lemma. □

Theorem 3.3.1 *Let $M = [d_1, \cdots, d_r, d_{r+1}, \cdots, d_n]'$ be nonsingular (n, n) matrix and $[d_1, \cdots, d_r]'$ be a (r, n) submatrix of A, where $r \le n$. If $(M^{-1})'CM^{-1}$ contains a (k, k) indefinite principal submatrix B induced by $\gamma_1, \cdots, \gamma_k$ rows and columns, then there exist two $(n, n-1)$ matrices D and Q with $\text{rank}(D) = n-1$ such that (3.16) holds and (3.4) can be decomposed into at most $(m - r + l)$ subproblems of dimension $(n-1)$ by using (3.16) and all the subproblems corresponding to the following indices may be omitted*

$$\{1, \cdots, r\} \setminus \{\gamma_i \mid 1 \le \gamma_i \le r, \ i = 1, \cdots, k\},$$

where l is the total number of elements of the set $\{\gamma_i \mid 1 \leq \gamma_i \leq r,\ i = 1, \cdots, k\}$.

Proof. By Lemma 3.3.1 and its proof, there exist two $(n, n-1)$ matrices D and Q with $\text{rank}(D) = n - 1$ such that (3.16) holds and the columns of D contains all d_j with j satisfying

$$j \in \{1, \cdots, r\} \setminus \{\gamma_i \mid 1 \leq \gamma_i \leq r,\ i = 1, \cdots, k\}.$$

From the definition of l, we know that D contains $r - l$ columns of $[d_1, \cdots, d_r]$. From Proposition 3.2.1, these $r - l$ subproblems may be omitted. Therefore (3.4) can be decomposed into at most $(m - r + l)$ subproblems of dimension $(n-1)$ by using the decomposition (3.16). This completes the proof of the theorem. □

From Theorem 3.3.1, we can see that the number of subproblems omitted will be bigger if l is smaller. So, the number of subproblems omitted will be big if B is located in the bottom of the right hand side of $(M')^{-1}CM^{-1}$ or close to the bottom of the right hand side of $(M')^{-1}CM^{-1}$. If $\text{rank}(A) = n$, then we can take M to be an (n, n) nonsingular submatrix of A. In this case, r and l of Theorem 3.3.1 will be n and k, respectively. If $\text{rank}(A) = r < n$, we can pick up r linearly independent rows of A with any other $n - r$ n-vectors to form nonsingular matrix M. In this case, we have $0 \leq l \leq k$. From the above theorem, we can see that the best case occurs when $(M^{-1})CM^{-1}$ contains a $(2, 2)$ indefinite principal submatrix for some submatrix M of A. In this case, (3.4) can be decomposed into $(m - n + 2)$ $(n-1)$-dimensional subproblems, i.e., the biggest number of subproblems omitted is $n - 2$. In fact, we can show that the biggest number of subproblems omitted can be $n - 1$ provided $(M^{-1})'CM^{-1}$ contains a zero diagonal entry for some submatrix M of A. This can be explained as follows. Let $(M^{-1})'CM^{-1} = [f_{ij}]$. Without loss of generality, assume $f_{nn} = 0$. Let

$$B = \begin{bmatrix} f_{(n-1)(n-1)} & f_{n(n-1)} \\ f_{(n-1)n} & 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad Q_2 = \begin{bmatrix} f_{(n-1)(n-1)} \\ 2f_{n(n-1)} \end{bmatrix}$$

It is easy to check that

$$B = \frac{1}{2}(D_2 Q_2' + Q_2 D_2').$$

By following the proof of Lemma 3.3.1, we have

$$\hat{D} = \begin{bmatrix} I_{n-2} & 0 \\ 0 & D_2 \end{bmatrix} = \begin{bmatrix} I_{n-1} \\ 0 \end{bmatrix}.$$

So, $D = M'\hat{D}$ will have $n-1$ columns same as columns of M. Hence the number of subproblems omitted will be $n-1$.

Based on the constructive proof of Lemma 3.3.1, we next give a detailed formulation for the decomposition (3.16) if there exists an (n, n) nonsingular submatrix M such that $(M^{-1})CM^{-1}$ contains a (k, k) indefinite principal submatrix. We refer to this as procedure Ψ_1.

Procedure $\Psi_1(C, M, k)$

Given an (n, n) symmetric matrix C and an (n, n) nonsingular matrix M such that $(M^{-1})'CM^{-1}$ contains an (k, k) indefinite principal submatrix, procedure $\Psi_1(C, M, k)$ produces two $(n, n-1)$ matrices D and Q such that $C = \frac{1}{2}[DQ' + QD']$, $\text{rank}(D) = n-1$ and at least $n-k$ columns of D are formed by some $n-k$ rows of M. The details of procedure $\Psi_1(C, M, k)$ are as follows.

Let $(M^{-1})'CM^{-1} = (f_{ij})$ and B be an indefinite submatrix induced by $\gamma_1, \cdots, \gamma_k$ columns and rows of (f_{ij}) with $\gamma_1 < \gamma_2 < \cdots < \gamma_k$. By procedure $\Psi(B)$, we can compute two $(k, k-1)$ matrices D_k and Q_k such that

$$B = \frac{1}{2}(D_k Q_k' + Q_k D_k').$$

Let $D_k = (\bar{d}_{\mu\nu})$, $Q_k = (\bar{q}_{\mu\nu})$, $\hat{D} = (\hat{d}_{ij})$ and $\hat{Q} = (\hat{q}_{ij})$. Let $J = \{\gamma_1, \cdots, \gamma_k\}$ and $K = \{1, \cdots, k\}$. The matrices $\hat{D}$ and $\hat{Q}$ are formulated as follows

$$\hat{d}_{ij} = \begin{cases} 1 & \text{if } i = j,\ i \notin J,\ i < \gamma_k, \\ 1 & \text{if } i = j+1,\ i > \gamma_k, \\ \bar{d}_{\mu\nu} & \text{if } j < \gamma_k,\ i = \gamma_\mu,\ j = \gamma_\nu \text{ for some } \mu, \nu \in K, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\hat{q}_{ij} = \begin{cases} f_{ij} & \text{if } i, j \notin J,\ j < \gamma_k, \\ 2f_{ij} & \text{if } j \notin J,\ i \in J,\ j < \gamma_k, \\ 2f_{i(j+1)} & \text{if } i \in J,\ j \geq \gamma_k, \\ f_{i(j+1)} & \text{if } i \notin J,\ j \geq \gamma_k, \\ \bar{q}_{\mu\nu} & \text{if } j < \gamma_k,\ i = \gamma_\mu,\ j = \gamma_\mu \text{ for some } \mu, \nu \in K, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, D and Q are obtained by setting $D = M'\hat{D}$ and $Q = M'\hat{Q}$.

Now let us demonstrate the above idea with a quadratic programming problem having box constraints; i.e., we assume (3.4) has the special form

$$\min \{c'x + x'Cx \mid 0 \leq x_i \leq 1,\ i = 1, \cdots, n\}. \tag{3.17}$$

Here, $A = (I_n, -I_n)'$, where I_n is an (n,n) identity matrix. So, each nonsingular (n,n) submatrix of A is a diagonal matrix with diagonal entries being 1 or -1. Hence in order to check conditions of Theorem 3.3.1, we only need to check if C has an indefinite principal submatrix. This gives the following corollary.

Corollary 3.3.1 *In (3.17), if C has an (k,k) indefinite submatrix, then there exist two $(n, n-1)$ matrices D and Q with rank$(D) = n-1$ such that*

$$C = \frac{1}{2}(DQ' + QD') \tag{3.18}$$

and at least $n-k$ columns of D are formed by some $n-k$ rows of I_n. Furthermore, (3.17) can be decomposed into at most $2k$ subproblems in $n-1$ dimensional space by using (3.18).

From Corollary 3.3.1, we can see that if C has a $(2,2)$ indefinite principal submatrix, then (3.17) can be decomposed into at most 4 subproblems in $n-1$ dimensional space. In fact, it is easy to see that if C has a zero diagonal entry, then (3.17) can be decomposed into at most 2 subproblems in $n-1$ dimensional space.

Example 3.3.1. Consider (3.17) with $n=6$ and

$$C = \begin{bmatrix} 17 & 11 & -5 & 21 & -14 & -5 \\ 11 & 12 & -5 & 16 & -11 & -3 \\ -5 & -5 & 2 & -7 & 5 & 1 \\ 21 & 16 & -7 & 27 & -18 & -6 \\ -14 & -11 & 5 & -18 & 12 & 4 \\ -5 & -3 & 1 & -6 & 4 & 2 \end{bmatrix}.$$

By the conjugate direction algorithm presented in the appendix, it can be checked that C has exactly two negative eigenvalues. Also C has two $(2,2)$ indefinite principal submatrices as follows:

$$\begin{bmatrix} 12 & -5 \\ -5 & 2 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & 5 \\ 5 & 12 \end{bmatrix}.$$

For these two principal submatrices, it can be checked by the conjugate direction algorithm that the first submatrix is the best choice to formulate D, Q and four subproblems because the Hessian of each of the four subproblems has exactly one negative eigenvalue. The method of Section 2.3 of Chapter 2 can be applied directly to these subproblems.

Remark 3.3.1. If the decomposition procedure in Section 3.2 is performed for (3.17), the subproblems will not be quadratic programming problems with only box constraints. But if we are only concerned about a global minimizer for (3.17), we can guarantee that the subproblems have only box constraints and the number of subproblems can be reduced by the above technique. For example, we can use Murty's method to generate subproblems and use our techniques to reduce the number of subproblems.

Finally, let us return to the general QP and consider how to check the conditions imposed in Theorem 3.3.1, that is, $(M^{-1})'CM^{-1}$ contains an (k, k) indefinite principal submatrix for some (n, n) nonsingular matrix M. From Corollary 3.3.1, we can see that even for the problem with box constraints, we need to check all principal submatrices to find the smallest size of such a matrix. If C has a (k, k) indefinite principal submatrix and k is a relatively small number (say, $k = 2$ or 3), then we can find the smallest size of indefinite principal submatrix by enumerating all principal submatrices starting with $(2, 2)$ submatrices. If k is relatively large, then this approach will consume a lot of time. If constraints are not box constraints, the situation even worse. What we can do is enumerate some of the (n, n) nonsingular submatrices such that each of them contain a maximal linearly independent columns of A' and check all $(2, 2)$ and $(3, 3)$ principal submatrices for $(M^{-1})'CM^{-1}$ for each enumerated M. Of course, this can not guarantee we can find a $(2, 2)$ or $(3, 3)$ indefinite principal submatrix. The following proposition tells us that we can guarantee to reduce certain number of subproblems although this number may be quite small.

Proposition 3.3.1 *Assume that C has k_1 positive eigenvalues and k_2 negative eigenvalues, then for any nonsingular (n, n) submatrix M, each $(n-l, n-l)$ principal submatrix of $(M^{-1})'CM^{-1}$ will be indefinite or singular, where $l = \min\{k_1, k_2\} - 1$.*

Proof. Since C has k_1 positive eigenvalues and k_2 negative eigenvalues, so too does $(M^{-1})'CM^{-1}$. Let

$$(M^{-1})'CM^{-1} = \begin{bmatrix} F_1 & F_2 \\ F_2' & B \end{bmatrix},$$

where B is an $(n-l, n-l)$ submatrix, F_1 is an (l, l) submatrix and F_2 is an $(l, n-l)$ submatrix. It is sufficient to show that B is indefinite or singular. If not, B will be positive definite or negative definite. Consequently, B is invertible. Hence

$$\begin{bmatrix} I & -F_2B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} F_1 & F_2 \\ F_2' & B \end{bmatrix} \begin{bmatrix} I & 0 \\ -B^{-1}F_2' & B^{-1} \end{bmatrix} = \begin{bmatrix} F_1 - F_2B^{-1}F_2' & 0 \\ 0 & B^{-1} \end{bmatrix}.$$

Let us assume B is positive definite (the negative definite case can be treated similarly). Then $(M^{-1})'CM^{-1}$ has at most l negative eigenvalues. But $l = \min \{k_1, k_2\} - 1 < k_2$, a contradiction. So B is indefinite or singular. This completes the proof of the proposition. □

Remark 3.3.2. It can be shown that each $(n-l, n-l)$ principal submatrix of $(M^{-1})'CM^{-1}$ in Proposition 3.3.1 is indefinite even if it is singular.

Since we can use a conjugate direction algorithm to check the number of positive eigenvalues and negative eigenvalues, Proposition 3.2.1 and Theorem 3.3.1 tell us that certain number of subproblems can always be reduced. Before concluding this section, let us apply the techniques we have developed to Example 3.2.1.

Example 3.3.2. Consider Example 3.2.1.

Since the Hessian C in Example 3.2.1 contains a principal submatrix

$$\begin{bmatrix} -1 & -1 \\ -1 & -5 \end{bmatrix},$$

we can decompose it as follows

$$\begin{bmatrix} -1 & -1 \\ -1 & -5 \end{bmatrix} = \frac{1}{2}[\begin{pmatrix} 1 \\ \sqrt{6}+1 \end{pmatrix}(-1, \sqrt{6}-1) + \begin{pmatrix} -1 \\ \sqrt{6}-1 \end{pmatrix}(1, \sqrt{6}+1)].$$

So, $D_2' = (-1, \sqrt{6}+1)$, $Q_2' = (-1, \sqrt{6}-1)$,

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1+\sqrt{6} \end{bmatrix} \text{ and } Q = \begin{bmatrix} 2 & 0 \\ -1 & -1 \\ 9 & \sqrt{6}-1 \end{bmatrix}.$$

Now by setting $D'x = t$, we can decompose the QP into four subproblems. It can be checked that two of them are infeasible. In fact, for $x_2 = 0$, the subproblem has the following constraints:

$$\begin{aligned} & 0 \le t_1 \le 1, \\ & 0 \le t_2 \le 1+\sqrt{6}, \\ & \frac{-1-2\sqrt{6}-(10+\sqrt{6})t_1 - 2\sqrt{6}t_2}{1+\sqrt{6}} \ge 0. \end{aligned}$$

Obviously, it is infeasible. Similarly, we can show that the subproblem corresponding to $x_3 = 1$ is also infeasible. So, there are only two subproblems formulated for this decomposition. Let us write down these two subproblems as follows.

From $x_2 = 1$, we have

$$\begin{aligned} & \min 2t_1^2 + \frac{9t_1t_2}{1+\sqrt{6}} + \frac{\sqrt{6}-1}{\sqrt{6}+1}t_2^2 - \frac{11+2\sqrt{6}}{1+\sqrt{6}}t_1 - \frac{1+2\sqrt{6}}{1+\sqrt{6}}t_2 - \frac{2\sqrt{6}+1}{1+\sqrt{6}}, \\ & \text{subject to } \; 0 \le t_1 \le 1, \; 1 \le t_2 \le \sqrt{6}+2. \end{aligned}$$

From $x_3 = 0$, we have

$$\min 2t_1^2 - t_1t_2 - t_2^2 - t_1 - 2t_2,$$

$$\text{subject to } \; 0 \le t_1 \le 1, \; 0 \le t_2 \le 1.$$

It can be checked that $(0, 0.3(\sqrt{6}+1)+1)$ and $(0.5, 1)$ are two local minimizers for the first subproblem, and $(0.5, 1)$ is a local minimizer for the second subproblem. So, $(0.5, 1)$ and $(0, 0.3(\sqrt{6}+1)+1)$ are two local minimizers. Since the corresponding points of $(0.5, 1)$ and $(0, 0.3(\sqrt{6}+1)+1)$ in the original space are $(0.5, 1, 0)$ and $(0, 1, 0.3)$, respectively, $(0.5, 1, 0)$ and $(0, 1, 0.3)$ are two local minimizers for QP.

Example 3.3.3 Each of the final two problems in Table 3.2 were reduced from 6 to 3 2-dimensional subproblems.

3.4 Conclusion

We have developed a decomposition method to locate the global minimum, all isolated local minima and some of the nonisolated local minima for a general indefinite QP. We have shown that the number of subproblems can be reduced by constructing a proper decomposition for the Hessian matrix and a corresponding algorithm is also established.

If the decomposition procedure terminates with the subproblem QP's each one of which has a Hessian matrix having exactly one negative eigenvalue, then the method will compute all isolated local minimizers and some nonisolated local minimizers of QP. In addition, the method will compute the global minimizer of QP, provided it exists, and will provide the information that QP is unbounded from below when that is the case.

Chapter 4

A Class of Copositivity Problems

4.1 Introduction

In this chapter, we will use the results developed in Chapter 2 and 3 to solve a special class of copositivity problems.

For a given (n, n) real symmetric matrix C and a polyhedron cone F, C is called copositive on F if $x'Cx \geq 0$ for any $x \in F$. The problem we are concerned with is to determine whether C is copositive on F whenever C and F are given. This is an NP-hard problem even for $F = \mathrm{E}^n_+$, the positive orthant of E^n (see [22]). There are several applications for determining whether a given matrix is copositive. For example, see [13] and [21]. Let $a_i' \in \mathrm{E}^n$ for $i = 1, \cdots, m$, $A = (a_1, \cdots, a_m)'$ and define

$$F = \{x \in \mathrm{E}^n \mid Ax \leq 0\}. \tag{4.1}$$

From the definition of C being copositive on F, it is easy to see that the problem can be solved by solving the following quadratic programming problem:

$$\text{QP} \qquad \alpha = \min \{x'Cx \mid x \in F\}.$$

Obviously, $\alpha = 0$ iff C is copositive on F.

When $F = \mathrm{E}^n_+$, Valiaho [29] pointed out that the copositivity of C on F can be solved by

$$\min \{x'Cx \mid x_i \geq 0,\ i = 1, \cdots, n,\ i \neq k,\ x_k = 1\};$$

i.e.,

$$\min \{c_{kk} + 2c(k)'x(k) + x(k)'C(k)x(k) \mid x(k) \geq 0\}, \tag{4.2}$$

where $x(k) = (x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n)'$, $C(k)$ is an $(n-1, n-1)$ principal submatrix of C obtained by deleting the k-th row and column from C, $c(k) = (c_{k1}, \cdots, c_{kk-1},\ c_{kk+1}, \cdots, c_{kn})'$, and k is any fixed index with $1 \leq k \leq n$. If C contains a maximal $(n-1, n-1)$ positive semidefinite principal submatrix, (4.2) is a convex problem for some k. So it can be solved by standard quadratic programming algorithms. If C contains a maximal $(n-2, n-2)$ positive semidefinite principal submatrix, then for some k (4.2) can be solved efficiently by parametric quadratic programming techniques. The reader may refer to [29] for details. Based on the above analysis, we can see that the copositivity problem can be solved efficiently for some special classes of C by quadratic programming techniques.

In this chapter we will also consider solving a special classes of copositivity problems. We assume that C has exactly one or two negative eigenvalues and F is a general polyhedral cone defined by (4.1). In Section 4.2, we will develop an algorithm for a class of copositivity problems in which C is an integral diagonal matrix. In Section 4.3, we will show how a general copositivity problem in which

C has exactly one negative eigenvalue can be transformed into the formulation of Section 4.2. In Section 4.4, we will report numerical results for the algorithm.

In what follows, we need the concept of a projected Hessian.

Definition 4.1.1 For any (n, n) symmetric matrix Q, any $h \in \mathrm{E}^n \setminus \{0\}$ and any real number ν, we call Q_h the projected Hessian of Q on $h'x = \nu$ if

$$Q_h = (v_1, \cdots, v_{n-1})' Q (v_1, \cdots, v_{n-1}),$$

where $\mathrm{rank}(v_1, \cdots, v_{n-1}) = n - 1$ and $v_i \in \{x \in \mathrm{E}^n \mid h'x = 0\}$ for $i = 1, \cdots, n-1$.

From this definition, we can see that the projected Hessian Q_h of Q on $h'x = \nu$ is independent of ν and there are an infinite number of projected Hessians for given Q and h. It can be shown that for given Q and h all projected Hessians have same number of positive and negative eigenvalues. So, $x'Cx$ is convex on $h'x = 1$ iff a projected Hessian C_h of C on $h'x = 1$ is positive semidefinite.

4.2 An Algorithm

In this section, we will present an algorithm for the class of copositivity problems in which A is a rational matrix, $\mathrm{rank}(A) = n$ and $C = \mathrm{diag}\ (-d_1, d_2, \cdots, d_n)$ with d_i being a positive integer for $i = 1, \cdots, n$. Later on, the case that $C = \mathrm{diag}\ (-d_1, -d_2, d_3, \cdots, d_n)$ with d_i being a positive integer for $i = 1, \cdots, n$ will also be discussed. The first result we will give is quite simple, but it will give us a useful idea for the construction of the algorithm.

Proposition 4.2.1 *The optimal value $\alpha = 0$ iff for any $h \in E^n$, if $\{x \in F \mid h'x = 1\}$ is nonempty and bounded, then*

$$\bar{\alpha} = \ min\ \{x'Cx \mid x \in F, h'x = 1\} \geq 0.$$

Proof. The necessary condition is obvious. We need only show the sufficient condition. Since $\{x \in F \mid h'x = 1\}$ is nonempty and bounded, $\{x \in F \mid h'x = 0\} = \{0\}$. Now we claim that $h'x > 0$ for any $x \in F \setminus \{0\}$. Assume on the contrary that the claim is not true, then there is an $x^1 \in F \setminus \{0\}$ such that $h'x^1 \leq 0$. If $h'x^1 = 0$, then $x^1 \in \{x \in F \mid h'x = 0\}$; i.e., $\{x \in F \mid h'x = 0\} \neq \{0\}$, a contradiction. So, $h'x^1 < 0$. Since $\{x \in F \mid h'x = 1\}$ is nonempty, let x^2 be any element of this set. Let $\lambda = -h'x^1$, then $\lambda > 0$, $x^1 + \lambda x^2 \in F$ and $h'(x^1 + \lambda x^2) = 0$. So, $\{x \in F \mid h'x = 0\} = \{0\}$ implies $x^1 + \lambda x^2 = 0$; i.e., $x^1 = -\lambda x^2$. Since $x^1, x^2 \in F$, $a_i'x^1 \leq 0$ and $a_i'x^1 = -\lambda a_i'x^2 \geq 0$ for $i = 1, \cdots, m$. So, $a_i'x^1 = 0$ for $i = 1, \cdots, m$. Since $x^1 \neq 0$, rank$(A) \leq n-1$, a contradiction. Therefore the claim is true. Hence for any $\bar{x} \in F \setminus \{0\}$, $\bar{x}/h'\bar{x} \in \{x \in F \mid h'x = 1\}$ and $(\bar{x}/h'\bar{x})'C(\bar{x}/h'\bar{x}) \geq 0$; i.e., $\bar{x}'C\bar{x} \geq 0$. So, $\alpha = 0$. The proof of the proposition is thus complete. □

Example 4.2.1 Let $n = 2$, $C = \text{diag}(-1, 1)$ and $F = \{x \in E^2 \mid -2x_1 + x_2 \leq 0, x_1 - x_2 \leq 0\}$.

For this example, it is easy to see that $x \in F$ implies $x_1 \geq 0$, $x_2 \geq 0$ and $x_2 \geq x_1$. So, $x'Cx = x_2^2 - x_1^2 \geq 0$ for any $x \in F$; i.e., C is copositive on F. In the following we will give a different approach for this example and hope this approach to be generalized to high dimensional problems. In doing so, consider $x_1 = 1$. Obviously $\{x \in F \mid x_1 = 1\} = \{x \in E^2 \mid x_1 = 1,\ 1 \leq x_2 \leq 2\}$ and $x'Cx = -1 + x_2^2$ on $\{x \in F \mid x_1 = 1\}$ is convex. By solving

$$\alpha = \min \{-1 + x_2^2 \mid 1 \leq x_2 \leq 2\},$$

we have $\alpha = 0$. So by Proposition 4.2.1, C is copositive on F.

This is a simple example but it helps to make the following point. If a hyperplane $h'x = 1$ can be constructed such that $\{x \in F \mid h'x = 1\}$ is nonempty, bounded and $x'Cx$ is convex on $\{x \in F \mid h'x = 1\}$, then the copositivity of C

on F can be solved by solving a convex quadratic programming problem. So it is necessary to investigate the conditions such that the proper hyperplane exists. The following characterization gives us a direct information concerning how to construct a hyperplane such that the projected Hessian of C on the hyperplane is positive semidefinite.

Proposition 4.2.2 *If $Q = diag(\beta_1, \cdots, \beta_n)$ and $1 < k < n$ with $\beta_i = -1$ for $i = 1, \cdots, k$ and $\beta_j = 1$ for $j = k+1, \cdots, n$, then for any given vector $h \in E^n \setminus \{0\}$, Q_h has exactly $k-1$ negative eigenvalues iff*

$$\sum_{i=1}^{k} h_i^2 - \sum_{i=k+1}^{n} h_i^2 \geq 0. \tag{4.3}$$

Proof. First of all, let us consider two special cases for h as follows.

(i) There exists a j with $1 \leq j \leq k$ such that $h_j \neq 0$ and $h_i = 0$ for $i = k+1, \cdots, n$.

(ii) There exists a l with $k+1 \leq l \leq n$ such that $h_l \neq 0$ and $h_i = 0$ for $i = 1, \cdots, k$.

Let us consider case (i) first. Since (4.3) always holds in this case, we only need to show that Q_h has exactly $k-1$ negative eigenvalues. Without loss of generality, let $j = 1$; i.e., $h_1 \neq 0$. From $h'x = 0$, we can write

$$x_1 = -\sum_{i=2}^{k} \frac{h_i}{h_1} x_i.$$

So, on $h'x = 0$, we have

$$x'Qx = -\left(\sum_{i=2}^{k} \frac{h_i}{h_1} x_i\right)^2 - \sum_{i=2}^{k} x_i^2 + \sum_{i=k+1}^{n} x_i^2,$$

which implies that Q_h has exactly $k-1$ negative eigenvalues. So the proposition is true for case (i).

Now let us consider case (ii). Since (4.3) does not hold in this case, we only need to show that Q_h has exactly k negative eigenvalues. Without loss of generality, let $l = n$ so that $h_n \neq 0$. From $h'x = 0$, we can write

$$x_n = -\sum_{i=k+1}^{n-1} \frac{h_i}{h_n} x_i,$$

which implies

$$x'Qx = -\sum_{i=1}^{k} x_i^2 + \sum_{i=k+1}^{n-1} x_i^2 + (\sum_{i=k+1}^{n-1} \frac{h_i}{h_n} x_i)^2.$$

Thus Q_h has exactly k negative eigenvalues. Hence the proposition is also true for this case.

In the following, we assume that there exist a j and an l with $1 \leq j \leq k$ and $k+1 \leq l \leq n$ such that $h_j \neq 0$ and $h_l \neq 0$. Let us treat case $k = 1$ first. In this case, we can write

$$x_1 = -\sum_{i=2}^{n} \frac{h_i}{h_1} x_i$$

from $h'x = 0$. So, on $h'x = 0$, we have

$$x'Qx = -(\sum_{i=2}^{n} \frac{h_i}{h_1} x_i)^2 + \sum_{i=2}^{n} x_i^2.$$

If (4.3) holds, then Cauchy's inequality implies

$$(\sum_{i=2}^{n} \frac{h_i}{h_1} x_i)^2 \leq (\sum_{i=2}^{n} \frac{h_i^2}{h_1^2})(\sum_{i=2}^{n} x_i^2) \leq \sum_{i=2}^{n} x_i^2.$$

So, $x'Qx \geq 0$ for any x satisfying $h'x = 0$. This implies that Q_h is positive semidefinite. Conversely, if Q_h is positive semidefinite, then $x'Qx \geq 0$ for any x satisfying $h'x = 0$. This implies

$$\sum_{i=2}^{n} x_i^2 \geq (\sum_{i=2}^{n} \frac{h_i}{h_1} x_i)^2$$

for any $(x_2, \cdots, x_n) \in \mathrm{E}^{n-1}$. By taking $x_i = h_i/h_1$ for $i = 2, \cdots, n$, we have

$$\sum_{i=2}^{n} \frac{h_i^2}{h_1^2} \geq (\sum_{i=2}^{n} \frac{h_i^2}{h_1^2})^2;$$

i.e.,

$$h_1^2 \geq \sum_{i=2}^{n} h_i^2.$$

Therefore, the proposition holds for $k = 1$. Now let us assume $k \geq 2$. Without loss of generality, assume that $h_1 \neq 0$ and $h_n \neq 0$. From

$$x_1 = -\sum_{i=2}^{n} \frac{h_i}{h_1} x_i,$$

we know that Q_h can be written as

$$\begin{bmatrix} -I_{k-1} & 0 \\ 0 & I_{n-k} \end{bmatrix} - bb',$$

where $b = (-h_2/h_1, \cdots, -h_n/h_1)'$. Since $h_n \neq 0$, $b_{n-1} \neq 0$. By mathematical induction, it is not hard to prove that

$$det(\lambda I - Q_h) =$$
$$(\lambda+1)^{k-2}(\lambda-1)^{n-k-1}[\lambda^2 + (\sum_{i=1}^{n-1} b_i^2)\lambda - 1 - \sum_{i=1}^{k-1} b_i^2 + \sum_{i=k}^{n-1} b_i^2].$$

So, Q_h has exactly $k-1$ negative eigenvalues iff

$$\lambda^2 + (\sum_{i=1}^{n-1} b_i^2)\lambda - 1 - \sum_{i=1}^{k-1} b_i^2 + \sum_{i=k}^{n-1} b_i^2 = 0$$

has exactly one negative solution which is equivalent to

$$1 + \sum_{i=1}^{k-1} b_i^2 - \sum_{i=k}^{n-1} b_i^2 \geq 0,$$

that is,

$$\sum_{i=1}^{k} h_i^2 \geq \sum_{i=k+1}^{n} h_i^2.$$

This completes the proof of the proposition. □

From Proposition 4.2.2, it is easy to get the following corollary.

Corollary 4.2.1 *For a given vector $h \in E^n \setminus \{0\}$, a projected Hessian C_h of C on $h'x = 1$ is positive semidefinite iff*

$$\frac{h_2^2}{d_2} + \cdots + \frac{h_n^2}{d_n} \leq \frac{h_1^2}{d_1}.$$

From Corollary 4.2.1 we can see that there are a lot of choices for the vector h such that C_h is positive semidefinite. Among all these h such that C_h is positive semidefinite we need to choose one such that $\{x \in F \mid h'x = 1\}$ is nonempty and bounded. The following example illustrates how to construct such h to solve the copositivity problem.

Example 4.2.2 Let $n = 2$, $C = \text{diag}(-1, 1)$ and $F = \{x \in \mathrm{E}^2 \mid x_1 - x_2 \leq 0,$ $-x_1 - 2x_2 \leq 0\}$.

Let us try the approach used in Example 4.2.1. For any $\alpha, \beta \in \mathrm{E}^1$ if $\{x \in F \mid \alpha x_1 + \beta x_2 = 1\}$ is nonempty and bounded, then $\beta \neq 0$. So, in order to make

$$x'Cx = \frac{1}{\beta^2} - \frac{2\alpha}{\beta^2}x_1 + \frac{\alpha^2 - \beta^2}{\beta^2}x_1^2$$

on $\{x \in F \mid \alpha x_1 + \beta x_2 = 1\}$ convex, we must have $\alpha^2 \geq \beta^2$; i.e., $|\alpha| \geq |\beta|$. Now consider a point $(-\beta/\alpha, 1)'$. Since $|\alpha| \geq |\beta|$, $|\beta/\alpha| \leq 1$. This implies $(-\beta/\alpha, 1)' \in F$. Because $\alpha(-\beta/\alpha) + \beta = 0$, we have

$$(-\beta/\alpha, 1)' \in \{x \in F \mid \alpha x_1 + \beta x_2 = 0\}.$$

Therefore $\{x \in F \mid \alpha x_1 + \beta x_2 = 1\}$ is unbounded, a contradiction. Hence the approach does not work. However if we divide F into two pieces

$$F_1 = \{x \in F \mid x_1 \geq 0\} = \{x \in \mathrm{E}^2 \mid x_1 \geq 0,\ x_1 - x_2 \leq 0\},$$

$$F_2 = \{x \in F \mid x_1 \leq 0\} = \{x \in \mathrm{E}^2 \mid x_1 \leq 0,\ -x_1 - 2x_2 \leq 0\}$$

and consider the copositivity of C on F_1 and F_2, respectively, then the approach will work. Obviously, C is copositive on F iff C is copositive on both of F_1 and F_2. Note that $\{x \in F_1 \mid 2x_1 + x_2 = 1\}$ and $\{x \in F_2 \mid -2x_1 + x_2 = 1\}$ are nonempty, bounded and $x'Cx$ is convex on both of them. So from

$$\begin{aligned}\alpha_2 &= \min\{x'Cx \mid x_1 \leq 0,\ -x_1 - 2x_2 \leq 0,\ -2x_1 + x_2 = 1\} \\ &= \min\{1 + 4x_1 + 3x_1^2 \mid -\frac{2}{5} \leq x_1 \leq 0\} = -\frac{3}{25} < 0,\end{aligned}$$

we know that C is not copositive on F.

Now let us return to the general problem formulated at the beginning of this section. From Example 4.2.2, we can see that it is necessary to consider consider the following two quadratic programming problems:

$$\text{QP}_1 \qquad \alpha_1 = \min\{x'Cx \mid x \in F,\ x_1 \geq 0\}$$

and

$$\text{QP}_2 \qquad \alpha_2 = \min\{x'Cx \mid x \in F,\ x_1 \leq 0\}.$$

Obviously, $\alpha_1 = \alpha_2 = 0$ iff $\alpha = 0$. The reason we transform QP into QP_1 and QP_2 is that a projected Hessian of C on $x_1 = 0$ is positive definite. Our purpose is to perturb $x_1 = 0$ a little bit to form vectors $\bar{h}$ and $\hat{h}$ such that $\{x \in F \mid x_1 \geq 0, (\bar{h})'x = 1\}$ and $\{x \in F \mid x_1 \leq 0, (\hat{h})'x = 1\}$ are nonempty and bounded and, $C_{\bar{h}}$ and $C_{\hat{h}}$ are positive semidefinite. In the following, we only consider how to construct $\bar{h}$ (as $\hat{h}$ can be constructed in same way).

We will proceed as follows. We state a complete algorithm for checking if C is copositive on $\{x \in F \mid x_1 \geq 0\}$ first. Then we will show that the algorithm is correct step by step.

Algorithm 4.2.1

Step 1. Let $r = e - \sum_{i=1}^{m} a_i$ and $F_3 = \{x \in F \mid x_1 = 0,\ r'x = 1\}$, where $e = (1, 0, \cdots, 0)' \in \mathrm{E}^n$. If $F_3 \neq \phi$, go to Step 2. Otherwise let $h = e$ and go to Step 3.

Step 2. Let

$$\epsilon = \min \{\frac{1}{|r_1| + 1}, \frac{1}{d_1 \sum_{i=2}^{n} |r_i|}\},\ h_1 = 1 \text{ and } h_i = \epsilon r_i$$

for $i = 2, \cdots, n$. Go to Step 3.

Step 3. Solve

$$\alpha_3 = \min \{x'Cx \mid x \in F,\ x_1 \geq 0,\ h'x = 1\}. \tag{4.4}$$

If $\alpha_3 \geq 0$ or (4.4) is infeasible, C is copositive on $\{x \in F \mid x_1 \geq 0\}$. Otherwise, C is not copositive on $\{x \in F \mid x_1 \geq 0\}$.

The fact that Algorithm 4.2.1 works for checking the copositivity of C on $\{x \in F \mid x_1 \geq 0\}$ is shown by the following 5 lemmas. First we assume $\{x \in F \mid x_1 \geq 0\} \neq \{0\}$. Later on we will discuss the case $\{x \in F \mid x_1 \geq 0\} = \{0\}$.

Lemma 4.2.1 *If $F_3 = \phi$ in Step 1, then by setting $h = e$, the set $\{x \in F \mid x_1 \geq 0,\ h'x = 1\}$ is nonempty and bounded and, a projected Hessian C_h of C on $h'x = 1$ is positive semidefinite.*

Proof. Since $\{x \in F \mid x_1 = 0,\ r'x = 1\} = \phi$, $\{x \in F \mid x_1 = 0\} = \{0\}$. Since $\{x \in F \mid x_1 \geq 0\} \neq \{0\}$, $\{x \in F \mid x_1 \geq 0,\ h'x = 1\} = \{x \in F \mid x_1 = 1\}$ is nonempty and bounded. The fact that a projected Hessian C_h of C on $h'x = 1$ is positive semidefinite is trivial. This completes the proof of the lemma. □

In the following, we consider the case $F_3 \neq \phi$ and assume $s^i \in \{x \in \mathrm{E}^n \mid x_1 = 0,\ r'x = 0\}$ for $i = 1, \cdots, n-2$ such that $s^1, \cdots, s^{n-2}$ are linearly independent.

Lemma 4.2.2 *For any $x^* \in F_3$, $-\epsilon e + x^*$, $-\epsilon e + x^* + s^1$, $\cdots$, $-\epsilon e + x^* + s^{n-2}$ are linearly independent.*

Proof. By adding ϵe to each of these vectors, we can see it is sufficient to show that x^*, $x^* + s^1, \cdots, x^* + s^{n-2}$ are linearly independent. For any real numbers $\lambda_1, \cdots, \lambda_{n-1}$ such that

$$\lambda_1 x^* + \lambda_2(x^* + s^1) + \cdots + \lambda_{n-1}(x^* + s^{n-2}) = 0,$$

we have

$$r'[\lambda_1 x^* + \lambda_2(x^* + s^1) + \cdots + \lambda_{n-1}(x^* + s^{n-2})] = \sum_{i=1}^{n-1} \lambda_i = 0.$$

Thus,

$$\lambda_2 s^1 + \cdots + \lambda_{n-1} s^{n-2} = 0.$$

Since $s^1, \cdots, s^{n-2}$ are linearly independent, $\lambda_2 = \cdots = \lambda_{n-1} = 0$ which implies $\lambda_1 = 0$. Hence $x^*, x^* + s^1, \cdots, x^* + s^{n-2}$ are linearly independent. This completes the proof of the lemma. □

Lemma 4.2.3 *For any $x^* \in F_3$, h is a solution to the following system*

$$\left.\begin{array}{l} h'(-\epsilon e + x^*) = 0 \\ h'(-\epsilon e + x^* + s^1) = 0 \\ \vdots \\ h'(-\epsilon e + x^* + s^{n-2}) = 0 \end{array}\right\} \tag{4.5}$$

Proof. Since $h_i = \epsilon r_i$ for $i = 2, \cdots, n$ and $x_1^* = 0$, we have $h'x^* = \epsilon r' x^* = \epsilon$. Similarly, since $s_1^i = 0$ for $i = 1, \cdots, n-2$, we have

$$h's^i = \epsilon r' s^i = 0 \quad \text{for } i = 1, \cdots, n-2.$$

Since $h_1 = 1$, $h'(\epsilon e) = \epsilon$. Hence h is a solution to (4.5) and the proof of the lemma is complete. □

Lemma 4.2.4 *The set $\{x \in F \mid x_1 \geq 0,\ h'x = 1\}$ is nonempty and bounded.*

Proof. Let $x^* \in F_3$ be any point. Since $h'x^* = \epsilon > 0$, we have

$$h'(\frac{1}{h'x^*}x^*) = 1.$$

This implies $\{x \in F \mid x_1 \geq 0,\ h'x = 1\}$ is nonempty. In order to show $\{x \in F \mid x_1 \geq 0,\ h'x = 1\}$ is bounded, it suffices to show

$$\{x \in F \mid x_1 \geq 0,\ h'x = 0\} = \{0\}. \tag{4.6}$$

Since $\{x \in F \mid x_1 \geq 0,\ r'x = 0\} = \{0\}$, for any $\hat{x} \in \{x \in F \mid x_1 \geq 0\}$ with $\hat{x} \neq 0$ we have $r'\hat{x} > 0$. Hence

$$r'(\frac{1}{r'\hat{x}}\hat{x}) = 1.$$

In order to establish (4.6), it is sufficient to show

$$\{x \in F \mid x_1 \geq 0,\ r'x = 1\} \bigcap \{x \in \mathrm{E}^n \mid h'x = 0\} = \phi. \tag{4.7}$$

If (4.7) does not hold, then there exists a vector y with

$$y \in \{x \in F \mid x_1 \geq 0,\ r'x = 1\} \bigcap \{x \in \mathrm{E}^n \mid h'x = 0\}.$$

Since h is a solution to (4.5) and $-\epsilon e + x^*$, $-\epsilon e + x^* + s^1$, ⋯, $-\epsilon e + x^* + s^{n-2}$ are linearly independent, there exist real numbers $\lambda_1, \cdots, \lambda_{n-1}$ such that

$$y = \lambda_1(-\epsilon e + x^*) + \sum_{i=1}^{n-2} \lambda_{i+1}(-\epsilon e + x^* + s^i). \tag{4.8}$$

There are two cases to consider: $y_1 = 0$ and $y_1 \neq 0$. If $y_1 = 0$, then from

$$y_1 = \lambda_1(-\epsilon) + \cdots + \lambda_{n-1}(-\epsilon) = (-\epsilon)\sum_{i=1}^{n-1} \lambda_i,$$

we have $\sum_{i=1}^{n-1} \lambda_i = 0$. In this case,

$$y = \sum_{i=1}^{n-1} \lambda_i x^* + \lambda_2 s^1 + \cdots + \lambda_{n-1} s^{n-2} = \lambda_2 s^1 + \cdots + \lambda_{n-1} s^{n-2}.$$

This implies $r'y = 0$. This contradicts $r'y = 1$. Hence we must have $y_1 \neq 0$. Since

$$y \in \{x \in F \mid x_1 \geq 0,\ r'x = 1\}, \tag{4.9}$$

we have $y_1 > 0$. Since $y_1 = (-\epsilon)\sum_{i=1}^{n-1}\lambda_i$, $\sum_{i=1}^{n-1}\lambda_i < 0$. From (4.8) and (4.9), we have

$$1 = r'y = -\epsilon r_1 \sum_{i=1}^{n-1}\lambda_i + \sum_{i=1}^{n-1}\lambda_i.$$

This implies

$$-\epsilon r_1 \sum_{i=1}^{n-1}\lambda_i + \sum_{i=1}^{n-1}\lambda_i - 1 = 0. \tag{4.10}$$

Since $\epsilon \leq 1/(|r_1| + 1)$,

$$-\epsilon r_1 \sum_{i=1}^{n-1}\lambda_i \leq \frac{|r_1|}{(|r_1|+1)}(-\sum_{i=1}^{n-1}\lambda_i) < -\sum_{i=1}^{n-1}\lambda_i.$$

So

$$-\epsilon r_1 \sum_{i=1}^{n-1}\lambda_i + \sum_{i=1}^{n-1}\lambda_i - 1 < -1.$$

This contradicts (4.10). Hence the validity of (4.7) is established and the proof of the lemma is complete. □.

Lemma 4.2.5 *A projected Hessian C_h of C on $h'x = 1$ is positive semidefinite.*

Proof. From the construction of h and ϵ, we have

$$\sum_{i=2}^{n}\frac{h_i^2}{d_i} \leq \sum_{i=2}^{n} h_i^2 = \epsilon^2 \sum_{i=2}^{n} r_i^2 \leq \frac{\sum_{i=2}^{n} r_i^2}{d_1^2(\sum_{i=2}^{n}|r_i|)^2} \leq \frac{1}{d_1^2} \leq \frac{h_1^2}{d_1}.$$

By Corollary 4.2.1, C_h is positive semidefinite, as required. □

In the above, we assumed that $\{x \in F \mid x_1 \geq 0\} \neq \{0\}$. If $\{x \in F \mid x_1 \geq 0\} = \{0\}$, then from Step 1, we have $F_3 = \phi$. So $h = e$ and in Step 3, (4.4) will be infeasible. The algorithm will then tell us that C is copositive on $\{x \in F \mid x_1 \geq$

$0\}$ and the algorithm does indeed give the correct information. In fact, (4.4) is infeasible iff $\{x \in F \mid x_1 \geq 0\} = \{0\}$.

In order to use Algorithm 4.2.1 to check if C is copositive on $\{x \in \mathrm{E}^n \mid Ax \leq 0,$ $x_1 \leq 0\}$, we only need to do the following. Let

$$\bar{a}_{i1} = -a_{i1} \text{ and } \bar{a}_{ij} = a_{ij}$$

for $i = 1, \cdots, m$ and $j = 2, \cdots, n$. Then checking if C is copositive on $\{x \in \mathrm{E}^n \mid Ax \leq 0,\ x_1 \leq 0\}$ is equivalent to checking if C is copositive on $\{x \in \mathrm{E}^n \mid \bar{A}x \leq 0,$ $x_1 \geq 0\}$, where $\bar{A} = [\bar{a}_{ij}]$. The later can be solved by using Algorithm 4.2.1 directly.

From Algorithm 4.2.1, we can see that h is computed by solving one feasibility problem. So h can be computed in polynomial time and the size of h can be bounded by a polynomial function of l, where l is the size for A, C and e. Since (4.4) is a convex quadratic programming problem, it can be solved in polynomial time and we have the following.

Theorem 4.2.1 *The copositivity problem formulated in this section can be solved in polynomial time.*

The following example is taken from [11] with a slight modification. The original problem is considered in Example 4.2.4.

Example 4.2.3 Let $n = 3$, $m = 5$, and consider

$$C = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } A = \begin{bmatrix} 2 & 0 & -5 & 0 & 2 \\ 1 & 5 & 0 & -5 & -1 \\ -2 & -4 & -3 & -4 & -2 \end{bmatrix}'.$$

Let us use Algorithm 4.2.1 to determine if C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \leq 0,$ $x_1 \geq 0\}$ first.

Step 1: $r = e - \sum_{i=1}^{5} a_i = (2,0,15)'$ and $(0,0,1/15)'$ is a feasible solution to $\{x \in \mathrm{E}^3 \mid Ax \le 0,\ x_1 = 0,\ r'x = 1\}$. Go to Step 2.

Step 2: $\epsilon = \min\{1/3, 1/15\} = 1/15$, $h_1 = 1$, $h_2 = 0$ and $h_3 = 1$. Go to Step 3.

Step 3: By setting $x_3 = 1 - x_1$,

$$\alpha_3 = \min\ \{x'Cx \mid Ax \le 0,\ x_1 \ge 0,\ h'x = 1\}$$

becomes

$$\alpha_3 = \min\ \{1 - 2x_1 + x_2^2 \mid \hat{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \le \hat{b}\}, \tag{4.11}$$

where

$$\hat{A} = \begin{bmatrix} 4 & 4 & -2 & 4 & 4 & -1 \\ 1 & 5 & 0 & -5 & -1 & 0 \end{bmatrix}' \text{ and } \hat{b} = (2,4,3,4,2,0)'.$$

By solving (4.11), we have $\alpha_3 = 0$. So, C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \le 0,$ $x_1 \ge 0\}$. Now let us determine if C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \le 0,\ x_1 \le 0\}$. By setting $\bar{a}_{i1} = -a_{i1}$ and $\bar{a}_{ij} = a_{ij}$ for $i = 1, \cdots, 5$ and $j = 2, 3$, it is sufficient to determine if C is copositive on $\{x \in \mathrm{E}^3 \mid \bar{A}x \le 0,\ x_1 \ge 0\}$, where $\bar{A} = [\bar{a}_{ij}]$. Similar as the above, Algorithm 4.2.1 tells us that C is copositive on $\{x \in \mathrm{E}^3 \mid \bar{A}x \le 0,$ $x_1 \ge 0\}$; i.e., C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \le 0,\ x_1 \le 0\}$. Hence C is copositive on $\{x \in \mathrm{E}^n \mid Ax \le 0\}$.

In what follows, we consider the case that C has two negative eigenvalues; i.e., $C = \mathrm{diag}(-d_1, -d_2, d_3, \cdots, d_n)$, where d_i is positive integer for $i = 1, \cdots, n$. From Proposition 4.2.2, we get the following corollary.

Corollary 4.2.2 *For a given vector $h \in E^n \setminus \{0\}$, a projected Hessian C_h of C on $h'x = 1$ has exactly one negative eigenvalue iff*

$$\frac{h_3^2}{d_3} + \cdots + \frac{h_n^2}{d_n} \le \frac{h_1^2}{d_1} + \frac{h_2^2}{d_2}.$$

From Corollaries 4.2.1 and 4.2.2, we can see that we can still use Algorithm 4.2.1 to solve this copositivity problem. But now (4.4) is a non-convex quadratic programming problem with Hessian having exactly one negative eigenvalue. So, it can be solved by the method of Section 2.3 of Chapter 2. By the generalized conjugate direction algorithm developed in Section 3.5 of Chapter 3, we can write down the parametric formulation for (4.4). However since (4.4) has a special structure, we can write down the parametric formulation for (4.4) directly as follows. From $h'x = 1$, we have

$$x_1 = \frac{1}{h_1} - \sum_{i=2}^{n} \frac{h_i}{h_1} x_i.$$

So,

$$\begin{aligned}
&-d_1x_1^2 - d_2x_2^2 + \sum_{i=3}^{n} d_ix_i^2 \\
&= -d_1\left(\frac{1}{h_1} - \sum_{i=2}^{n} \frac{h_i}{h_1}x_i\right)^2 - d_2x_2^2 + \sum_{i=3}^{n} d_ix_i^2 \\
&= \psi(x_2) + (\bar{c} + x_2q)'\bar{x} + \bar{x}'\bar{C}\bar{x},
\end{aligned}$$

where

$$\begin{aligned}
\bar{x} &= (x_3, \cdots, x_n)', \\
\bar{c} &= \left(\frac{2d_1h_3}{h_1}, \cdots, \frac{2d_1h_n}{h_1}\right)', \\
q &= \left(\frac{-2d_1h_2h_3}{h_1^2}, \cdots, \frac{-2d_1h_2h_n}{h_1^2}\right)', \\
\psi(x_2) &= -\frac{d_1}{h_1^2} + \frac{2d_1h_2}{h_1^2}x_2 - \left(\frac{d_1h_2^2}{h_1^2} + d_2\right)x_2^2, \\
\bar{C} &= \operatorname{diag}(d_3, \cdots, d_n) - \left(\frac{d_1h_3}{h_1}, \cdots, \frac{d_1h_n}{h_1}\right)'\left(\frac{h_3}{h_1}, \cdots, \frac{h_n}{h_1}\right).
\end{aligned}$$

From $Ax \leq 0$, we have

$$\left\{ \begin{bmatrix} a_{13} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m3} & \cdots & a_{mn} \end{bmatrix} - \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} \left(\frac{h_3}{h_1}, \cdots, \frac{h_n}{h_1}\right) \right\} \bar{x}$$

$$\leq \begin{bmatrix} -a_{11} \\ \vdots \\ -a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} \frac{a_{11}h_2}{h_1} - a_{12} \\ \vdots \\ \frac{a_{m1}h_2}{h_1} - a_{m2} \end{bmatrix}.$$

From $x_1 \geq 0$ and $h_1 > 0$, we have

$$(\frac{h_3}{h_1}, \cdots, \frac{h_n}{h_1})\bar{x} \leq \frac{1}{h_1} + x_2(-\frac{h_2}{h_1}).$$

So, (4.4) can be written as

$$\begin{aligned} &\min \psi(x_2) + (\bar{c} + x_2 q)'\bar{x} + \bar{x}'\bar{C}\bar{x} \\ &\text{s.t.} \qquad \bar{A}\bar{x} \leq \bar{b} + x_2 p, \end{aligned}$$

where

$$\bar{A} = \begin{bmatrix} a_{13} - \frac{a_{11}h_3}{h_1} & \cdots & a_{1n} - \frac{a_{11}h_n}{h_1} \\ \vdots & & \vdots \\ a_{m3} - \frac{a_{m1}h_3}{h_1} & \cdots & a_{mn} - \frac{a_{m1}h_n}{h_1} \\ \\ \frac{h_3}{h_1} & \cdots & \frac{h_n}{h_1} \end{bmatrix},$$

$$\bar{b} = (-a_{11}, \cdots, -a_{m1}, \frac{1}{h_1})'$$

and

$$p = (\frac{a_{11}h_2}{h_1} - a_{12}, \cdots, \frac{a_{m1}h_2}{h_1} - a_{m2}, -\frac{h_2}{h_1})'.$$

From Algorithm 4.2.1, we know

$$\frac{h_2^2}{d_2} + \cdots + \frac{h_n^2}{d_n} \leq \frac{h_1^2}{d_1}.$$

So,

$$\frac{h_3^2}{d_3} + \cdots + \frac{h_n^2}{d_n} \leq \frac{h_1^2}{d_1}.$$

Therefore, $\bar{C}$ is positive semidefinite. This implies that the above problem is a convex parametric quadratic programming problem if we take x_2 as a parameter. Hence it can be solved efficiently by the method of Section 2.3 of Chapter 2.

Now let us see one more example to illustrate the above ideas. This example is taken from [11].

Example 4.2.4 Let $n = 3$, $m = 5$, and consider

$$C = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & 0 & -5 & 0 & 2 \\ 1 & 5 & 0 & -5 & -1 \\ -2 & -4 & -3 & -4 & -2 \end{bmatrix}'.$$

This example is almost the same as Example 4.2.3 except $c_{22} = -1$ here, the element in the second row and column of C. As in Example 4.2.3, let us determine if C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \leq 0,\ x_1 \geq 0\}$ first. In Example 4.2.3, we know $h = (1, 0, 1)'$. By the above discussion, we know that this h can be used to obtain a parametric formulation for (4.4). In fact, by setting $x_1 = 1 - x_3$, (4.4) can be transformed to

$$\begin{array}{ll} \min & -1 + 2x_3 - x_2^2 \\ \text{s.t.} & x_2 - 4x_3 \leq -2,\ 5x_2 - 4x_3 \leq 0, \\ & 2x_3 \leq 5,\ -5x_2 - 4x_3 \leq 0, \\ & -x_2 - 4x_3 \leq -2,\ x_3 \leq 1. \end{array}$$

By taking x_2 as a parameter, the problem becomes a parametric linear programming problem. The optimal value of the problem is zero. So, C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \leq 0,\ x_1 \geq 0\}$. Now let us determine if C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \leq 0,\ x_1 \leq 0\}$. By setting $\bar{a}_{i1} = -a_{i1}$ and $\bar{a}_{ij} = a_{ij}$ for $i = 1, \cdots, 5$ and $j = 2, 3$, it is sufficient to determine if C is copositive on $\{x \in \mathrm{E}^3 \mid \bar{A}x \leq 0,\ x_1 \geq 0\}$, where $\bar{A} = [\bar{a}_{ij}]$. Similar as the above, Algorithm 4.2.1 tells us that C is copositive on

$\{x \in \mathrm{E}^3 \mid \bar{A}x \leq 0,\ x_1 \geq 0\}$; i.e., C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \leq 0,\ x_1 \leq 0\}$. Hence C is copositive on $\{x \in \mathrm{E}^3 \mid Ax \leq 0\}$.

So far, we have established an algorithm for a class of copositivity problems in which C is a diagonal integral matrix with exactly one or two negative eigenvalues. Therefore, in order to solve a class of copositivity problems in which C is an (n,n) symmetric matrix with exactly one or two negative eigenvalues, it is sufficient to transform the problem into the formulation of this section. This will be discussed in the next section.

4.3 A Polynomial Transformation

In this section, we assume that both A and C are integral matrices, $\mathrm{rank}(A) = \mathrm{rank}(C) = n$ and C has exactly one negative eigenvalue. Here C need not be a diagonal matrix. The purpose of this section is to explain how to transform the copositivity problem into the formulation of Section 4.2 in polynomial time.

If a symmetric matrix is positive definite matrix, we know that it can be factored into LDL', where D is a diagonal matrix and L is a unit lower triangular matrix. However when the matrix is not positive definite, the factorization may not work as the determinant of some principal minor of the matrix may equal to zero. In what follows, we will give an algorithm to diagonalize C first. The algorithm is developed baded on some modifications of LDL' factorization. Now let us start the algorithm. First of all, let us introduce the concept of the unit triangular matrix.

Definition 4.3.1 An (n,n) matrix $L = [l_{ij}]$ is called a unit lower triangular matrix if

$$l_{ij} = \begin{cases} 0 & \text{for } j > i,\ i = 1,\cdots,n \text{ and } j = 2,\cdots,n, \\ 1 & \text{for } i = j,\ i = 1,\cdots,n. \end{cases}$$

We attempt to factor C into LDL', where $D = \text{diag}(d_{11}, \cdots, d_{nn})$ and L is a unit lower triangular matrix. Let C_j denote the j-th principal minor of C for $j = 1, \cdots, n$. If $\det(C_j) \neq 0$ for all j with $1 \leq j \leq n$, the process of factoring of C into LDL' can be accomplished. In this case, L and D can be computed in polynomial time and, the size of L and D are bounded by a polynomial function of the size of C. The reader may refer Chapter 2 of [31] for details. In the following, let us consider how to deal with the case that there is a j with $1 \leq j \leq n$ such that $\det(C_j) = 0$.

Let k be the smallest number such that $\det(C_k) = 0$. Then the factorization of C into LDL' can not be continued after $k-1$ iterations. In this case, we partition C as follows

$$C = \begin{bmatrix} C_{k-1} & B \\ B' & E \end{bmatrix}.$$

Let

$$\begin{aligned} \hat{C} &\equiv \begin{bmatrix} I_{k-1} & 0 \\ -B'C_{k-1}^{-1} & I_{n-k+1} \end{bmatrix} \begin{bmatrix} C_{k-1} & B \\ B' & E \end{bmatrix} \begin{bmatrix} I_{k-1} & -C_{k-1}^{-1}B \\ 0 & I_{n-k+1} \end{bmatrix} \\ &= \begin{bmatrix} C_{k-1} & 0 \\ 0 & E - B'C_{k-1}^{-1}B \end{bmatrix}, \end{aligned}$$

where I_i is an (i,i) identity matrix for $i = 1, \cdots, n$. Let $\hat{C} = [\hat{c}_{ij}]$. Since $\det(C_k) = 0$, we have $\hat{c}_{kk} = 0$. Since C is nonsingular, $\hat{C}$ is nonsingular. So, there exists a $j \geq k+1$ such that $\hat{c}_{kj} \neq 0$. Define

$$z = \begin{cases} -\frac{|\hat{c}_{jj}|}{\hat{c}_{kj}} e_k + e_j & \text{if } \hat{c}_{jj} \neq 0, \\ -\frac{1}{\hat{c}_{kj}} e_k + e_j & \text{otherwise,} \end{cases}$$

where $e_i \in \mathrm{E}^n$ is a unit vector with i-th entry being equal to one for $i = 1, \cdots, n$.

It is easy to check that $z'\hat{C}z < 0$. By setting

$$\hat{y} = \begin{bmatrix} I_{k-1} & -C_{k-1}^{-1}B \\ 0 & I_{n-k+1} \end{bmatrix} z,$$

we have $\hat{y}'C\hat{y} < 0$. Since $\hat{y} \neq 0$, it must have at least one nonzero component. For simplicity, suppose $\hat{y}_1 \neq 0$. Set

$$p_i = e_i - \frac{e_i'C\hat{y}}{\hat{y}'C\hat{y}}\hat{y} \text{ for } i = 2, \cdots, n$$

and $P = (\hat{y}, p_2, \cdots, p_n)$. Then P is nonsingular and

$$P'CP = \begin{bmatrix} \hat{y}'C\hat{y} & 0 \\ 0 & F \end{bmatrix},$$

where F is an $(n-1, n-1)$ symmetric matrix. Since $\text{rank}(C) = n$ and C has exactly one negative eigenvalue, F is positive definite. So F can be factored into $\bar{L}\bar{D}\bar{L}'$, where $\bar{L}$ is an $(n-1, n-1)$ unit triangular matrix and $\bar{D} = \text{diag}(\bar{d}_{11}, \cdots, \bar{d}_{nn})$ with $d_{ii} > 0$ for $i = 1, \cdots, n$. Thus,

$$P'CP = \begin{bmatrix} 1 & 0 \\ 0 & \bar{L} \end{bmatrix} \begin{bmatrix} \hat{y}'C\hat{y} & 0 \\ 0 & \bar{D} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \bar{L}' \end{bmatrix}.$$

Let

$$\hat{y}'C\hat{y} = -\frac{\lambda_1}{\beta_1}, \ \bar{d}_{ii} = \frac{\lambda_{i+1}}{\beta_{i+1}} \text{ for } i = 1, \cdots, n-1,$$

where λ_i and β_i are positive integers for $i = 1, \cdots, n$. Then by setting

$$M = P \begin{bmatrix} 1 & 0 \\ 0 & \bar{L}' \end{bmatrix}^{-1} \text{diag}(\beta_1, \cdots, \beta_n)$$

and $d_i = \lambda_i \beta_i$ for $i = 1, \cdots, n$, we have

$$M'CM = \text{diag}(-d_1, d_2, \cdots, d_n).$$

Now, by putting $x = My$, it follows that

$$C \text{ is copositive on } \{x \in \mathrm{E}^n \mid Ax \leq 0\} \tag{4.12}$$

if and only if

$$\mathrm{diag}(-d_1, d_2, \cdots, d_n) \text{ is copositive on } \{x \in \mathrm{E}^n \mid (AM)y \leq 0\}. \tag{4.13}$$

Note (4.13) is precisely the formulation of the copositivity problem discussed in Section 4.2. Since M is computed by matrix inverse with matrix multiplications and LDL' factorizations, M can be computed in polynomial time and the size of M is bounded by a polynomial function of the size of C. Therefore the size of (4.13) is bounded by a polynomial function of the size of (4.12). Now by combining this with Theorem 4.2.1, we have the following result.

Theorem 4.3.1 *Assume that rank(A) = rank$(C) = n$. If C has exactly one negative eigenvalue, then determining if C is copositive on $\{x \in E^n \mid Ax \leq 0\}$ can be done in polynomial time.*

If C has exactly two negative eigenvalues, it can also be diagonalized by the algorithm similar to the above. It can also be diagonalized by the general conjugate direction algorithm developed in Section 3.5 of Chapter 3. So the problem can be transformed into the formulation of Section 4.2. Hence this type of problem can be efficiently solved by parametric quadratic programming techniques.

We conclude this section with two more examples to illustrate the above transformation process.

Example 4.3.1 Let

$$C = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

and diagonalize C by the above transformation process.

By putting

$$C = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} 1 & l_{21} & l_{31} \\ 0 & 1 & l_{32} \\ 0 & 0 & 1 \end{bmatrix}$$

we have $d_{11} = 1$, $l_{21} = 2$, $l_{31} = 1$, $d_{22} = -3$, $l_{32} = 1/3$, $d_{33} = 4/3$. So, $d_1 = 1$, $d_2 = 3$, $d_3 = 12$, $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = 3$,

$$M = (L')^{-1} \operatorname{diag}(\beta_1, \beta_2, \beta_3) = \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

and

$$M'CM = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 12 \end{bmatrix}.$$

Example 4.3.2 Let

$$C = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \\ 2 & 3 & 2 \end{bmatrix}$$

and diagonalize C by the above transformation process.

By the process of factoring C into LDL', we obtain $\det(C_2) = 0$. So, we have $k = 2$, $C_1 = [1]$, $B = (1, 2)'$ and

$$E = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}.$$

Now, let us summarize our computations as follows:

$$\hat{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -3 \end{bmatrix}, \; z = -\frac{|\hat{c}_{33}|}{\hat{c}_{23}} e_2 + e_3 = \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}, \; \hat{y} = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix},$$

$$P = \begin{bmatrix} 1 & 1/9 & -2/3 \\ -3 & 2/3 & 2 \\ 1 & 1/9 & 1/3 \end{bmatrix}, \; F = \begin{bmatrix} 10/9 & 7/3 \\ 7/3 & 5 \end{bmatrix}, \; \bar{L} = \begin{bmatrix} 1 & 0 \\ -21/10 & 1 \end{bmatrix},$$

$$d_1 = 9,\; d_2 = 90,\; d_3 = 10,\; \beta_1 = 1,\; \beta_2 = 9,\; \beta_3 = 10,$$

$$M = P \begin{bmatrix} 1 & 0 \\ 0 & \bar{L}' \end{bmatrix}^{-1} \operatorname{diag}(\beta_1, \beta_2, \beta_3) = \begin{bmatrix} 1 & 1 & -9 \\ -3 & 6 & 6 \\ 1 & 1 & 1 \end{bmatrix},$$

$$M'CM = \begin{bmatrix} 1 & -3 & 1 \\ 1 & 6 & 1 \\ -9 & 6 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -9 \\ -3 & 6 & 6 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -9 & 0 & 0 \\ 0 & 90 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

4.4 Numerical Results

We obtained numerical results by taking $C = \operatorname{diag}(-1, 1, \cdots, 1)$, and A, an (m, n) matrix with randomly generated elements. The code was executed on a 486-66 PC and the computation time is measured in seconds. The results of applying our method to such data are summarized in Table 4.1. The quadratic programming algorithm used is that of Best and Ritter [10] as implemented in [9].

Table 4.1: Numerical Experiments for Algorithm 4.2.1

m	30	70	110	150	190	250	300	350	400
n	15	35	55	75	95	125	150	175	200
time	0.05	0.09	0.10	0.12	0.16	0.22	0.39	0.43	0.57

4.5 Conclusion

We have established a polynomial algorithm for a class of copositivity problem in which C has one negative eigenvalue and $n-1$ positive eigenvalues. The algorithm is extended to solve a class of copositivity problems in which C has two negative eigenvalues and $n-2$ positive eigenvalues. The algorithm is illustrated by numerical examples and randomly generated data.

Chapter 5

Global and Local Non-convex Minimization

5.1 Introduction

In this chapter, we generalize the results established in Chapter 2 to a large class of non-convex minimization problems. Here we consider the following nonlinear programming model

NP $\quad \min \{g(x,t) \mid (x,t) \in \Omega,\ g_j(x,t) \leq 0,\ j = 1, \cdots, m\},$

where $g(\cdot,\cdot),\ g_j(\cdot,\cdot) : \mathrm{E}^{n+k} \to \mathrm{E}^1 \bigcup \{-\infty\} \bigcup \{+\infty\}$ for $j = 1, \cdots, m$ and $\Omega \subseteq \mathrm{E}^{n+k}$. Let $R = \{(x,t) \in \Omega \mid g_j(x,t) \leq 0,\ j = 1, \cdots, m\}$ and $R(t) = \{x \in \mathrm{E}^n \mid (x,t) \in \Omega,\ g_j(x,t) \leq 0,\ j = 1, \cdots, m\}$ for each $t \in \mathrm{E}^k$.

This model is quite similar to the model used by Geoffrion [16].

If $g(\cdot,t)$ is quasi-convex or convex, then, we may solve NP by solving the following main nonlinear programming problem

MNP $\qquad \min \{f(t) \mid t \in \mathrm{E}^k\},$

where

$$f(t) = \begin{cases} \inf \{g(x,t) \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

This approach was used in Chapters 2 and 3 for solving non-convex quadratic programming problems. Also see Best and Ding [6] and [7]. Related approaches were also used by Geoffrion [16], Kough [18] and Benders [2]. We will discuss these later.

In order to find all isolated local minima, some non-isolated local minima and the global minimum for NP, we must evaluate $f(t)$ and then solve MNP. We refer to this as our *parametric local optimization procedure.*

From the definition of f, we can see that the connection between NP and MNP is the following parametric nonlinear programming problem

NP(t) $\qquad \min \{g(x,t) \mid x \in R(t)\} \quad \text{for } t \in \mathrm{E}^k.$

Let arg min{NP(t)} denote the set of all optimal solutions for NP(t). Since we want to solve NP by solving MNP, it is necessary to study the relationships between NP and MNP. For example, can MNP represent NP or how much information can MNP retain from NP? We will answer these questions partly in this paper. First, let us see some examples.

Example 5.1.1 Consider the indefinite quadratic problem

$$\min \{x'x - y'y \mid Ax + By \leq b\},$$

where $x \in \mathrm{E}^n$, $y \in \mathrm{E}^k$, $A \in \mathrm{E}^{m\times n}$, $B \in \mathrm{E}^{m\times k}$ and $b \in \mathrm{E}^m$.

This problem was studied by Kough [18]. By setting $t = y$, $g(x,t) = x'x - t't$ and $R(t) = \{x \in \mathrm{E}^n \mid Ax + Bt \leq b\}$, MNP and NP$(t)$ can be formulated accordingly.

NP(t) and $f(t)$ can be written as follows

$$\text{NP}(t) \qquad \min\ \{x'x \mid x \in R(t)\} \quad \text{for } t \in \mathrm{E}^k;$$

$$f(t) = t't + \begin{cases} \inf\{x'x \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

In this example, NP(t) is a multi-parametric convex quadratic programming problem.

Example 5.1.2 Consider the nonlinear programming problem

$$\min\ \{c'x + \psi(y) \mid a_j'x + g_j(y) + b_j \leq 0,\ j = 1, \cdots, m,\ x \geq 0,\ y \in S\},$$

where $c \in \mathrm{E}^n$, $\psi(\cdot) : \mathrm{E}^k \to \mathrm{E}^1$, $a_j \in \mathrm{E}^n$, $g_j(\cdot) : \mathrm{E}^k \to \mathrm{E}^1$, $b_j \in \mathrm{E}^1$ for $j = 1, \cdots, m$ and S is a subset of E^k.

This model was studied by Benders [2]. Obviously, for each fixed y, the problem is a linear programming. So, we may put $t = y$, $g(x,t) = c'x + \psi(t)$ and $R(t) = \{x \in \mathrm{E}^n \mid a_j'x + g_j(y) + b_j \leq 0, i = 1, \cdots, m, x \geq 0\}$ for each $t \in S$, and formulate NP(t) and MNP accordingly. NP(t) and $f(t)$ can be written as follows:

$$\text{NP}(t) \qquad \min\ \{c'x \mid x \in R(t)\} \quad \text{for } t \in \mathrm{E}^k,$$

$$f(t) = \psi(t) + \begin{cases} \inf\ \{c'x \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

The problem is a nonlinear programming problem, but NP(t) is a linear programming. Hence it makes the problem much easier to solve in some sense.

Example 5.1.3 Consider the nonlinear programming problem

$$\min\ \{\bar{g}(x) + \hat{g}(y) \mid \bar{g}_j(x) + \hat{g}_j(y) \leq 0 \text{ for } j = 1, \cdots, m,\ (x,y) \in \Omega\},$$

where Ω is a subset of E^{n+k}, $\bar{g}(\cdot)$, $\bar{g}_j(\cdot) : \mathrm{E}^n \to \mathrm{E}^1$, $\hat{g}(\cdot)$, $\hat{g}_j(\cdot) : \mathrm{E}^k \to \mathrm{E}^1$, $\bar{g}$ and $\bar{g}_j$ are convex on E^n for $j = 1, \cdots, m$.

Obviously, for each fixed y, the problem is a convex programming problem. So, we may set $t = y$, $g(x,t) = \bar{g}(x) + \hat{g}(t)$ and $R(t) = \{x \in \mathrm{E}^n \mid \bar{g}_j(x) + \hat{g}_j(y) \le 0, j = 1, \cdots, m, (x,t) \in \Omega\}$, and formulate NP($t$) and MNP accordingly. NP(t) and $f(t)$ can be written as follows

$$\text{NP}(t) \qquad \min\{\bar{g}(x) \mid x \in R(t)\} \quad \text{for } t \in \mathrm{E}^k;$$

$$f(t) = \hat{g}(t) + \begin{cases} \inf\{\bar{g}(x) \mid x \in R(t)\}, & \text{if } R(t) \ne \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

Note that Example 5.1.3 is a general formulation of Examples 5.1.1 and 5.1.2.

We will organize this paper as follows. In Section 5.2, we will develop the relationships between NP, NP(t) and MNP. In particular, we will establish a one to one correspondence between isolated local minimizers of NP and MNP for a large class of non-convex programming problems. In Section 5.3, we will discuss how to apply the results established in Section 5.2 to some special class of non-convex problems.

5.2 The Relationships Between NP and MNP

In this section, we will discuss some relationships between NP and MNP. We will generalize some results of Chapter 2 and also present some new results. First of all we will give a result concerning convexity.

Proposition 5.2.1 *If Ω is a convex subset of E^{n+k}, g_j $(j = 1, \cdots, m)$ is a convex vector function on E^{n+k}, then (i) g convexity on E^{n+k} implies f convexity on E^k and (ii) g quasi-convexity on E^{n+k} implies f quasi-convexity on E^k.*

Proof. The proofs of (i) and (ii) are quite similar, so we only prove (i) here. Let (t^1, γ^1) and (t^2, γ^2) be any two points in $\{(t, \gamma) \in \mathrm{E}^{k+1} \mid f(t) \leq \gamma,\ t \in \mathrm{E}^k, \gamma \in \mathrm{E}^1\}$. We need to show $f(\alpha t^1 + (1-\alpha)t^2 \leq \alpha\gamma^1 + (1-\alpha)\gamma^2$ for all α with $0 < \alpha < 1$. For any $\epsilon > 0$, since

$$f(t^1) < \gamma^1 + \epsilon \text{ and } f(t^2) < \gamma^2 + \epsilon,$$

there exist two points $x^1 \in R(t^1)$ and $x^2 \in R(t^2)$ such that

$$g(x^1, t^1) < \gamma^1 + \epsilon \text{ and } g(x^2, t^2) < \gamma^2 + \epsilon.$$

Since g_j $(j = 1, \cdots, m)$ is convex and Ω is also convex, it is easy to check $\alpha x^1 + (1-\alpha)x^2 \in R(\alpha t^1 + (1-\alpha)t^2)$. So,

$$\begin{aligned} f(\alpha t^1 + (1-\alpha)t^2) &\leq g(\alpha x^1 + (1-\alpha)x^2, \alpha t^1 + (1-\alpha)t^2) \\ &\leq \alpha g(x^1, t^1) + (1-\alpha)g(x^2, t^2) \leq \alpha\gamma^1 + (1-\alpha)\gamma^2 + \epsilon. \end{aligned}$$

Since ϵ is any positive number, we have

$$f(\alpha t^1 + (1-\alpha)t^2) \leq \alpha f(t^1) + (1-\alpha)f(t^2).$$

The proof of the proposition is thus complete. □

The following result is concerned about one to one correspondence between the global minimizer of NP and MNP.

Theorem 5.2.1 *If a point $t^* \in E^k$ with $f(t^*) > -\infty$ is a global minimizer of MNP, then for any $x^* \in$ arg min$\{NP(t^*)\}$, (x^*, t^*) is a global minimizer for NP. Conversely, if (x^*, t^*) is a global minimizer for NP, then t^* is a global minimizer for MNP.*

In the following, we will establish one to one correspondence results between isolated local minima of NP and MNP for several classes of non-convex programming problems. The results will be formulated in Theorems 5.2.2–5.2.5 and will be the consequences of Propositions 5.2.2–5.2.8, following.

Proposition 5.2.2 *If t^* is a local minimizer for MNP, then for any $x^* \in$ arg min$\{NP(t^*)\}$, (x^*, t^*) is a local minimizer for NP.*

Proof. Since t^* is a local minimizer for MNP, there exists an $\epsilon > 0$ such that

$$f(t) \geq f(t^*) \quad \text{for each } t \in B_\epsilon(t^*), \tag{5.1}$$

where $B_\epsilon(t^*) = \{t \in E^k \mid \|t - t^*\| \leq \epsilon\}$. Assume to the contrary, that (x^*, t^*) is not a local minimizer for NP. Then there exists a sequence $\{(x^i, t^i)\} \subset R$ with $(x^i, t^i) \to (x^*, t^*)$ and $(x^i, t^i) \neq (x^*, t^*)$ for all i satisfying

$$g(x^i, t^i) < g(x^*, t^*) = f(t^*). \tag{5.2}$$

Since $(x^i, t^i) \in R$ for each i, $x^i \in R(t^i)$. Since $t^i \to t^*$, there is a $M > 0$ such that $t^i \in B_\epsilon(t^*)$ when $i \geq M$. Now from $(x^i, t^i) \in R$ and (5.2), we have $f(t^i) < f(t^*)$ for each i, this contradicts (5.1) for $i \geq M$. The proof of the proposition is thus complete. □

Proposition 5.2.3 *Assume that t^* is an isolated local minimizer for MNP. For each $x^* \in$ arg min$\{NP(t^*)\}$, if there is a $\delta > 0$ such that $\{x \in B_\delta(x^*) \bigcap R(t^*) \mid g(x, t^*) = f(t^*)\} = \{x^*\}$, then (x^*, t^*) is an isolated local minimizer for NP, where $B_\delta(x^*) = \{x \in E^n \mid \|x - x^*\| \leq \delta\}$.*

Proof. From Proposition 5.2.2, (x^*, t^*) is a local minimizer for NP. Assume to the contrary, that (x^*, t^*) is not an isolated local minimizer. Then there exists a sequence $\{(x^i, t^i)\} \subset R$ with $(x^i, t^i) \to (x^*, t^*)$ and $(x^i, t^i) \neq (x^*, t^*)$ for all i satisfying

$$g(x^i, t^i) = g(x^*, t^*). \tag{5.3}$$

If $\{i \mid t^i \neq t^*,\ i = 1, 2, \cdots\}$ is finite, then there is a $M > 0$ such that $t^i = t^*$ for all $i \geq M$. From $x^i \to x^*$, there exists a $N > 0$ such that $x^i \in B_\delta(x^*)$ for all $i \geq N$. Since $\{x \in B_\delta(x^*) \bigcap R(t^*) \mid g(x^*, t^*) = f(t^*)\} = \{x^*\}$, (5.3) implies $x^i = x^*$ for all $i \geq \max\{M, N\}$, a contradiction. So $\{i \mid t^i \neq t^*,\ i = 1, 2, \cdots\}$ is infinite. Without loss of generality, let $t^i \neq t^*$ for all i. Again from (5.3), we have

$$f(t^i) \leq g(x^i, t^i) = g(x^*, t^*) = f(t^*) \text{ for all } i. \tag{5.4}$$

But $t^i \to t^*$, and thus (5.4) is in contradiction to t^* being an isolated local minimizer. Hence (x^*, t^*) must be an isolated local minimizer for NP. The proof of the proposition is thus complete. □

Lemma 5.2.1 *Assume that $R(t)$ is convex for each $t \in E^k$ and $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. If (x^*, t^*) is an isolated local minimizer for NP, then there exists a $\delta > 0$ such that*

$$g(x^*, t^*) < g(x, t) \quad \textit{for each } (x, t) \in R \bigcap B_\delta(x^*, t^*) \setminus \{x^*, t^*\} \tag{5.5}$$

and

$$g(x^*, t^*) < g(x, t^*) \quad \textit{for each } x \in R(t^*) \setminus \{x^*\}; \tag{5.6}$$

i.e., arg min$\{NP(t^*)\} = \{x^*\}$.

Proof. (5.5) is exactly the definition of an isolated local minimizer for NP. From (5.5), we have

$$g(x^*, t^*) < g(x, t^*) \text{ for each } (x, t^*) \in R \bigcap B_\delta(x^*, t^*) \setminus \{x^*, t^*\}.$$

This implies

$$g(x^*, t^*) < g(x, t^*) \text{ for each } x \in R(t^*) \bigcap B_\delta(x^*) \setminus \{x^*\}.$$

Since g is quasi-convex on $R(t^*)$ and $R(t^*)$ is convex, we have

$$g(x^*, t^*) < g(x, t^*) \text{ for each } x \in R(t^*) \setminus \{x^*\};$$

i.e., (5.6) holds and $\arg\min\{\text{NP}(t^*)\} = \{x^*\}$. This completes the proof of the lemma. □

Proposition 5.2.4 *Assume that g and g_j $(j = 1, \cdots, m)$ are continuous on Ω, $X \equiv \{x \in E^n \mid (x,t) \in \Omega,\ g_j(x,t) \leq 0, j = 1, \cdots, m\}$ is bounded, Ω is closed, $R(t)$ is convex for each $t \in E^k$ and $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. If (x^*, t^*) is an isolated local minimizer for NP, then t^* is an isolated local minimizer for MNP and $\arg\min\{NP(t^*)\} = \{x^*\}$.*

Proof. By Lemma 5.2.1, there exists a $\delta > 0$ such that (5.5) and (5.6) hold. Now assume that on the contrary, t^* is not an isolated local minimizer for MNP. Then there exists a sequence $\{t^i\} \subset \mathrm{E}^k$ with $t^i \neq t^*$ for all i and $t^i \to t^*$ such that

$$f(t^i) \leq f(t^*). \tag{5.7}$$

From (5.7) and the definition of f, we have $R(t^i) \neq \phi$ for all i. Since Ω is closed and g_j $(j = 1, \cdots, m)$ is continuous on Ω, $R(t^i)$ is a closed subset for all i. From $R(t^i) \subset X$ and the boundedness of X, we know that $R(t^i)$ is compact for all i. Since g is continuous on Ω, a global minimizer of $\text{NP}(t^i)$ is attained for all i. So there is a sequence $\{x^i\} \subset X$ with $x^i \in R(t^i)$ such that

$$g(x^i, t^i) = f(t^i) \leq f(t^*) = g(x^*, t^*) \text{ for all } i. \tag{5.8}$$

Since X is bounded and $\{x^i\} \subset X$, $\{x^i\}$ has a convergent subsequence. Without loss of generality, let $x^i \to x^0$. Because $t^i \to t^*$, there exist an $M > 0$ such that $\|t^i - t^*\| < \frac{\delta}{2}$ for all $i \geq M$. So (5.5) and (5.8) imply $\|x^i - x^*\| \geq \frac{\delta}{2}$ for all $i \geq M$.

Therefore $x^0 \neq x^*$. Since Ω is closed and g_j is continuous on Ω $(j = 1, \cdots, m)$, $(x^0, t^*) \in R$; i.e., $x^0 \in R(t^*)$. Again from (5.8) and the continuity of g on Ω, we have $g(x^0, t^*) \leq g(x^*, t^*)$. This contradicts (5.6). Hence t^* is an isolated local minimizer for MNP. The proof of the proposition is complete. □

Now by combining Proposition 5.2.3 and Proposition 5.2.4, we can get the following one to one correspondence on isolated local minimizers for NP and MNP.

Theorem 5.2.2 *Assume that g and g_j $(j = 1, \cdots, m)$ are continuous on Ω, $\{x \in E^n \mid (x, t) \in \Omega,\ g_j(x, t) \leq 0, j = 1, \cdots, m\}$ is bounded, Ω is closed, $R(t)$ is convex and $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. Then (x^*, t^*) is an isolated local minimizer for NP iff t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$.*

In Proposition 5.2.4, the assumption that X is bounded is quite strong. In the following we are going to establish several results which are same type as Proposition 5.2.4 without the boundedness assumption of X. To this end, let $T = \{t \in \mathrm{E}^k \mid R(t) \neq \phi\}$.

Proposition 5.2.5 *Assume that the set-valued map $R(\cdot)$ is lower semi-continuous on T relative to T, g and g_j are continuous on Ω $(j = 1, \cdots, m)$, Ω is closed, $R(t)$ is convex and $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. If (x^*, t^*) is an isolated local minimizer for NP, then t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$.*

Proof. By Lemma 5.2.1, there exists a $\delta > 0$ such that (5.5) and (5.6) hold. Assume that on the contrary, t^* is not an isolated local minimizer for MNP. Then there exists a sequence $\{t^i\} \subset \mathrm{E}^k$ with $t^i \to t^*$ and $t^i \neq t^*$ for all i satisfying

$$f(t^i) \leq f(t^*). \tag{5.9}$$

From the definition of f and (5.9), we know $R(t^i) \neq \phi$ for all i. So, from the lower semi-continuity of $R(\cdot)$ on T relative to T, there exists a sequence $\{y^i\} \subset E^n$ with $y^i \in R(t^i)$ for all i such that $y^i \to x^*$. We claim that NP(t^i) has an optimal solution for sufficiently large i. If this is not the case, then without loss of generality, we may assume that NP(t^i) has no optimal solutions for all i. So for each i, there is a sequence $\{x^{i_t}\} \subset R(t^i)$ with $\|x^{i_t}\| \to +\infty$ such that $g(x^{i_t}, t^i) \to f(t^i)$. Hence there is a sequence $\{x^i\}$ with $x^i \in R(t^i)$ for all i and $\|x^i\| \to +\infty$ such that

$$g(x^i, t^i) \leq \begin{cases} f(t^i) + \frac{1}{i} & \text{if } f(t^i) \text{ is finite,} \\ f(t^*) & \text{otherwise.} \end{cases}$$

Since $\{x^i/\|x^i\|\}$ is bounded, it has a convergent subsequence. Without loss of generality, let $x^i/\|x^i\| \to x^0$. So, for sufficiently large i we have

$$\frac{1}{\|x^i\|}x^i + (1 - \frac{1}{\|x^i\|})y^i \in R(t^i)$$

and

$$g(\frac{1}{\|x^i\|}x^i + (1 - \frac{1}{\|x^i\|})y^i, t^i) \leq \max\{g(x^i, t^i), g(y^i, t^i)\}.$$

By the continuity of g on Ω, we have

$$g(x^0 + x^*, t^*) \leq g(x^*, t^*) = f(t^*). \tag{5.10}$$

Since Ω is closed and g_j is continuous on Ω $(j = 1, \cdots, m)$, $x^0 + x^* \in R(t^*)$. Obviously, $x^0 + x^* \neq x^*$. So, (5.10) contradicts (5.6). Hence NP(t^i) has an optimal solution for sufficiently large i. Without loss of generality, let x^i be an optimal solution for NP(t^i) for all i. From (5.9), we have

$$g(x^i, t^i) \leq f(t^*) \text{ for all } i. \tag{5.11}$$

Similar to the above we can prove that $\{x^i\}$ is bounded. So $\{x^i\}$ has a convergent subsequence. Without loss of generality, let $x^i \to \hat{x}$. Because $t^i \to t^*$, there exist an $M > 0$ such that $\|t^i - t^*\| < \frac{\delta}{2}$ for all $i \geq M$. So (5.5) and (5.11) imply $\|x^i - x^*\| \geq \frac{\delta}{2}$ for all $i \geq M$. Therefore $\hat{x} \neq x^*$. Since Ω is closed and g_j is continuous on Ω $(j = 1, \cdots, m)$, $(\hat{x}, t^*) \in R$; i.e., $\hat{x} \in R(t^*)$. Again from (5.11) and the continuity of g on Ω, we have $g(\hat{x}, t^*) \leq g(x^*, t^*)$. This contradicts (5.6). Hence t^* is an isolated local minimizer for MNP. The proof of the proposition is complete. □

By combining Propositions 5.2.3 and 5.2.5, we have the following one to one correspondence result on isolated local minimizers between NP and MNP.

Theorem 5.2.3 *Let the set-valued map $R(\cdot)$ be lower semi-continuous on T relative to T, g and g_j are continuous on Ω $(j = 1, \cdots, m)$, Ω is closed, $R(t)$ is convex and $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. Then (x^*, t^*) is an isolated local minimizer for NP iff t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$.*

In the following, we are going to consider the assumption of lower semi-continuity of $R(\cdot)$ in Proposition 5.2.5; i.e., we will consider some sufficient conditions which make $R(\cdot)$ a lower semi-continuous set-valued map on T relative to T. In doing so, let

$$\psi_{x^*}(t) = \begin{cases} \inf\{\|x - x^*\| \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $\|y\| = (\sum_{i=1}^{n} y_i^2)^{\frac{1}{2}}$ and x^* is any fixed point in E^n.

Lemma 5.2.2 *Assume that Ω is a closed subset of E^{n+k}, R is a convex subset of E^{n+k}, g_j is a continuous function on Ω $(j = 1, \cdots, m)$ and $T \neq \phi$. Then ψ_{x^*} is a properly closed convex function on E^k with dom$(\psi_{x^*}) = T$.*

Proof. From the definitions of T and ψ_{x^*}, we have $\text{dom}(\psi_{x^*}) = T$. Since each g_j is continuous on Ω and Ω is closed, $R(t)$ is closed subset of E^n for each $t \in E^k$. So, for any $t^1, t^2 \in T$ with $t^1 \neq t^2$, there are two points $x^1 \in R(t^1)$ and $x^2 \in R(t^2)$ such that

$$\psi_{x^*}(t^1) = \|x^* - x^1\| \text{ and } \psi_{x^*}(t^2) = \|x^* - x^2\|.$$

Since $\alpha(x^1, t^1) + (1-\alpha)(x^2, t^2) \in R$ for any α with $0 < \alpha < 1$, $\alpha x^1 + (1-\alpha)x^2 \in R(\alpha t^1 + (1-\alpha)t^2)$. So T is convex and

$$\begin{aligned}
\psi_{x^*}(\alpha t^1 + (1-\alpha)t^2) &\leq \|\alpha x^1 + (1-\alpha)x^2 - x^*\| \\
&= \|\alpha(x^1 - x^*) + (1-\alpha)(x^2 - x^*)\| \\
&\leq \alpha\|x^1 - x^*\| + (1-\alpha)\|x^2 - x^*\| \\
&= \alpha\psi_{x^*}(t^1) + (1-\alpha)\psi_{x^*}(t^2).
\end{aligned}$$

Hence ψ_{x^*} is a properly convex function on E^k. Now let us prove that ψ_{x^*} is closed. For any $r \in R$ and any sequence $\{t^i\}$ with $\psi_{x^*}(t^i) \leq r$ for all i and $t^i \to t^*$, we need to show $\psi_{x^*}(t^*) \leq r$. Since $\psi_{x^*}(t^i) \leq r$ for all i, there is a point $x^i \in R(t^i)$ such that $\psi_{x^*}(t^i) = \|x^i - x^*\| \leq r$ for all i. So $\{x^i\}$ is bounded. Without loss of generality, let $x^i \to x^0$. Since Ω is closed, $(x^i, t^i) \to (x^0, t^0) \in \Omega$. From the continuity of g_j, we have $g_j(x^i, t^i) \to g_j(x^0, t^0) \leq 0$ for $j = 1, \cdots, m$. So $x^0 \in R(t^*)$ and $\|x^* - x^i\| \to \|x^* - x^0\| \leq r$; i.e., $\psi_{x^*}(t^*) \leq r$. Hence ψ_{x^*} is closed. The proof of the lemma is complete. □

Lemma 5.2.3 *Under assumptions of Lemma 5.2.2, $R(\cdot)$ is lower semi-continuous on $ri(T)$ relative to T. Furthermore, if T is locally simplicial at $t^* \in T$, then $R(\cdot)$ is lower semi-continuous at t^* relative to T.*

Proof. For any $t^* \in \mathrm{ri}(T)$, any $x^* \in R(t^*)$ and any sequence $\{t^i\} \subset T$ with $t^i \to t^*$, we need to show that there exists a sequence $\{x^i\}$ with $x^i \in R(t^i)$ for each i such that $x^i \to x^*$. By Lemma 5.2.2, ψ_{x^*} is convex and $\mathrm{dom}(\psi_{x^*}) = T$.

By Theorem 5.2.2 of Rockafellar [27], ψ_{x^*} is continuous on ri(T) relative to T. So, $\psi_{x^*}(t^i) \to \psi_{x^*}(t^*) = 0$; i.e., there is a sequence $\{x^i\}$ with $x^i \in R(t^i)$ for all i satisfying $\psi_{x^*}(t^i) = \|x^i - x^*\| \to 0$. Therefore $x^i \in R(t^i)$ for all i and $x^i \to x^*$. Hence $R(\cdot)$ is lower semi-continuous on ri(T) relative to T. Similarly, if T is locally simplicial at $t^* \in T$, then by Theorem 10.2 of Rockafellar [27], we know that ψ_{x^*} is continuous at t^* relative to T for any $x^* \in R(t^*)$. Hence $R(\cdot)$ is lower semi-continuous at t^* relative to T. The proof of the lemma is complete. □

Proposition 5.2.6 *Assume that Ω is a closed subset of E^{n+k}, g_j and g are continuous on Ω $(j = 1, \cdots, m)$, R is a convex subset of E^{n+k}, $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. If (x^*, t^*) is an isolated local minimizer for NP, then arg min$\{NP(t^*)\} = \{x^*\}$ and t^* is an isolated local minimizer for MNP if $t^* \in ri(T)$ or T is locally simplicial at t^*.*

Proof. By Lemma 5.2.1, we have arg min$\{\text{NP}(t^*)\} = \{x^*\}$. From Lemma 5.2.3, we know that $t^* \in$ ri(T) (or T is locally simplicial at t^*) implies $R(\cdot)$ is lower semi-continuous at t^* relative to T. So, all conditions of Proposition 5.2.5 are satisfied. Hence t^* is an isolated local minimizer for MNP. This completes the proof of the proposition. □

Remember that k is the dimension of parameter t. If $k = 1$, we have the following corollary from Proposition 5.2.6.

Corollary 5.2.1 *Assume that $k = 1$, Ω is a closed subset of E^{n+k}, g_j and g are continuous on Ω $(j = 1, \cdots, m)$, R is a convex subset of E^{n+k}, $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. If (x^*, t^*) is an isolated local minimizer for NP, then t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$.*

Proof. By Lemma 5.2.1, we have arg min$\{\text{NP}(t^*)\} = \{x^*\}$. Since $k = 1$, T is an interval in E^1. So, T is locally simplicial at t for any $t \in T$. By Proposition 5.2.6, t^*

is an isolated local minimizer for MNP. The proof of the corollary is thus complete. □

Similar to Corollary 5.2.1, the following corollary also follows from Proposition 5.2.6.

Corollary 5.2.2 *Assume that Ω is a polyhedron in E^{n+k}, $g_j(x,t) = a_j'x + b_j't + c_j$, $a_j \in E^n$, $b_j \in E^k$, $c_j \in E^1$ $(j = 1, \cdots, m)$, g is continuos on Ω and $g(\cdot, t)$ is quasi-convex on $R(t)$ for each $t \in E^k$. If (x^*, t^*) is an isolated local minimizer for NP, then t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$.*

In the following, we will consider the convex case; i.e., $g(\cdot, t)$ is convex on E^n for each $t \in \mathrm{E}^k$. We need some assumptions as follows.

Assumption A Let Ω be a closed subset of E^{n+k}, g and g_j $(j = 1, \cdots, m)$ are continuous on Ω, $R(t)$ is a convex subset of E^n for each $t \in \mathrm{E}^k$, $g(\cdot, t)$ is convex on E^n and for each $t^* \in T$ there are $\epsilon = \epsilon(t^*) > 0$ and $M = M(t^*) > 0$ such that $B_M(0) \bigcap R(t) \neq \phi$ for each $t \in B_\epsilon(t^*) \bigcap T$, where $B_M(0) = \{x \in \mathrm{E}^n \mid \| x \| \leq M\}$.

Remark 5.2.1 The assumption that $B_M(0) \bigcap R(t) \neq \phi$ for each $t \in B_\epsilon(t^*) \bigcap T$ is much weaker than the lower semi-continuity of $R(\cdot)$ on T relative to T. Also, this condition is much easier to check than lower semi-continuity.

Proposition 5.2.7 *Under Assumption A, if (x^*, t^*) is an isolated local minimizer for NP, then t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$..*

Proof. By Lemma 5.2.1, there exists a $\delta > 0$ such that (5.5) and (5.6) hold. Assume that on the contrary, t^* is not an isolated local minimizer for MNP. Then there exists a sequence $\{t^i\} \subseteq B_\epsilon(t^*) \bigcap T$ with $t^i \to t^*$ and $t^i \neq t^*$ for i satisfying

$$f(t^i) \leq f(t^*). \tag{5.12}$$

Without loss of generality, we may take $0 < \epsilon \leq \frac{1}{2}\delta$. From (5.5), we have

$$g(x^*, t^*) < g(x, t^i) \text{ for each } x \in R(t^i) \bigcap B_\epsilon(x^*).$$

Since $g(\cdot, t^i)$ is continuous on $R(t^i) \bigcap B_\epsilon(x^*)$ and $R(t^i) \bigcap B_\epsilon(x^*)$ is compact, there is an $\epsilon^i > 0$ such that

$$g(x^*, t^*) + \epsilon^i < g(x, t^i) \text{ for each } x \in R(t^i) \bigcap B_\epsilon(x^*). \tag{5.13}$$

Take a sequence $\{\gamma^i\} \subseteq \mathrm{E}^1$ with $0 < \gamma^i < \epsilon^i$ for all i and $\gamma^i \to 0$. Then from (5.12) there is a sequence $\{x^i\} \subseteq \mathrm{E}^n$ with $x^i \in R(t^i)$ for all i satisfying

$$g(x^i, t^i) < f(t^*) + \gamma^i. \tag{5.14}$$

From (5.13), (5.14) and $g(x^*, t^*) = f(t^*)$, we have $\| x^i - x^* \| > \epsilon$ for each i. From (5.6), we know $\{x \in R(t^*) \mid g(x, t^*) \leq g(x^*, t^*)\} = \{x^*\}$. So, by Corollary 8.7.1 of Rockafellar [27], $\{x \in R(t^*) \mid g(x, t^*) \leq \lambda\}$ is bounded for any $\lambda \in \mathrm{E}^1$. Since $B_M(0) \bigcap R(t) \neq \phi$ for each $t \in B_\epsilon(t^*) \bigcap T$, there exists a sequence $\{y^i\} \subseteq R(t^i)$ such that $\| y^i \| \leq M$ for all i. Without loss of generality, let $y^i \to y^0$. Now we claim that $\{x^i\}$ is bounded. If this is not true, we may assume that $\| x^i \| \to +\infty$ and $x^i / \| x^i \| \to x^0$. So for any $\alpha > 0$, we have

$$\frac{\alpha}{\| x^i \|} x^i + \left(1 - \frac{\alpha}{\| x^i \|}\right) y^i \in R(t^i)$$

and

$$\begin{aligned} & g\left(\frac{\alpha}{\| x^i \|} x^i + \left(1 - \frac{\alpha}{\| x^i \|}\right) y^i, t^i\right) \\ \leq\ & \frac{\alpha}{\| x^i \|} g(x^i, t^i) + \left(1 - \frac{\alpha}{\| x^i \|}\right) g(y^i, t^i) \\ <\ & \frac{\alpha}{\| x^i \|} (f(t^*) + \gamma^i) + \left(1 - \frac{\alpha}{\| x^i \|}\right) g(y^i, t^i) \end{aligned}$$

for sufficiently large i. Therefore the continuity of g on Ω implies

$$g(\alpha x^0 + y^0, t^*) \leq g(y^0, t^*). \tag{5.15}$$

The continuity of g_j ($j = 1, \cdots, m$) on Ω and the closedness of Ω imply $\alpha x^0 + y^0 \in R(t^*)$. Since α is any positive number, (5.15) implies $\{x \in R(t^*) \mid g(x, t^*) \leq g(y^0, t^*)\}$ is unbounded, a contradiction. Hence $\{x^i\}$ is bounded. Without loss of generality, let $x^i \to \bar{x}$. From $\| x^i - x^* \| > \epsilon$ for all i and (5.14), we have

$$g(\bar{x}, t^*) \leq g(x^*, t^*) \text{ and } \bar{x} \neq x^*.$$

This contradicts (5.6). Hence t^* is an isolated local minimizer for MNP and the proof of the proposition is complete. □

The following example shows that the condition $B_M(0) \bigcap R(t) \neq \phi$ for each $t \in B_\epsilon(t^*) \bigcap T$ is necessary in Proposition 5.2.7.

Example 5.2.1 Let $\Omega = \{(0,0)'\} \bigcup \{(x,t) \in E^2 \mid x > 0, t > 0, tx = 1\}$, $m = 0$, $k = n = 1$ and $g(x,t) \equiv 1$ for each $(x,t) \in E^2$.

Obviously, Ω is closed, g is continuous on Ω, g is convex on $R(t)$ for each $t \in E^1$, $R(0) = \{0\}$ and $R(t) = \{1/t\}$ for each $t > 0$. From Figure 5.1 (a), we can see that $(0,0)'$ is an isolated feasible solution. So, $(0,0)'$ is an isolated local minimizer for NP. But $f(t) \equiv 1$ for $t \geq 0$ implies that $t^* = 0$ is not an isolated local minimizer for MNP. It is easy to check that $B_M(0) \bigcap R(t) \neq \phi$ for each $t \in B_\epsilon(t^*) \bigcap T$ does not hold. So this condition is necessary in Proposition 5.2.7.

Now by combining Proposition 5.2.3 and 5.2.7, we have the following one to one correspondence theorem for the convex case.

Theorem 5.2.4 *Let Assumption A be satisfied. Then (x^*, t^*) is an isolated local minimizer for NP iff t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\}$ = $\{x^*\}$.*

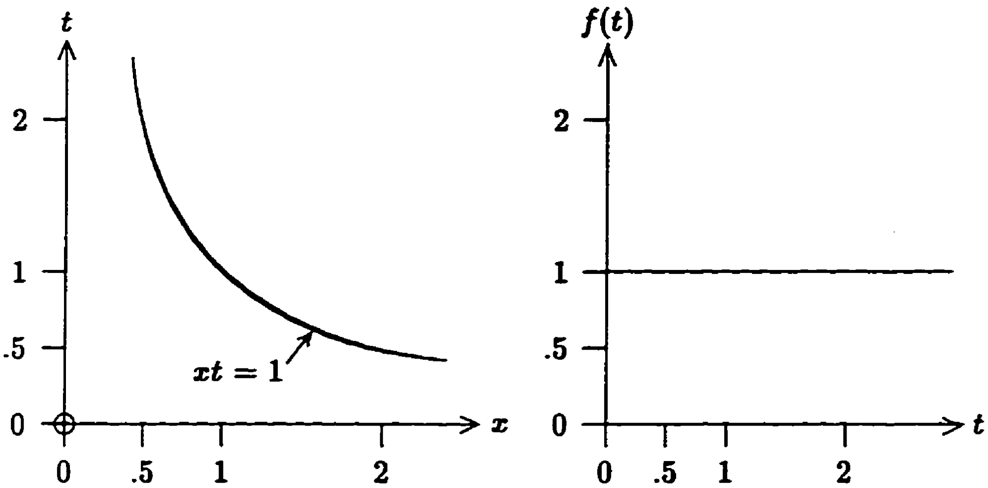

Figure 5.1: (a) Example 5.2.1 (b) $f(t)$ for Example 5.2.1

In the following, we will consider a special class of problems, namely those for which $g(\cdot,t)$ is a quadratic convex function on E^n for each fixed t. In doing so, we need the following assumptions.

Assumption B Let $g(x,t) = c(t)'x + \frac{1}{2}x'C(t)x + \psi(t)$, $g_j(x,t) = a_j(t)'x + b_j(t)$, $j = 1,\cdots,m$, $\Omega = E^{n+k}$, where $C(\cdot) : E^k \to E^{n\times n}$, $c(\cdot) : E^k \to E^n$, $\psi(\cdot) : E^k \to E^1$, $a_j(\cdot) : E^k \to E^n$ and $b_j(\cdot) : E^k \to E^1$ ($j = 1\cdots,m$) are all continuous on E^k and, $C(t)$ is is a symmetric positive semidefinite matrix for each $t \in T$.

Proposition 5.2.8 *Under the Assumption B, if (x^*,t^*) is an isolated local minimizer for NP, then t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\} = \{x^*\}$.*

Proof. Since (x^*,t^*) is an isolated minimizer for NP, there exists a $\delta > 0$ such that

$$c(t^*)'x^* + \frac{1}{2}(x^*)'C(t^*)x^* + \psi(t^*) < c(t)'x + \frac{1}{2}x'C(t)x + \psi(t) \tag{5.16}$$

for each $(x,t) \in B_\delta(x^*,t^*) \bigcap R \setminus \{(x^*,t^*)\}$. So,

$$c(t^*)'x^* + \frac{1}{2}(x^*)'C(t^*)x^* + \psi(t^*) < c(t^*)'x + \frac{1}{2}x'C(t^*)x + \psi(t^*) \qquad (5.17)$$

for each $x \in B_\delta(x^*) \bigcap R(t^*) \setminus \{x^*\}$. Since $c(t^*)'x + \frac{1}{2}x'C(t^*)x + \psi(t^*)$ is convex on $R(t^*)$, (5.17) also holds for any $x \in R(t^*) \setminus \{x^*\}$. So, arg min$\{\text{NP}(t^*)\} = \{x^*\}$. Let

$$f_1(t) = \begin{cases} \inf \{c'(t)x + \frac{1}{2}x'C(t)x \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise,} \end{cases}$$

and consider

QP(t) $\qquad$ $\min \{c'(t)x + \frac{1}{2}x'C(t)x \mid x \in R(t)\}$.

Since $\{x^*\}$ is the solution set of QP(t^*), there is no nonzero solution for the following system

$$a_j'(t^*)s \leq 0 \ (j = 1, \cdots, m), \ C(t^*)s = 0, \ c'(t^*)s \leq 0. \qquad (5.18)$$

By Theorem 1.2.1, f_1 is lower semi-continuous at t^*. Therefore $f = f_1 + \psi$ is lower semi-continuous at t^*. So, for any $\gamma > 0$, there exists an $\epsilon > 0$ with $\epsilon < \frac{\delta}{2}$ such that $f(t) \geq f(t^*) - \gamma$ for each $t \in B_\epsilon(t^*)$. Now assume that on the contrary, t^* is not an isolated local minimizer for MNP, then there exists a sequence $\{t^i\} \subset B_\epsilon(t^*)$ with $t^i \neq t^*$ for all i and $t^i \to t^*$ such that

$$f(t^*) - \gamma \leq f(t^i) \leq f(t^*). \qquad (5.19)$$

Therefore $R(t^i) \neq \phi$ for each i. Thus, QP(t^i) has an optimal solution x^i for all i; i.e.,

$$f(t^i) = c(t^i)'x^i + \frac{1}{2}(x^i)'C(t^i)x^i + \psi(t^i) \ \text{ for all } i. \qquad (5.20)$$

We claim that $\{x^i\}$ is bounded. If this is not true, we may assume

$$\lim_{i\to\infty} \frac{x^i}{\| x^i \|} = s \text{ and } \lim_{i\to\infty} \| x^i \| = +\infty.$$

So, the continuity of ψ and $t^i \to t^*$ imply

$$\lim_{i\to+\infty} \frac{\psi(t^i)}{\| x_i \|} = \lim_{i\to+\infty} \frac{\psi(t^i)}{\| x_i \|^2} = 0.$$

From (5.19) and (5.20), we have

$$\frac{1}{2}s'C(t^*)s = \lim_{i\to+\infty} \frac{f(t^i)}{\| x^i \|^2} = 0$$

and

$$c(t^*)'s + \lim_{i\to\infty} \frac{(x^i)'C(t^i)x^i}{2 \| x^i \|} + \lim_{i\to\infty} \frac{\psi(t^i)}{\| x^i \|} = \lim_{i\to\infty} \frac{f(t^i)}{\| x^i \|} = 0,$$

i.e.,

$$C(t^*)s = 0 \text{ and } c(t^*)'s = -\lim_{i\to\infty} \frac{(x^i)'C(t^i)x^i}{2 \| x^i \|} \leq 0.$$

From $a_j(t^i)'x^i + b_j(t^i) \leq 0$ for $j = 1, \cdots, m$, we have $a_j(t^*)'s \leq 0$ for $j = 1, \cdots, m$. So, (5.18) has a nonzero solution, a contradiction. Hence $\{x^i\}$ is bounded and has a convergent subsequence. Without loss of generality, let $x^i \to x^0$. Since $t^i \in B_\epsilon(t^i)$ and $\epsilon < \frac{\delta}{2}$, (5.16), (5.19) and (5.20) imply $x^i \notin B_{\frac{\delta}{2}}(x^*)$. So, $x^0 \neq x^*$. It is easy to show $x^0 \in R(t^*)$. Hence (5.19) and (5.20) imply

$$c(t^*)'x^0 + \frac{1}{2}(x^0)'C(t^*)x^0 + \psi(t^*) \leq c(t^*)'x^* + \frac{1}{2}(x^*)'C(t^*)x^* + \psi(t^*).$$

This contradicts (5.17). Hence t^* is an isolated local minimizer for MNP. The proof of the proposition is complete. □

Example 5.2.2. Consider the following nonlinear programming problem

$$\begin{aligned} &\min\ -x + x^2(2 - sin(t)), \\ &\text{subject to } 1/4 \leq x \leq 1/(1 + cos^2(t)), \\ &\qquad\qquad 0 \leq t \leq 2\pi. \end{aligned}$$

Obviously, the problem is non-convex, but when t is fixed, it is a convex quadratic programming problem. So, we can take t as a parameter. Therefore, we have

$$f(t) = \begin{cases} \frac{-1}{4(2-sin(t))} & \text{for } 0 \le t \le \pi \text{ with } x = \frac{1}{2(2-sin(t))}, \\ -\frac{1}{8} - \frac{sin(t)}{16} & \text{for } \pi \le t \le 2\pi \text{ with } x = \frac{1}{4}, \\ +\infty & \text{otherwise.} \end{cases}$$

Example 5.2.2 is illustrated in Figures 5.2(a) and (b). Figure 5.2(a) shows the feasible region of the example. Figure 5.2(b) shows $f(t)$, a piece-wise nonlinear function which by inspection, has two isolated local minimizers $\pi/2$ and 2π. So, by Proposition 5.2.2, the problem also has two local minimizers $(\pi/2, 1/2)$ and $(2\pi, 1/4)$. By Proposition 5.2.3, we can see that $(\pi/2, 1/2)$ and $(2\pi, 1/4)$ are two isolated local minimizers. By Proposition 5.2.8, the problem has exactly two isolated local minimizers.

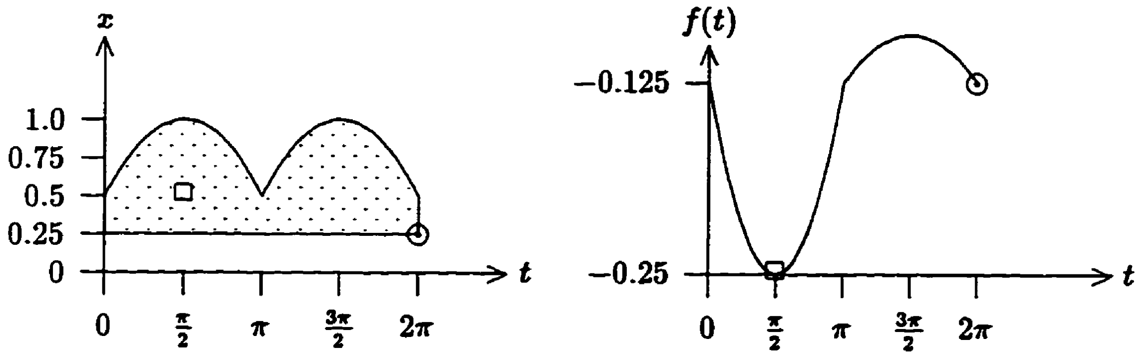

Figure 5.2: (a) Example 5.2.2 (b) $f(t)$ for Example 5.2.2

Now by combining Proposition 5.2.3 and 5.2.8, we have the following one to one correspondence result on isolated local minimizers between NP and MNP.

Theorem 5.2.5 *Let Assumption B be satisfied. Then (x^*, t^*) is an isolated local minimizer for NP iff t^* is an isolated local minimizer for MNP and arg min$\{NP(t^*)\}$ = $\{x^*\}$.*

From Theorem 5.2.2 to 5.2.5, we can see that for a large class of functions g, isolated local minimizers for NP and MNP always has one to one correspondence. In the following section, we are going to consider some applications of the results established in this section.

5.3 Applications

In this section, we will discuss some applications of the results established in Section 2. First of all, we will discuss how to formulate MNP and f. Then we will give several concrete examples. Usually, a nonlinear programming problem is given as follows

$$\min \{\bar{g}(x) \mid x \in X\}, \tag{5.21}$$

where $\bar{g}(\cdot) : \mathrm{E}^n \to \mathrm{E}^1$ and $X \subset \mathrm{E}^n$. It is possible that $\bar{g}$ has no any special properties (for example, convex or quasi-convex) on $R(t)$ for any $x_I = t$, where $I \subset \{1, \cdots, n\}$ and x_I represents a vector induced by all components corresponding to the indices of I. However $\bar{g}$ may have some useful properties on $R(t)$ if $R(t)$ is induced by some k-dimensional vector function h, that is, $R(t) = \{x \in X \mid h(x) = t\}$. So, MNP may be formulated as

$$\min \{f(t) \mid t \in \mathrm{E}^k\}, \tag{5.22}$$

where

$$f(t) = \begin{cases} \inf \{\bar{g}(x) \mid x \in R(t)\} & \text{if } R(t) \neq \phi \\ +\infty & \text{otherwise} \end{cases}$$

In Section 5.2, parameter t is the part of components of variable (x, t) in NP. But in (5.21) and (5.22), parameter t may not be any part of components of x.

However if we define $\bar{g}(x)$ as $g(x,t)$; i.e., $g(x,t) \equiv \bar{g}(x)$, all results established in Section 5.2 hold for (5.21) and (5.22). Let us state one of them for (5.21) and (5.22) to illustrate this, for example, Proposition 5.2.4.

Proposition 5.3.1 *Assume that $\bar{g}$ and h are continuous on X, X is a compact subset of E^n, $R(t)$ is convex and $\bar{g}$ is quasi-convex on $R(t)$ for each $t \in E^k$. If x^* is an isolated local minimizer for (5.21), then $t^* = h(x^*)$ is an isolated local minimizer for (5.22) and $\{x \in R(t^*) \mid \bar{g}(x) = f(t^*)\} = \{x^*\}$.*

Proof. Define $g(x,t) = \bar{g}(x)$, $\Omega = X \times \mathrm{E}^k$, $g_j(x,t) = h_j(x) - t_j$ and $g_{j+k}(x,t) = t_j - h_j(x)$ for $j = 1, \cdots, k$ and $m = 2k$. Since x^* is an isolated local minimizer for (5.21), (x^*, t^*) is an isolated local minimizer for NP. Also it is straightforward to check that all other conditions of Proposition 5.2.4 hold. So, t^* is an isolated local minimizer for (5.22) and $\{x \in R(t^*) \mid \bar{g}(x) = f(t^*)\} = \{x^*\}$. The proof of the proposition is thus complete. □

In the following, we are going to give several examples. For all these examples, the function f can be computed efficiently by parametric linear programming or parametric quadratic programming technique. So, by the results of Section 5.2 and the parametric local optimization procedure, we know that a global minimizer (if it exists), all isolated local minimizers and some local minimizers of these examples can be computed efficiently.

Example 5.3.1 Consider

$$\min \{q'x + \gamma + \frac{c'x + \alpha}{d'x + \beta} \mid Ax \leq b\}, \tag{5.23}$$

where $c, d, q \in \mathrm{E}^n$, $\alpha, \beta, \gamma \in \mathrm{E}^1$, $b \in \mathrm{E}^m$, $A \in \mathrm{E}^{m \times n}$ and $d'x + \beta > 0$ for any $x \in \{x \in \mathrm{E}^n \mid Ax \leq b\}$.

For this problem, a suitable choice for h is $d'x + \beta$. So,

$$f(t) = \begin{cases} \frac{\alpha + t\gamma}{t} + \frac{1}{t} \inf \{c'x + tq'x \mid x \in R(t)\} & \text{if } R(t) \neq \phi \\ +\infty & \text{otherwise} \end{cases}$$

where $R(t) = \{x \in \mathrm{E}^n \mid Ax \leq b,\ d'x = t - \beta\}$. Therefore, $f(t)$ can be formulated by solving the following parametric programming problem

$$\min \{c'x + tq'x \mid Ax \leq b,\ d'x = t - \beta\}.$$

This is a parametric linear programming problem. So, it can be solved efficiently for all t. The reader may refer to Best and Ritter [3] or [12] for more details. Hence a global minimizer, all isolated local minimizers and some local minimizers of (5.23) can be computed efficiently by the parametric local optimization procedure.

Now let us illustrate Example 5.3.1 by taking $\gamma = 2$, $\alpha = -2$, $\beta = 1$, $c = (1, -1)'$, $q = (-1, 0)'$, $d = (1, 1)'$,

$$A = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \text{ and } b = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix};$$

i.e.,

$$\min \{2 - x_1 + \frac{x_1 - x_2 - 2}{x_1 + x_2 + 1} \mid 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\} \tag{5.24}$$

Then $\min \{c'x + tq'x \mid x \in R(t)\}$ have the following solution

$$x(t) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{for } 1 \leq t \leq 2,$$

$$x(t) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{for } 2 \leq t \leq 3$$

and

$$R(t) = \phi \text{ for } t < 1 \text{ or } t > 3.$$

This implies

$$f(t) = \begin{cases} \frac{t-1}{t}, & \text{if } 1 \le t \le 2, \\ \frac{1}{t}, & \text{if } 2 \le t \le 3, \\ +\infty, & \text{otherwise} \end{cases}$$

It is straightforward to check that $t = 1$ and $t = 3$ are two local minima for f and $t = 3$ is a global minimum. So,

$$x(1) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } x(3) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

are local and global minima for (5.24), respectively.

Example 5.3.2 Consider

$$\min \left\{\frac{q'x + \gamma}{(d'x + \beta)^{l-1}} + \frac{\alpha + c'x + x'Cx}{(d'x + \beta)^{l}} + \psi(d'x + \beta) \mid Ax \le b\right\}, \tag{5.25}$$

where $C \in E^{n \times n}$ is a symmetric positive semidefinite matrix, $A \in E^{m \times n}$, $c, d, q \in E^n$, $b \in E^m$, $\alpha, \beta, \gamma \in E^1$, $\psi(\cdot) : E^1 \to E^1$ is a continuous real valued function, $d'x + \beta > 0$ for each $x \in \{x \in E^n \mid Ax \le b\}$ and l is a positive integer.

As in Example 5.3.1, a suitable choice for function h is $d'x + \beta$, that is, $R(t) = \{x \in E^n \mid Ax \le b,\ d'x = t - \beta\}$. So,

$$f(t) = \begin{cases} \psi(t) + \frac{\alpha + t\gamma}{t^l} + \frac{1}{t^l} \inf \{c'x + tq'x + x'Cx \mid x \in R(t)\} & \text{if } R(t) \ne \phi \\ +\infty & \text{otherwise} \end{cases}$$

So, f may be formulated by solving the following parametric programming problem

$$\min \{c'x + tq'x + x'Cx \mid Ax \le b,\ d'x = t - \beta\}.$$

This is a parametric quadratic programming problem and can be solved efficiently. The reader may refer to Best [3] for more details. For some class of functions ψ, $\min \{f(t) \mid t \in \mathrm{E}^1\}$ can be solved efficiently; i.e., (5.25) can be solved efficiently for a global minimizer, all isolated local minimizers and some local minimizers by the parametric local optimization procedure.

Remark 5.3.1 Obviously, (5.25) will reduce to (5.23) if $l = 1$, $C \equiv 0$ and $\psi \equiv 0$. So (5.25) is a generalized formulation of (5.23).

When $l = 1$ in Example 5.3.2, the model was used for a portfolio optimization problem in finance by Speakman [28].

Example 5.3.3 Consider the indefinite quadratic programming problem

$$\min \{c'x + \frac{1}{2}x'Cx + x'DQ'x \mid Ax \leq b\}, \tag{5.26}$$

where $C \in \mathrm{E}^{n\times n}$ is a symmetric positive semidefinite matrix, $D, Q \in \mathrm{E}^{n\times k}$, $A \in \mathrm{E}^{m\times n}$, $c \in \mathrm{E}^n$ and $b \in \mathrm{E}^m$.

This problem was studied in Chapter 2 by setting $h(x) = D'x$, that is, $R(t) = \{x \in \mathrm{E}^n \mid Ax \leq b,\ D'x = t\}$. The reader may refer to Chapter 2 for details.

Finally, let us analyze a class of cubic minimization problems with linear constraints.

Example 5.3.4 Consider the following cubic minimization problem

$$\min \{c' \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}x'(C + y\hat{C})x \mid A \begin{pmatrix} x \\ y \end{pmatrix} \leq b\}, \tag{5.27}$$

where $c \in \mathrm{E}^{n+1}$, C and $\hat{C}$ are (n, n) symmetric matrices, $A \in \mathrm{E}^{m\times(n+1)}$, $b \in \mathrm{E}^m$, $x \in \mathrm{E}^n$ and $y \in \mathrm{E}^1$ are variables. We assume that $C + y\hat{C}$ is positive semidefinite for each y such that $\{x \in \mathrm{E}^n \mid A(x', y')' \leq b\} \neq \phi$.

Obviously, we should take y as the parameter t, that is, $R(t) = \{x \in \mathrm{E}^n \mid A(x', t)' \leq b\}$. So, we have

$$f(t) = \begin{cases} \inf \{c'(x', t)' + \frac{1}{2}x'(C + t\hat{C})x \mid x \in R(t)\}, & \text{if } R(t) \neq \phi, \\ +\infty, & \text{otherwise.} \end{cases}$$

So, f may be formulated by solving the following parametric programming problem

$$\min \{c' \begin{pmatrix} x \\ t \end{pmatrix} + \frac{1}{2}x'(C + t\hat{C})x \mid x \in R(t)\}.$$

This is a parametric quadratic programming problem with a parameterized Hessian. Some special cases of the problem can be solved efficiently, for example, $\operatorname{rank}(\hat{C}) = 1$ or 2. The reader may refer to Best and Caron [4] for more details. Hence some special cases of (5.27) can be solved efficiently for a global minimizer, all isolated local minimizers and some local minimizers by the parametric local optimization procedure.

5.4 Conclusion

We have investigated the relationships between the original nonlinear programming problem NP and its main problem MNP. In particular, we have established a one to one correspondence between isolated local minimizers for NP and those for MNP for a large class of nonlinear programming problems. These provide some ideas and a theoretical background concerning the computation of a global minimizer, all isolated local minimizers and some local minimizers for this class of nonlinear programming problems.

We have shown that for linear fractional programming problems, some quadratic fractional programming problems, some indefinite quadratic programming problems and some cubic programming problems, a global minimizer (if it exists), all isolated local minimizers and some local minimizers can be computed efficiently.

Chapter 6

Conclusion

6.1 Introduction

In this chapter we summarize the contributions of this thesis and also outline future research projects.

6.2 Contribution of Thesis

The main contributions of the thesis are listed below.

(1) We have developed relationships between a given non-convex quadratic programming problem QP and a derived unconstrained (but non-differentiable) quadratic problem MQP. We established that any local minimum of MQP gives a corresponding local minimum of QP. Also we established that the isolated local minima (including the global minimum) of both QP and MQP are in one to one correspondence. For the case that the Hessian of QP has exactly one negative eigenvalue, we have developed an algorithm to compute all isolated local minima and some

non-isolated local minima of QP by parametric quadratic programming techniques. The algorithm will compute the global minimizer of QP, provided it exists, and will provide the information that QP is unbounded from below when that is the case.

(2) Based on the results stated in (1) and parametric linear programming technique we established a decomposition procedure which when applied to indefinite quadratic programming problem will locate all isolated local minima and some non-isolated local minima. The decomposition procedure will also locate the global minimum of the indefinite quadratic programming problem if it exists, and will provide the information that the problem is unbounded from below when that is the case.

(3) We established a polynomial algorithm for a class of coposotivity problems in which the matrix has exactly one negative eigenvalue. A slight modification of the algorithm provides an efficient method for a class of copositivity problems in which the matrix has exactly two negative eigenvalues.

(4) We generalized the results established for non-convex quadratic programming problems to general non-convex minimization problems, that is, the objective function is not quadratic and constraints are not linear.

6.3 Further Research Directions

In this section we outline some research topics and open questions related to this thesis.

(1) In Chapter 2 we have developed an efficient numerical procedure for the non-convex quadratic programming problem in which the Hessian has exactly one negative eigenvalue. The procedure is designed to locate all isolated local minima,

some non-isolated local minima and the global minimum. For the same purpose, we have developed a decomposition procedure for indefinite quadratic programming problems. However the decomposition procedure may not be efficient from the computational point of view. Even the indefinite QP with a Hessian having exactly two negative eigenvalues is quite different from the non-convex QP with a Hessian having exactly one negative eigenvalue. Does there exist an efficient algorithm which when applied to the indefinite QP with a Hessian having exactly two negative eigenvalues can locate all isolated local minima, some non-isolated local minima and the global minimum? This appears to be an interesting and challenging open question.

(2) In Section 3 of Chapter 3, we have shown that the number of subproblems can be reduced if $(M')^{-1}CM^{-1}$ contains an indefinite principal submatrix. So we need an efficient algorithm to find the smallest indefinite principal submatrix of $(M')^{-1}CM^{-1}$. Does there exist such an efficient algorithm?

(3) In Chapter 3 we have developed a decomposition procedure based on the parametric LP technique and a decomposition for indefinite symmetric matrices for indefinite quadratic programming problems. Do there exist other decomposition procedures for this type of problem?

(4) In Chapter 5 we have extended the one to one correspondence result for non-convex QP to a large classes of non-convex programming problems. In Section 3 of Chapter 5 we have investigated several applications of the one to one correspondence result to several classes of non-convex programming problems. However in these applications the parametric programming problem NP(t) is still a parametric quadratic programming problem; i.e., we still use the parametric QP technique to compute $f(t)$. Therefore, are there other techniques other than parametric QP which we can use to compute $f(t)$ so that we can solve a large classes of non-convex

programming problems for all isolated local minima, some non-isolated local minima and the global minimum?

(5) In Chapter 4 we have shown that the copositivity problem with a Hessian having exactly one negative eigenvalue can be solved in polynomial time if $\mathrm{rank}(C) = n$. Can the problem be solved in polynomial time if $\mathrm{rank}(C) < n$?

Bibliography

[1] Bank, B., Guddat, T., Klatte, D., Kummer, B., and Tammer, K., Nonlinear Parametric Optimization, Akademie–Verlag, Berlin, Germany, 1982.

[2] Benders, J.F., Partitioning procedures for solving mixed–variables programming problems, *Numerische Mathematick*, **4** (1962) 238–252.

[3] Best, M. J., An algorithm for the solution of the parametric quadratic programming algorithm, in Applied Mathematics and Parallel Computing – Festschrift for Klaus Ritter, H.Fischer, B. Riedmüller and S. Schäffler (editors), Heidelburg: Physica–Verlag, (1996) 57–76.

[4] Best, M. J., and Caron, R. J., A parameterized Hessian quadratic programming problem, *Annals of Operations Research*, **5** (1985) 373–394.

[5] Best. M. J., and Ding, B., On the continuity of the minimum in parametric quadratic programs, *Journal of Optimization Theory and Applications*, **86** (1995) 245–250.

[6] Best, M. J., and Ding, B., Global and local quadratic minimization, *Journal of Global Optimization*, (1996) to appear.

[7] Best, M. J., and Ding, B., A decomposition procedure for global and local quadratic minimization (submitted).

[8] Best, M. J., and Ritter, K., Linear Programming: Active Set Analysis and Computer Programs, Prentice–Hall Inc., Englewood Cliffs New Jersey, 1985.

[9] Best, M. J., and Ritter, K., Quadratic Programming: Active Set Analysis and Computer Programs, Prentice–Hall, Englewood Cliffs, New Jersey (to appear).

[10] Best, M. J., and Ritter, K., A quadratic programming algorithm, *Zeitschrift für Operations Research*, **32** (1988) 271–297.

[11] Bomze, I. M., and Danninger, G., A global optimization algorithm for concave quadratic programming problems, *SIAM J. Optimization*, **3** (1993) 826–842.

[12] Dinkelbach, W., On nonlinear fractional programming, *Management Sciences*, **13** (1967) 492–498.

[13] Danninger, G., Role of copositivity in optimality criteria for nonconvex optimization problems, *Journal of Optimization Theory and Applications*, **75** (1992) 535–558.

[14] Floudas, C. A., and Pardalos, P. M., A Collection of Test Problems for Constrained Global Optimization Algorithms, volume 455 of Lecture Notes in Computer Science. Springer–Verlag, Berlin, Germany, 1990.

[15] Gill, P. E., and Murray, W., Numerically stable methods for quadratic programming, *Mathematical Programming*, **14** (1978) 349–372.

[16] Geoffrion, A. M., Generalized Benders decomposition, *J. of Optimization Theory and Applications*, **10** (1972) 237–260.

[17] Konno, H., Kuno, T., and Yajima, T., Parametric simplex algorithms for solving a special class of nonconvex minimization problems, *Journal of Global Optimization*, **1** (1991) 65–82.

[18] Kough, F., The indefinite quadratic programming problem, *Operations Research*, **27** (1979) 516–533.

[19] Mangsarian, O. L., Nonlinear Programming, McGraw-Hill, New York, 1969.

[20] Monteiro, R. D. C., and Adler, I., Interior path following primal–dual algorithms, Part II: Convex quadratic programming, *Mathematical Programming*, **44** (1989) 43–66.

[21] Murty, K. G., Linear Complementarity, Linear and Nonlinear Programming, Berlin: Heldermann, 1988.

[22] Murty, K. G., and Kabadi, S. N., Some NP–complete problems in quadratic and nonlinear programming, *Mathematical Programming*, **39** (1987) 117–129.

[23] Pardalos, P. M., Polynomial time algorithms for some classes of constrained nonconvex quadratic problems, *Optimization*, **21** (1990) 843–853.

[24] Pardalos, P. M. and Schnitger, G., Checking local optimality in constrained quadratic programming is NP–hard, *Operations Research Letters*, **7** (1988) 33–35.

[25] Pardalos, P. M., Glick, J. H., and Rosen, J. B., Global minimization of indefinite quadratic problems, *Computing*, **39** (1978) 281–291.

[26] Pardalos, P. M., and Vavasis, S. A., Quadratic programming with one negative eigenvalue is NP–hard, *Journal of Global Optimization*, **1** (1991) 15–22.

[27] Rockafellar, R. T., Convex Analysis, Princeton Press, Princeton, 1970.

[28] Speakman, G. J, A Portfolio Tracking Problem, A master thesis, Department of Combinatorics and Optimization, University of Waterloo, 1995.

[29] Valiaho, H., Quadratic–programming criteria for copositive matrices, *Linear Algebra and its Applications*, **119** (1989) 163–182.

[30] Van de Panne, C., and Whinstoon, A., The symmetric formulation of the simplex method for quadratic programming, *Econometrica*, **37** (1969) 507-–527.

[31] Vavasis, S. A., Nonlinear Optimization: Complexity Issues, Oxford, Oxford University Press, 1991.

[32] Vavasis, S. A., On approximation algorithms for concave quadratic programming, in: C. A. Floudas and P. M. Pardalos, eds., Recent Advances in Global Optimization (Princeton University Press, Princeton, NJ, (1992a) 3–18.

[33] Vavasis, S. A., Approximation algorithms for indefinite quadratic programming, *Mathematical Programming*, **57** (1992) 279–311.

[34] Ye, Y., and Tse, E., An extension of Karmarkar's projective algorithm for convex quadratic programming, *Mathematical Programming*, **44** (1989) 157-–179.