QUADRATIC PROGRAMMING AND AFFINE VARIATIONAL INEQUALITIES

A Qualitative Study

Nonconvex Optimization and Its Applications

VOLUME 78

QUADRATIC PROGRAMMING AND AFFINE VARIATIONAL INEQUALITIES

A Qualitative Study

By

GUE MYUNG LEE
Pukyong National University, Republic of Korea

NGUYEN NANG TAM
Hanoi Pedagogical Institute No. 2, Vietnam

NGUYEN DONG YEN
Vietnamese Academy of Science and Technology, Vietnam

Springer

Library of Congress Cataloging-in-Publication Data

A C.I.P. record for this book is available from the Library of Congress.

ISBN 0-387-24277-5 e-ISBN 0-387-24278-3 Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11375562

springeronline.com

Contents

Preface

Quadratic programs and affine variational inequalities represent two fundamental, closely-related classes of problems in the theories of mathematical programming and variational inequalities, respectively. This book develops a unified theory on qualitative aspects of nonconvex quadratic programming and affine variational inequalities. The first seven chapters introduce the reader step-by-step to the central issues concerning a quadratic program or an affine variational inequality, such as the solution existence, necessary and sufficient conditions for a point to belong to the solution set, and properties of the solution set. The subsequent two chapters discuss briefly two concrete models (linear fractional vector optimization and the traffic equilibrium problem) whose analysis can benefit a lot from using the results on quadratic programs and affine variational inequalities. There are six chapters devoted to the study of continuity and/or differentiability properties of the characteristic maps and functions in quadratic programs and in affine variational inequalities where all the components of the problem data are subject to perturbation. Quadratic programs and affine variational inequalities under linear perturbations are studied in three other chapters. One special feature of the presentation is that when a certain property of a characteristic map or function is investigated, we always try first to establish necessary conditions for it to hold, then we go on to study whether the obtained necessary conditions are sufficient ones. This helps to clarify the structures of the two classes of problems under consideration. The qualitative results can be used for dealing with algorithms and applications related to quadratic programming problems and affine variational inequalities.

This book can be useful for postgraduate students in applied mathematics and for researchers in the field of nonlinear programming and equilibrium problems. It can be used for some advanced courses on nonconvex quadratic programming and affine variational inequalities.

Among many references in the field discussed in this monograph, we would like to mention the following well-known books: *"Linear and Combinatorial Programming"* by K. G. Murty (1976), *"Non-Linear Parametric Optimization"* by B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer (1982), and *"The Linear Complementarity Problem"* by R. W. Cottle, J.-S. Pang and R. E. Stone (1992).

As for prerequisites, the reader is expected to be familiar with the basic facts of Linear Algebra, Functional Analysis, and Convex Analysis.

We started writing this book in Pusan (Korea) and completed our writing in Hanoi (Vietnam). This book would not be possible without the financial support from the Korea Research Foundation (Grant KRF 2000-015-DP0044), the Korean Science and Engineering Foundation (through the APEC Postdoctoral Fellowship Program and the Brain Pool Program), the National Program in Basic Sciences (Vietnam).

We would like to ask the international publishers who have published some of our research papers in their journals or proceedings volumes for letting us to use a re-edited form of these papers for this book. We thank them a lot for their kind permission.

We would like to express our sincere thanks to the following experts for their kind help or generous encouragement at different times in our research related to this book: Prof. Y. J. Cho, Dr. N. H. Dien, Prof. P. H. Dien, Prof. F. Giannessi, Prof. J. S. Jung, Prof. P. Q. Khanh, Prof. D. S. Kim, Prof. J. K. Kim, Prof. S. Kum, Prof. M. Kwapisz, Prof. B. S. Lee, Prof. D. T. Luc, Prof. K. Malanowski, Prof. C. Malivert, Prof. A. Maugeri, Prof. L. D. Muu, Prof. A. Nowakowski, Prof. S. Park, Prof. J.-P. Penot, Prof. V. N. Phat, Prof. H. X. Phu, Dr. T. D. Phuong, Prof. B. Ricceri, Prof. P. H. Sach, Prof. N. K. Son, Prof. M. Studniarski, Prof. M. Théra, Prof. T. D. Van. Also, it is our pleasant duty to thank Mr. N. Q. Huy for his efficient cooperation in polishing some arguments in the proof of Theorem 8.1.

The late Professor W. Oettli had a great influence on our research on quadratic programs and affine variational inequalities. We always remember him with sympathy and gratefulness.

We would like to thank Professor P. M. Pardalos for supporting our plan of writing this monograph.

This book is dedicated to our parents. We thank our families for patience and encouragement.

Any comment on this book will be accepted with sincere thanks.

May 2004

Guc Myung Lee, Nguyen Nang Tam, and Nguyen Dong Yen

Notations and Abbreviations

N	the set of the positive integers
R	the real line
$\overline{R}$	the extended real line
R^n	the $n-$dimensional Euclidean space
R^n_+	the nonnegative orthant in R^n
$\emptyset$	the empty set
x^T	the transpose of vector x
$\|x\|$	the norm of vector x
$\langle x, y\rangle$	the scalar product of x and y
A^T	the transpose of matrix A
rank$A$	the rank of matrix A
$\|A\|$	the norm of matrix A
$R^{m\times n}$	the set of the $m \times n$-matrices
$R^{n\times n}_S$	the set of the symmetric $n \times n$-matrices
det$A$	the determinant of a square matrix A
E	the unit matrix in $R^{n\times n}$
$B(x, \delta)$	the open ball centered at x with radius δ
$\bar{B}(x, \delta)$	the closed ball centered at x with radius δ
$\bar{B}_{R^n}$	the closed unit ball in R^n
intΩ	the interior of Ω
$\overline{\Omega}$	the closure of Ω
bdΩ	the boundary of Ω
coΩ	the convex hull of Ω
dist(x, Ω)	the distance from x to Ω
cone$M$	the cone generated by M
riΔ	the relative interior of a convex set Δ
affΔ	the affine hull of Δ
extrΔ	the set of the extreme points of Δ
$0^+\Delta$	the recession cone of Δ
$T_\Delta(\bar{x})$	the tangent cone to Δ at $\bar{x}$
$N_\Delta(\bar{x})$	the normal cone to Δ at $\bar{x}$
$M^\perp$	the linear subspace of R^n orthogonal to $M \subset R^n$
$\mathrm{Pr}_K(\cdot)$ or $\mathrm{P}_K(\cdot)$	the metric projection from R^n onto a closed convex subset $K \subset R^n$

$\mathrm{dom}f$	the effective domain of function f
$f'(\bar{x};v)$	the directional derivative of f at $\bar{x}$ in direction v
$f^0(\bar{x};v)$	the Clarke generalized directional derivative of f at $\bar{x}$ in direction v
$\partial f(\bar{x})$	the subdifferential of a convex function f at $\bar{x}$, or the Clarke generalized gradient of a locally Lipschitz function f at $\bar{x}$
$\nabla f(\bar{x})$	the gradient of f at $\bar{x}$
$\nabla^2 f(\bar{x})$	the Hessian matrix of f at $\bar{x}$
$\mathrm{Sol}(P)$	the solution set of problem (P)
$\mathrm{loc}(P)$	the local solution set of problem (P)
$S(P)$	the KKT point set of problem (P)
$v(P)$	the optimal value of problem (P)
QP	quadratic programming
$QP(D,A,c,b)$	quadratic program defined by matrices D, A and vectors c, b
$S(D,A,c,b)$	the KKT point set of a quadratic program
$\mathrm{Sol}(D,A,c,b)$	the solution set of a quadratic program
$\mathrm{Sol}(c,b)$	the solution set of a quadratic program
$\varphi(D,A,c,b)$ or $\varphi(c,b)$	the optimal value function of a quadratic program
VI	variational inequality
$\mathrm{loc}(D,A,c,b)$	the local-solution set of a quadratic program
$\mathrm{VI}(\phi,\Delta)$	the VI defined by operator ϕ and set Δ
AVI	affine variational inequality
$\mathrm{Sol}(\mathrm{VI}(\phi,\Delta))$	the solution set of $\mathrm{VI}(\phi,\Delta)$
$\mathrm{AVI}(M,q,\Delta)$	the AVI defined by matrix M, vector q, and set Δ
$\mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$	the solution set of $\mathrm{AVI}(M,q,\Delta)$
$\mathrm{Sol}(M,A,q,b)$	the solution set of $\mathrm{AVI}(M,q,\Delta)$ where $\Delta=\{x\ :\ Ax\geq b\}$
LCP	linear complementarity
$\mathrm{LCP}(M,q)$	the LCP problem defined by matrix M and vector q
$\mathrm{Sol}(M,q)$	the solution set of $\mathrm{LCP}(M,q)$

NCP	nonlinear complementarity
NCP(ϕ, Δ)	the NCP problem defined by ϕ and Δ
LFVO	linear fractional vector optimization
VVI	vector variational inequality
Sol(VP)	the efficient solution set of the LFVO problem (VP)
Sol(VP)^w	the weakly efficient solution set of the LFVO problem (VP)
lsc	lower semicontinuous
lsc property	lower semicontinuity property
usc	upper semicontinuous
usc property	upper semicontinuity property
$OD-$pair	origin-destination pair
plq	piecewise linear-quadratic
PVI	parametric variational inequality

Chapter 1

Quadratic Programming Problems

Quadratic programming problems constitute a special class of nonlinear mathematical programming problems. This chapter presents some preliminaries related to mathematical programming problems including the quadratic programming problems. The subsequent three chapters will provide a detailed exposition of the basic facts on quadratic programming problems, such as the solution existence, first-order optimality conditions, second-order optimality conditions, and properties of the solution sets.

1.1 Mathematical Programming Problems

Many practical and theoretical problems can be modeled in the form

$$(P) \qquad \text{Minimize } f(x) \quad \text{subject to} \quad x \in \Delta,$$

where $f : R^n \to \overline{R}$ is a given function, $\Delta \subset R^n$ is a given subset. Here and subsequently, $\overline{R} = [-\infty, +\infty] = R \cup \{-\infty\} \cup \{+\infty\}$ denotes the extended real line, R^n stands for the n-dimensional Euclidean space with the norm

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

for all $x = (x_1, \ldots, x_n) \in R^n$ and the scalar product

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T y$$

for all $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in R^n$. Here and subsequently, the apex ^{T} denotes the matrix transposition. In the text, vectors are expressed as rows of real numbers; while in the matrix computations they are understood as columns of real numbers. The open ball in R^n centered at x with radius $\delta > 0$ is denoted by $B(x, \delta)$. The corresponding closed ball is denoted by $\bar{B}(x, \delta)$. Thus

$$B(x,\delta) = \{y \in R^n : \|y-x\| < \delta\},\ \bar{B}(x,\delta) = \{y \in R^n : \|y-x\| \leq \delta\}.$$

The unit ball $\bar{B}(0,1)$ will be frequently denoted by $\bar{B}_{R^n}$. For a set $\Omega \subset R^n$, the notations intΩ, $\overline{\Omega}$ and bdΩ, respectively, are used to denote the topological *interior*, the topological *closure* and the *boundary* of Ω. Thus $\overline{\Omega}$ is the smallest closed subset in R^n containing Ω, and

$$\text{int}\Omega = \{x \in \Omega \,:\, \exists \varepsilon > 0 \text{ s.t. } B(x,\varepsilon) \subset \Omega\}, \quad \text{bd}\Omega = \overline{\Omega} \setminus (\text{int}\Omega).$$

We say that $U \subset R^n$ is a *neighborhood* of $x \in R^n$ if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset U$. Sometimes instead of (P) we write the following

$$\min\{f(x) \,:\, x \in \Delta\}.$$

Definition 1.1. We call (P) a *mathematical programming problem.* We call f the *objective function* and Δ the *constraint set* (also the *feasible region*) of (P). Elements of Δ are said to be the *feasible vectors* of (P). If $\Delta = R^n$ then we say that (P) is an unconstrained problem. Otherwise (P) is called a constrained problem.

Definition 1.2 (cf. Rockafellar and Wets (1998), p. 4) A feasible vector $\bar{x} \in \Delta$ is called a (global) *solution* of (P) if $f(\bar{x}) \neq +\infty$ and $f(x) \geq f(\bar{x})$ for all $x \in \Delta$. We say that $\bar{x} \in \Delta$ is a *local solution* of (P) if $f(\bar{x}) \neq +\infty$ and there exists a neighborhood U of $\bar{x}$ such that

$$f(x) \geq f(\bar{x}) \quad \text{for all} \quad x \in \Delta \cap U. \tag{1.1}$$

The set of all the solutions (resp., the local solutions) of (P) is denoted by Sol(P) (resp., loc(P)). We say that two mathematical programming problems are *equivalent* if the solution set of the first problem coincides with that of the second one.

Definition 1.3. The *optimal value* $v(P)$ of (P) is defined by setting

$$v(P) = \inf\{f(x) \,:\, x \in \Delta\}. \tag{1.2}$$

If $\Delta = \emptyset$ then, by convention, $v(P) = +\infty$.

Remark 1.1. It is clear that $\mathrm{Sol}(P) \subset \mathrm{loc}(P)$. It is also obvious that

$$\mathrm{Sol}(P) = \{x \in \Delta \,:\, f(x) \neq +\infty,\ f(x) = v(P)\}.$$

Remark 1.2. It may happen that $\mathrm{loc}(P) \setminus \mathrm{Sol}(P) \neq \emptyset$. For example, if we choose $\Delta = [-1, +\infty)$ and $f(x) = 2x^3 - 3x^2 + 1$ then $\bar{x} = 1$ is a local solution of (P) which is not a global solution.

Remark 1.3. Instead of the minimization problem (P), one may encounter with the following maximization problem

$$(P_1) \qquad \text{Maximize} \quad f(x) \quad \text{subject to} \quad x \in \Delta.$$

A point $\bar{x} \in \Delta$ is said to be a (global) solution of (P_1) if $f(\bar{x}) \neq -\infty$ and $f(x) \leq f(\bar{x})$ for all $x \in \Delta$. We say that $\bar{x} \in \Delta$ is a local solution of (P_1) if $f(\bar{x}) \neq -\infty$ and there exists a neighborhood U of $\bar{x}$ such that $f(x) \leq f(\bar{x})$ for all $x \in \Delta \cap U$. It is clear that $\bar{x}$ is a solution (resp., a local solution) of (P_1) if and only if $\bar{x}$ is a solution (resp., a local solution) of the following minimization problem

$$\text{Minimize} \quad -f(x) \quad \text{subject to} \quad x \in \Delta.$$

Thus any maximization problem of the form (P_1) can be reduced to a minimization problem of the form (P).

Remark 1.4. Even in the case $v(P)$ is a finite real number, it may happen that $\mathrm{Sol}(P) = \emptyset$. For example, if $\Delta = [1, +\infty) \subset R$ and

$$f(x) = \begin{cases} \dfrac{1}{|x|} & \text{for } x \neq 0 \\ +\infty & \text{for } x = 0 \end{cases}$$

then $v(P) = 0$, while $\mathrm{Sol}(P) = \emptyset$.

There are different ways to classify mathematical programming problems:

- Convex vs. Nonconvex
- Smooth vs. Nonsmooth
- Linear vs. Nonlinear.

1.2 Convex Programs and Nonconvex Programs

Definition 1.4. We say that $\Delta \subset R^n$ is a *convex set* if

$$tx + (1-t)y \in \Delta \quad \text{for every } x \in \Delta,\ y \in \Delta \text{ and } t \in (0,1). \tag{1.3}$$

The smallest convex set containing a set $\Omega \subset R^n$ is called the *convex hull* of Ω and it is denoted by $\mathrm{co}\Omega$.

Definition 1.5. A function $f : R^n \to \overline{R}$ is said to be *convex* if its epigraph

$$\mathrm{epi}f := \{(x,\alpha) : x \in R^n,\ \alpha \in R,\ \alpha \geq f(x)\} \tag{1.4}$$

is a convex subset of the product space $R^n \times R$. A convex function f is said to be *proper* if $f(x) < +\infty$ for at least one $x \in R^n$ and $f(x) > -\infty$ for all $x \in R^n$. A function $f : R^n \to \overline{R}$ is said to be *concave* if the function $-f$ defined by the formula $(-f)(x) = -f(x)$ is convex.

By the usual convention (see Rockafellar (1970), p. 24),

$$\begin{aligned}
&\alpha + (+\infty) = (+\infty) + \alpha = +\infty \quad \text{for } -\infty < \alpha \leq +\infty,\\
&\alpha + (-\infty) = (-\infty) + \alpha = -\infty \quad \text{for } -\infty \leq \alpha < +\infty,\\
&\alpha(+\infty) = (+\infty)\alpha = +\infty,\ \ \alpha(-\infty) = (-\infty)\alpha = -\infty,\\
&\qquad \text{for } 0 < \alpha \leq +\infty,\\
&\alpha(+\infty) = (+\infty)\alpha = -\infty,\ \ \alpha(-\infty) = (-\infty)\alpha = +\infty,\\
&\qquad \text{for } -\infty \leq \alpha < 0,\\
&0(+\infty) = (+\infty)0 = 0 = 0(-\infty) = (-\infty)0,\\
&-(-\infty) = +\infty,\ \ \inf\emptyset = +\infty,\ \ \sup\emptyset = -\infty.
\end{aligned}$$

The combinations $(+\infty)+(-\infty)$ and $(-\infty)+(+\infty)$ have no meaning and will be avoided.

Note that a function $f : R^n \to R \cup \{+\infty\}$ is convex if and only if

$$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y),\ \forall x, y \in R^n,\ \forall t \in (0,1). \tag{1.5}$$

Indeed, by definition, f is convex if and only if the set $\mathrm{epi}f$ defined in (1.4) is convex. This means that

$$t(x,\alpha) + (1-t)(y,\beta) \in \mathrm{epi}f$$

for all $t \in (0,1)$ and for all $x, y \in R^n$, $\alpha, \beta \in R$ satisfying $\alpha \geq f(x)$, $\beta \geq f(y)$. It is a simple matter to show that the latter is equivalent to (1.5).

More generally, a function $f : R^n \to R \cup \{+\infty\}$ is convex if and only if

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k) \quad \text{(Jensen's Inequality)}$$

whenever $x_1, \ldots, x_k \in R^n$ and $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0$, $\lambda_1 + \cdots + \lambda_k = 1$. (See Rockafellar (1970), Theorem 4.3).

Definition 1.6. We say that (P) is a *convex program* (a convex mathematical programming problem) if Δ is a convex set and f is a convex function.

Proposition 1.1. *If (P) is a convex program then*

$$\mathrm{Sol}(P) = \mathrm{loc}(P). \tag{1.6}$$

Proof. It suffices to show that $\mathrm{loc}(P) \subset \mathrm{Sol}(P)$ whenever (P) is a convex program. Let $\bar{x} \in \mathrm{loc}(P)$ and let U be a neighborhood of $\bar{x}$ such that (1.1) holds. If $\bar{x} \notin \mathrm{Sol}(P)$ then there must exist $\hat{x} \in \Delta$ such that $f(\hat{x}) < f(\bar{x})$. Since $f(\bar{x}) \neq +\infty$, this implies that $f(\hat{x}) \in R \cup \{-\infty\}$.

We first consider the case $f(\hat{x}) \neq -\infty$. For any $t \in (0,1)$, we have

$$\begin{aligned} f(t\hat{x} + (1-t)\bar{x}) &\leq tf(\hat{x}) + (1-t)f(\bar{x}) \\ &< tf(\bar{x}) + (1-t)f(\bar{x}) = f(\bar{x}). \end{aligned} \tag{1.7}$$

Since $t\hat{x} + (1-t)\bar{x} = \bar{x} + t(\hat{x} - \bar{x})$ belongs to $\Delta \cap U$ for sufficiently small $t \in (0,1)$, (1.7) contradicts (1.1).

We now consider the case $f(\hat{x}) = -\infty$. Fix any $t \in (0,1)$. For every $\alpha \in R$, since $(\hat{x}, \alpha) \in \mathrm{epi}f$ and $(\bar{x}, f(\bar{x})) \in \mathrm{epi}f$, we have

$$t(\hat{x}, \alpha) + (1-t)(\bar{x}, f(\bar{x})) \in \mathrm{epi}f.$$

Hence $f(t\hat{x} + (1-t)\bar{x}) \leq t\alpha + (1-t)f(\bar{x})$ for all $\alpha \in R$. This implies that $f(t\hat{x} + (1-t)\bar{x}) = -\infty$. Since the last equality is valid for all $t \in (0,1)$ and $t\hat{x} + (1-t)\bar{x} \in \Delta \cap U$ if $t \in (0,1)$ is sufficiently small, (1.1) cannot hold. We have arrived at a contradiction. □

Definition 1.7. If Δ is nonconvex (= not convex) or f is nonconvex then we say that (P) is a *nonconvex program* (a nonconvex mathematical programming problem).

Example 1.1. Consider the problem

$$\min\{f(x) = (x_1 - c_1)^2 + (x_2 - c_2)^2 \; : \; x \in \Delta\}, \tag{1.8}$$

where $\Delta = \{x = (x_1, x_2) : x_1 \geq 0\} \cup \{x = (x_1, x_2) : x_2 \geq 0\}$ and $c = (c_1, c_2) = (-2, -1)$. Note that f is convex, while Δ is nonconvex. It is clear that (1.8) is equivalent to the following problem

$$\min\{\|x - c\| \; : \; x \in \Delta\}. \tag{1.9}$$

On can easily verify that the solution set of (1.8) and (1.9) consists of only one point $(-2, 0)$, and the local solution set contains two points: $(-2, 0)$ and $(0, -1)$.

Example 1.2. Let $f_1(x) = -x + 2$, $f_2(x) = x + \dfrac{3}{2}$, $x \in R$. Define $f(x) = \min\{f_1(x), \; f_2(x)\}$ and choose $\Delta = [0, 2] \subset R$. For these f and Δ, we have

$$\mathrm{Sol}(P) = \{2\}, \quad \mathrm{loc}(P) = \{0, \; 2\}.$$

Note that in this example f is a nonconvex function, while Δ is a convex set.

Convex functions have many nice properties. For example, a convex function is continuous at any interior point of its effective domain and it is directionally differentiable at any point in the domain.

Definition 1.8. For a function $f : R^n \to \overline{R}$, the set

$$\mathrm{dom} f := \{x \in R^n \; : \; -\infty < f(x) < +\infty\} \tag{1.10}$$

is called the *effective domain* of f. For a point $\bar{x} \in \mathrm{dom} f$ and a vector $v \in R^n$, if the limit

$$f'(\bar{x}; v) := \lim_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t} \tag{1.11}$$

(which may have the values $+\infty$ and $-\infty$) exists then f is said to be *directionally differentiable* at $\bar{x}$ in direction v and the value $f'(\bar{x}; v)$ is called the *directional derivative* of f at $\bar{x}$ in direction v. If $f'(\bar{x}; v)$ exists for all $v \in R^n$ then f is said to be directionally differentiable at $\bar{x}$.

In the next two theorems, $f : R^n \to R \cup \{+\infty\}$ is a proper convex function.

Theorem 1.1. (See Rockafellar (1970), Theorem 10.1) *If $\bar{x} \in R^n$ and $\delta > 0$ are such that the open ball $B(\bar{x}, \delta)$ is contained in* dom$f$, *then the restriction of f to $B(\bar{x}, \delta)$ is a continuous real function.*

Theorem 1.2. (See Rockafellar (1970), Theorem 23.1) *If $\bar{x} \in$* dom$f$ *then for any $v \in R^n$ the limit*

$$f'(\bar{x}; v) := \lim_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}$$

exists, and one has

$$f'(\bar{x}; v) = \inf_{t > 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}.$$

Definition 1.9. The *normal cone* $N_\Delta(\bar{x})$ to a convex set $\Delta \subset R^n$ at a point $\bar{x} \in R^n$ is defined by the formula

$$N_\Delta(\bar{x}) = \begin{cases} \{x^* \in R^n \, : \, \langle x^*, x - \bar{x} \rangle \le 0 \text{ for all } x \in \Delta\} & \text{if } \bar{x} \in \Delta \\ \emptyset \quad \text{if } \bar{x} \notin \Delta. \end{cases} \tag{1.12}$$

Definition 1.10. The *subdifferential* $\partial f(\bar{x})$ of a convex function $f : R^n \to \overline{R}$ at a point $\bar{x} \in R^n$ is defined by setting

$$\partial f(\bar{x}) = \{x^* \in R^n \, : \, f(\bar{x}) + \langle x^*, x - \bar{x} \rangle \le f(x) \text{ for every } x \in R^n\}. \tag{1.13}$$

Definition 1.11. A subset $M \subset R^n$ is called an *affine set* if $tx + (1-t)y \in M$ for every $x \in M$, $y \in M$ and $t \in R$. For a convex set $\Delta \subset R^n$, the *affine hull* affΔ of Δ is the smallest affine set containing Δ. The *relative interior* of Δ is defined by the formula

$$\text{ri}\Delta = \{x \in \Delta \, : \, \exists \delta > 0 \text{ such that } B(x, \delta) \cap \text{aff}\Delta \subset \Delta\}.$$

The following statement describes the relation between the directional derivative and the subdifferential of convex functions.

Theorem 1.3. (See Rockafellar (1970), Theorem 23.4) *Let f be a proper convex function on R^n. If $x \notin$* dom$f$ *then $\partial f(x)$ is empty. If $x \in$* ri(dom$f$) *then $\partial f(x)$ is nonempty and*

$$f'(x; v) = \sup\{\langle x^*, v \rangle \, : \, x^* \in \partial f(x)\}, \quad \forall v \in R^n.$$

Besides, $\partial f(x)$ is a nonempty bounded set if and only if

$$x \in \text{int}(\text{dom}f),$$

in which case $f'(x;v)$ is finite for every $v \in R^n$.

The following result is called the Moreau-Rockafellar Theorem.

Theorem 1.4. (See Rockafellar (1970), Theorem 23.8) *Let $f = f_1 + \cdots + f_k$, where $f_1, \ldots, f_k$ are proper convex functions on R^n. If*

$$\bigcap_{i=1}^{k} \mathrm{ri}(\mathrm{dom} f_i) \neq \emptyset$$

then

$$\partial f(x) = \partial f_1(x) + \cdots + \partial f_k(x), \quad \forall x \in R^n.$$

First-order necessary and sufficient optimality conditions for convex programs can be stated as follows.

Theorem 1.5. (See Rockafellar (1970), Theorem 27.4) *Suppose that f is a proper convex function on R^n and $\Delta \subset R^n$ is a nonempty convex set. If the inclusion*

$$0 \in \partial f(\bar{x}) + N_\Delta(\bar{x}) \tag{1.14}$$

holds for some $\bar{x} \in R^n$, then $\bar{x}$ is a solution of (P). Conversely, if

$$\mathrm{ri}(\mathrm{dom} f) \cap \mathrm{ri}\Delta \neq \emptyset \tag{1.15}$$

then (1.14) *is a necessary and sufficient condition for $\bar{x} \in R^n$ to be a solution of (P). In particular, if $\Delta = R^n$ then $\bar{x}$ is a solution of (P) if and only if $0 \in \partial f(\bar{x})$.*

Inclusion (1.14) means that there exist $x^* \in \partial f(\bar{x})$ and $u^* \in N_\Delta(\bar{x})$ such that $0 = x^* + u^*$. Note that (1.15) is a regularity condition for convex programs of the type (P).

The facts stated in Proposition 1.1 and Theorem 1.5 are the most characteristic properties of convex mathematical programming problems.

Theorem 1.5 can be used for solving effectively many convex programs. For illustration, let us consider the following example.

Example 1.3. (The Fermat point) Let A, B, C be three points in the two-dimensional space R^2 with the coordinates

$$a = (a_1, a_2), \quad b = (b_1, b_2), \quad c = (c_1, c_2),$$

respectively. Assume that there exists no straight line containing all the three points. The problem consists of finding a point M in R^2

with the coordinates $\bar{x} = (\bar{x}_1, \bar{x}_2)$ such that the sum of the distances from M to A, B and C is minimal. This amounts to saying that $\bar{x}$ is a solution of the following unconstrained convex program:

$$\min\{f(x) := \|x-a\| + \|x-b\| + \|x-c\| \; : \; x \in R^2\}. \tag{1.16}$$

In Lemma 1.1 below it will be proved that problem (1.16) has solutions and the solution set is a singleton. Note that $f = f_1 + f_2 + f_3$, where $f_1(x) = \|x-a\|$, $f_2(x) = \|x-b\|$, $f_3(x) = \|x-c\|$. By Theorem 1.5, $\bar{x}$ is a solution of (1.16) if and only if $0 \in \partial f(\bar{x})$. As $\mathrm{dom} f_i = R^2$ $(i = 1,2,3)$, using Theorem 1.4 we can write the last inclusion in the following equivalent form

$$0 \in \partial f_1(\bar{x}) + \partial f_2(\bar{x}) + \partial f_3(\bar{x}). \tag{1.17}$$

We first consider the case where $\bar{x}$ coincides with one of the three vectors a, b, c. Let $\bar{x} = a$, i.e. $M \equiv A$. In this case,

$$\partial f_1(\bar{x}) = \bar{B}_{R^2}, \;\; \partial f_2(\bar{x}) = \left\{ \frac{a-b}{\|a-b\|} \right\}, \;\; \partial f_3(\bar{x}) = \left\{ \frac{a-c}{\|a-c\|} \right\}.$$

Hence (1.17) is equivalent to saying that there exists $u^* \in \bar{B}_{R^2}$ such that

$$0 = u^* - v^* - w^*, \tag{1.18}$$

where $v^* := (b-a)/\|b-a\|$, $w^* := (c-a)/\|c-a\|$. From (1.18) it follows that

$$\begin{aligned} 1 \geq \|u^*\|^2 &= \langle u^*, u^* \rangle \\ &= \langle v^* + w^*, v^* + w^* \rangle \\ &= \|v^*\|^2 + \|w^*\|^2 + 2\langle v^*, w^* \rangle. \end{aligned}$$

As $\|v^*\| = 1$ and $\|w^*\| = 1$, this yields $\langle v^*, w^* \rangle \leq -\frac{1}{2}$. Denoting by α the geometric angle between the vectors v^* and w^* (which is equal to angle $\hat{A}$ of the triangle ABC), we deduce from the last inequality that

$$\cos\alpha = \frac{\langle v^*, w^* \rangle}{\|v^*\| \|w^*\|} = \langle v^*, w^* \rangle \leq -\frac{1}{2}.$$

Hence

$$\frac{2\pi}{3} \leq \alpha < \pi. \tag{1.19}$$

(The case $\alpha = \pi$ is excluded because there exists no straight line containing A, B and C.) It is easy to show that (1.19) implies that

$\widetilde{u}^* := v^* + w^*$ belongs to $\bar{B}_{R^2}$. Thus (1.19) is equivalent to (1.17). This means that (1.19) holds if and only if $\bar{x} = a$ is a solution of (1.16).

We now turn to the case where $\bar{x} \neq a$, $\bar{x} \neq b$ and $\bar{x} \neq c$, i.e. M does not coincide with anyone from the three vertexes A, B, C of the triangle ABC. In this case, as

$$\partial f_1(\bar{x}) = \left\{ \frac{\bar{x} - a}{\|\bar{x} - a\|} \right\}, \ \partial f_2(\bar{x}) = \left\{ \frac{\bar{x} - b}{\|\bar{x} - b\|} \right\}, \ \partial f_3(\bar{x}) = \left\{ \frac{\bar{x} - c}{\|\bar{x} - c\|} \right\},$$

(1.17) is equivalent to the equality

$$0 = u^* + v^* + w^*, \tag{1.20}$$

where $u^* := (a - \bar{x})/\|a - \bar{x}\|$, $v^* := (b - \bar{x})/\|b - \bar{x}\|$ and $w^* := (c - \bar{x})/\|c - \bar{x}\|$. By (1.20),

$$\begin{aligned} 1 = \|u^*\|^2 &= \langle u^*, u^* \rangle \\ &= \langle -v^* - w^*, -v^* - w^* \rangle \\ &= \|v^*\|^2 + \|w^*\|^2 + 2\langle v^*, w^* \rangle. \end{aligned}$$

Since $\|v^*\| = 1$ and $\|w^*\| = 1$, this implies that $\langle v^*, w^* \rangle = -\dfrac{1}{2}$. Hence the geometric angle α between v^* and w^* is $2\pi/3$. Similarly, we deduce from (1.20) that the geometric angle β (resp., γ) between u^* and w^* (resp., between u^* and v^*) is equal to $2\pi/3$. (Geometrically, we have shown that M sees the edges BC, AC and AB of the triangle ABC under the same angle 120^o.) It is easily seen that if

$$\alpha = \beta = \gamma = \frac{2\pi}{3}$$

then (1.20) is satisfied; hence (1.17) is valid and $\bar{x}$ is a solution of (1.16).

Summarizing all the above in the language of Euclidean Geometry, we have the following conclusions:

(i) If one of the three angles of the triangle ABC, say $\hat{A}$, is larger than or equal to 120^o, then $M \equiv A$ is the unique solution of our problem.

(ii) If all the three angles of the triangle ABC are smaller than 120^o, then the unique solution of our problem is the point M seeing the edges BC, AC and AB of the triangle ABC

under the same angle 120^o. (This special point M is called the *Fermat point* or the *Torricelli point* (see Weisstein (1999)). It can be proved that the Fermat point belongs to the interior of the triangle ABC.)

If the necessary and sufficient optimality condition stated in Theorem 1.5 yields a unique point $\bar{x}$ which can be expressed explicitly via the data of the optimization problem (see, for instance, the situation in Example 1.6 below) then the problem has solutions and the solution set is a singleton. In the other case, information about the solution existence and uniqueness can be obtained by analyzing furthermore the structure of the problem under consideration.

For the illustrative problem described in Example 1.3, the following statement is valid.

Lemma 1.1. *Let $a = (a_1, a_2)$, $b = (b_1, b_2)$, $c = (c_1, c_2)$ be given points in R^2 such that there exists no straight line containing all the three points. Then problem* (1.16) *has solutions and the solution set is a singleton.*

Proof. In order to show that (1.16) has solutions, we observe that

$$f(x) \geq 3\|x\| - \|a\| - \|b\| - \|c\|.$$

Therefore $\lim\limits_{\|x\|\to+\infty} f(x) = +\infty$. Fix any $z \in R^2$ and put $\gamma = f(z)$. Let $\varrho \in [\|z\|, +\infty)$ be such that

$$f(x) > \gamma \quad \text{for every } x \in R^2 \setminus \bar{B}(0, \varrho).$$

By the Weierstrass Theorem, the restriction of the continuous function $f(x)$ on the compact set $\bar{B}(0, \varrho)$ achieves minimum at some point $\bar{x} \in \bar{B}(0, \varrho)$, that is $f(\bar{x}) \leq f(y)$ for every $y \in \bar{B}(0, \varrho)$. Since

$$f(\bar{x}) \leq f(z) = \gamma < f(x) \quad \text{for all } x \in R^2 \setminus \bar{B}(0, \varrho),$$

it follows that $\bar{x}$ is a solution of (1.16).

We now prove that $f(x)$ is a *strictly convex function*, that is

$$f(tx + (1-t)y)) < tf(x) + (1-t)f(y)$$

for all x, y in R^2 with $x \neq y$ and for all $t \in (0, 1)$. Given any $x = (x_1, x_2)$, $y = (y_1, y_2)$ in R^2 with $x \neq y$ and $t \in (0, 1)$, we consider the following vector systems

$$\{x - a,\ y - a\}, \quad \{x - b,\ y - b\}, \quad \{x - c,\ y - c\}. \tag{1.21}$$

We claim that at least one of the three systems is linearly independent. Suppose the claim were false. Then we would have

$$\det \begin{pmatrix} x_1 - a_1 & y_1 - a_1 \\ x_2 - a_2 & y_2 - a_2 \end{pmatrix} = 0, \quad \det \begin{pmatrix} x_1 - b_1 & y_1 - b_1 \\ x_2 - b_2 & y_2 - b_2 \end{pmatrix} = 0,$$

$$\det \begin{pmatrix} x_1 - c_1 & y_1 - c_1 \\ x_2 - c_2 & y_2 - c_2 \end{pmatrix} = 0,$$

where $\det Z$ denotes the *determinant* of a square matrix Z. These equalities imply that

$$\begin{aligned} (x_1 - y_1)a_2 - (x_2 - y_2)a_1 &= x_1y_2 - x_2y_1, \\ (x_1 - y_1)b_2 - (x_2 - y_2)b_1 &= x_1y_2 - x_2y_1, \\ (x_1 - y_1)c_2 - (x_2 - y_2)c_1 &= x_1y_2 - x_2y_1. \end{aligned} \tag{1.22}$$

Since $x \neq y$, we have $(x_1 - y_1)^2 + (x_2 - y_2)^2 \neq 0$. So the set

$$L := \{z = (z_1, z_2) \in R^2 \, : \, (x_1 - y_1)z_2 - (x_2 - y_2)z_1 = x_1y_2 - x_2y_1\}$$

is a straight line in R^2. By (1.22), L contains all the points a, b, c. This contradicts our assumption. We have thus proved that at least one of the three vector systems in (1.21) is linearly independent. Without loss of generality, we can assume that the system $\{x-a,\, y-a\}$ is linearly independent. Then the system $\{t(x-a),\, (1-t)(y-a)\}$ is also linearly independent. This implies that

$$\|t(x-a) + (1-t)(y-a)\| < t\|x-a\| + (1-t)\|y-a\|.$$

So we have

$$\begin{aligned} f(tx + (1-t)y) &= \|tx + (1-t)y - a\| + \|tx + (1-t)y - b\| \\ &\quad + \|tx + (1-t)y - c\| \\ &= \|t(x-a) + (1-t)(y-a)\| \\ &\quad + \|t(x-b) + (1-t)(y-b)\| \\ &\quad + \|t(x-c) + (1-t)(y-c)\| \\ &< t\|x-a\| + (1-t)\|y-a\| \\ &\quad + t\|x-b\| + (1-t)\|y-b\| \\ &\quad + t\|x-c\| + (1-t)\|y-c\| \\ &= tf(x) + (1-t)f(y). \end{aligned}$$

The strict convexity of f has been established. From this property it follows immediately that (1.16) cannot have more than one solution.

Indeed, if there were two different solutions x and y of the problem, then by the strict convexity of f we would have

$$f\left(\frac{1}{2}x+\frac{1}{2}y\right) < \frac{1}{2}f(x)+\frac{1}{2}f(y)=f(x).$$

This contradicts the fact that x is a solution of (1.16). The proof of the lemma is complete. □

Remark 1.5. It follows from the above results that (1.16) admits a unique solution belonging to the convex hull of the set $\{a,\, b,\, c\}$. Hence (1.16) is equivalent to the following constrained convex program

$$\min\{\|x-a\|+\|x-b\|+\|x-c\| \,:\, x\in \text{co}\{a,\, b,\, c\}\}.$$

In problem (P), if Δ is the solution set of a system of inequalities and equalities then first-order optimality conditions can be written in a form involving some *Lagrange multipliers.*

Let us consider problem (P) under the assumptions that $f : R^n \to R$ is a convex function and

$$\Delta=\{x\in R^n : g_1(x)\le 0,\dots,g_m(x)\le 0,\, h_1(x)=0,\dots,h_s(x)=0\}, \tag{1.23}$$

where $g_i : R^n \to R$ for $i=1,\dots,m$ is a convex function, $h_j : R^n \to R$ for $j=1,\dots,s$ is an *affine function*, i.e. there exist $a_j\in R^n$ and $\alpha_j\in R$ such that $h_j(x)=\langle a_j,x\rangle+\alpha_j$ for every $x\in R^n$. It is admitted that the equality constraints (resp., the equality constraints) can be absent in (1.23). For abbreviation, we use the formal writing $m=0$ (resp., $s=0$) to indicate that all the inequality constraints (resp., all the equality constraints) in (1.23) are absent.

Theorem 1.6. (Kuhn-Tucker Theorem for convex programs; see Rockafellar (1970), p. 283) *Let (P) be a convex program where Δ is given by* (1.23). *Let the above assumptions on f, g_i $(i=1,\dots,m)$ and h_j $(j=1,\dots s)$ be satisfied. Assume that there exists a vector $z\in R^n$ such that*

$$g_i(z)<0 \text{ for } i=1,\dots,m \quad\text{and}\quad h_j(z)=0 \text{ for } j=1,\dots,s. \tag{1.24}$$

Then $\bar{x}$ is a solution of (P) if and only if there exist $m+s$ real numbers $\lambda_1,\dots,\lambda_m$, $\mu_1,\dots,\mu_s$, which are called the Langrange multipliers corresponding to $\bar{x}$, such that the following Kuhn-Tucker conditions are fulfilled:

(a) $\lambda_i \geq 0,\ g_i(\bar{x}) \leq 0$ *and* $\lambda_i f_i(\bar{x}) = 0$ *for* $i = 1, \ldots, m$,

(b) $h_j(\bar{x}) = 0$ *for* $j = 1, \ldots, s$,

(c) $0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial g_i(\bar{x}) + \sum_{j=1}^{s} \mu_j a_j$.

Note that (1.24) is a *constraint qualification* for convex programs. If $s = 0$ then it becomes

$$\exists z \in R^n \quad \text{s.t. } g_i(z) < 0 \text{ for } i = 1, \ldots, m. \quad \text{(The Slater condition)}$$

If $m = 0$ then (1.24) is equivalent to the requirement that Δ is nonempty. Actually, in that case condition (1.24) can be omitted in the formulation of Theorem 1.6.

1.3 Smooth Programs and Nonsmooth Programs

For brevity, if $f : R^n \to R$ is a continuously Fréchet differentiable function then we shall say that f is a C^1*-function.* Similarly, if f is twice continuously Fréchet differentiable function then we shall say that f is a C^2*-function.* The vector

$$\nabla f(\bar{x}) = \begin{pmatrix} \dfrac{\partial f(\bar{x})}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(\bar{x})}{\partial x_n} \end{pmatrix},$$

where $\dfrac{\partial f(\bar{x})}{\partial x_i}$ for $i = 1, \ldots, n$ denotes the partial derivative of f at $\bar{x}$ with respect to x_i, is called the *gradient* of f at $\bar{x}$. The matrix

$$\nabla^2 f(\bar{x}) = \begin{pmatrix} \dfrac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f(\bar{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\bar{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{x})}{\partial x_n \partial x_n} \end{pmatrix},$$

where $\dfrac{\partial^2 f(\bar{x})}{\partial x_j \partial x_i}$ denotes the second-order partial derivative of f at $\bar{x}$ w.r.t. x_j and x_i, is called the *Hessian matrix* of f at $\bar{x}$. It is well-known that if f is a C^1-function on R^n then f is directionally differentiable on R^n (see Definition 1.8) and

$$f'(\bar{x}; v) = \nabla f(\bar{x})v = \sum_{i=1}^{n} \frac{\partial f(\bar{x})}{\partial x_i} v_i,$$

for every $\bar{x} \in R^n$ and $v = (v_1, \ldots, v_n) \in R^n$.

Definition 1.12. We say that (P) is a *smooth program* (a smooth mathematical programming problem) if $f : R^n \to R$ is a C^1-function and Δ can be represented in the form (1.23) where $g_i : R^n \to R$ $(i = 1, \ldots, m)$ and $h_j : R^n \to R$ $(j = 1, \ldots, s)$ are C^1-functions. Otherwise, (P) is called a *nonsmooth program.*

We have considered problem (1.16) of finding the Fermat point. It is an example of nonsmooth programs. Function $f(x)$ in (1.16) is not a C^1-function. However, it is a *Lipschitz function* because

$$|f(x) - f(y)| \leq 3\|x - y\| \quad \text{for all } x, y \text{ in } R^2.$$

Definition 1.13. A function $f : R^n \to R$ is said to be a *locally Lipschitz* near $\bar{x} \in R^n$ if there exist a constant $\ell \geq 0$ and a neighborhood U of $\bar{x}$ such that

$$|f(x') - f(x)| \leq \ell\|x' - x\| \quad \text{for all } x, x' \text{ in } U.$$

If f is locally Lipschitz near every point in R^n then f is said to be a *locally Lipschitz function* on R^n. If f is locally Lipschitz near $\bar{x}$ then the *generalized directional derivative* of f at $\bar{x}$ in direction $v \in R^n$ is defined by

$$\begin{aligned} f^0(\bar{x}; v) \ &:= \limsup_{x \to \bar{x},\ t \downarrow 0} \frac{f(x + tv) - f(x))}{t} \\ &= \sup\{\xi \in R : \exists \text{ sequences } x_k \to \bar{x} \text{ and } t_k \to 0+ \\ &\qquad \text{such that } \xi = \lim_{k \to +\infty} \frac{f(x_k + t_k v) - f(x_k)}{t_k}\}. \end{aligned}$$

The *Clarke generalized gradient* of f at $\bar{x}$ is given by

$$\partial f(\bar{x}) := \{x^* \in R^n : f^0(\bar{x}; v) \geq \langle x^*, v\rangle \text{ for all } v \in R^n\}.$$

Theorem 1.7. (See Clarke (1983), Propositions 2.1.2, 2.2.4, 2.2.6 and 2.2.7) *Let $f : R^n \to R$ be a real function. Then the following assertions hold:*

(a) *If f is locally Lipschitz near $\bar{x} \in R^n$ then*

$$f^0(\bar{x}; v) = \max\{\langle x^*, v\rangle \ : \ x^* \in \partial f(\bar{x})\}$$

for every $v \in R^n$.

(b) *If f is a C^1-function then f is a locally Lipschitz function and $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$, $f^0(\bar{x}; v) = \langle \nabla f(\bar{x}), v\rangle$ for all $\bar{x} \in R^n$ and $v \in R^n$.*

(c) *If f is convex then f is a locally Lipschitz function and, for every $\bar{x} \in R^n$, the Clarke generalized gradient $\partial f(\bar{x})$ coincides with the subdifferential of f at $\bar{x}$ defined by formula* (1.13). *Besides, $f^0(\bar{x}; v) = f'(\bar{x}; v)$ for every $v \in R^n$.*

As concerning the above assertion (c), we note that the directional derivative $f'(\bar{x}; v)$ exists according to Theorem 1.2.

Definition 1.14. Let $C \subset R^n$ be a nonempty subset. The *Clarke tangent cone* $T_C(x)$ to C at $x \in C$ is the set of all $v \in R^n$ satisfying $d_C^0(x; v) = 0$, where $d_C^0(x; v)$ denotes the generalized directional derivative of the Lipschitzian function $d_C(z) := \inf\{\|y - z\| \ : \ y \in C\}$ at x in direction v. The *Clarke normal cone* $N_C(x)$ to C at x is defined as the *dual cone* of $T_C(x)$, i.e.

$$N_C(x) = \{x^* \in R^n \ : \ \langle x^*, v\rangle \leq 0 \quad \text{for all } v \in T_C(x)\}.$$

Theorem 1.8. (See Clarke (1983), Propositions 2.4.3, 2.4.4 and 2.4.5) *For any nonempty subset $C \subset R^n$ and any point $x \in C$, the following assertions hold:*

(a) $N_C(x) = \overline{\left\{ \cup_{t \geq 0} \, t\partial d_C(x) \right\}}$.

(b) *If C is convex then $N_C(x)$ coincides with the normal cone to C at x defined by formula* (1.12), *and $T_C(x)$ coincides with the topological closure of the set* $\text{cone}(C - x) := \{tz \ : \ t \geq 0, \ z \in C - x\}$.

(c) *The inclusion $v \in T_C(x)$ is valid if and only if, for every sequence x_k in C converging to x and sequence t_k in $(0, +\infty)$ converging to 0, there exists a sequence v_k in R^n converging to v such that $x_k + t_k v_k \in C$ for all k.*

We now consider problem (P) under the assumptions that $f : R^n \to R$ is a locally Lipschitz function and

$$\Delta = \{x \in C : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, \ h_1(x) = 0, \ldots, h_s(x) = 0\}, \tag{1.25}$$

where $C \subset R^n$ is a nonempty subset, $g_i : R^n \to R$ $(i = 1, \ldots, m)$ and $h_j : R^n \to R$ $(j = 1, \ldots, s)$ are locally Lipschitz functions.

Theorem 1.9. (See Clarke (1983), Theorem 6.1.1 and Remark 6.1.2) *If $\bar{x}$ is a local solution of (P) then there exist $m + s + 1$ real numbers $\lambda_0 \geq 0, \ \lambda_1 \geq 0, \ldots, \lambda_m \geq 0, \ \mu_1, \ldots, \mu_s$, not all zero, such that*

$$0 \in \lambda_0 \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial g_i(\bar{x}) + \sum_{j=1}^{s} \mu_j \partial h_j(\bar{x}) + N_C(\bar{x}) \tag{1.26}$$

and

$$\lambda_i g_i(\bar{x}) = 0 \quad \textit{for all } \ i = 1, 2, \ldots, m. \tag{1.27}$$

The preceding theorem expresses the first-order necessary optimality condition for a class of nonsmooth programs in the *Fritz-John form.* Under some suitable *constraint qualifications*, the multiplier λ_0 corresponding to the objective function f is positive. In that case, dividing both sides of the inclusion in (1.26) and the equalities in (1.27) by λ_0 and setting $\widetilde{\lambda}_i = \lambda_i/\lambda_0$ for $i = 1, \ldots, m$, $\widetilde{\mu}_j = \mu_j/\lambda_0$ for $j = 1, \ldots, s$, we obtain

$$0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} \widetilde{\lambda}_i \partial g_i(\bar{x}) + \sum_{j=1}^{s} \widetilde{\mu}_j \partial h_j(\bar{x}) + N_C(\bar{x}) \tag{1.28}$$

and

$$\widetilde{\lambda}_i g_i(\bar{x}) = 0 \quad \text{for all } \ i = 1, 2, \ldots, m. \tag{1.29}$$

Similarly as in the case of convex programs (see Theorem 1.6), if (1.28) and (1.29) are fulfilled then the numbers $\widetilde{\lambda}_1 \geq 0, \ldots, \widetilde{\lambda}_m \geq 0$, $\widetilde{\mu}_1 \in R, \ldots, \widetilde{\mu}_s \in R$ are called the *Lagrange multipliers* corresponding to $\bar{x}$.

It is a simple matter to obtain the following two *Lagrange multiplier rules* from Theorem 1.9. (See Clarke (1983), pp. 234–236).

Corollary 1.1. *If $\bar{x}$ is a local solution of (P) and if the constraint qualification*

$$\begin{cases} \Big[0 \in \sum_{i=1}^{m} \lambda_i \partial g_i(\bar{x}) + \sum_{j=1}^{s} \mu_j \partial h_j(\bar{x}) + N_C(\bar{x}), \\ \lambda_1 \geq 0, \dots, \lambda_m \geq 0,\ \mu_1 \in R, \dots, \mu_s \in R; \\ \lambda_i g_i(\bar{x}) = 0 \ \text{for } i = 1, \dots, m\Big] \\ \Longrightarrow \Big[\lambda_1 = \dots = \lambda_m = 0,\ \mu_1 = \dots = \mu_s = 0\Big] \end{cases}$$

holds, then there exist Lagrange multipliers $\lambda_1 \geq 0, \dots, \lambda_m \geq 0$, $\mu_1 \in R, \dots, \mu_s \in R$ such that $\lambda_i g_i(\bar{x}) = 0$ for $i = 1, 2, \dots, m$, and

$$0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial g_i(\bar{x}) + \sum_{j=1}^{s} \mu_j \partial h_j(\bar{x}) + N_C(\bar{x}).$$

Corollary 1.2. *Assume that $\bar{x}$ is a local solution of a smooth program (P) where Δ is given by formula (1.23). If the following Mangasarian-Fromovitz constraint qualification*

$$\begin{cases} \text{The vectors } \{\nabla h_j(\bar{x}) : j = 1, \dots, s\} \text{ are linearly independent,} \\ \text{and there exists } v \in R^n \text{ such that } \langle \nabla h_j(\bar{x}), v \rangle = 0 \\ \text{for } j = 1, \dots, s, \text{ and } \langle \nabla g_i(\bar{x}), v \rangle < 0 \\ \text{for every } i = 1, \dots, m \text{ satisfying } g_i(\bar{x}) = 0 \end{cases}$$

is satisfied, then there exist Lagrange multipliers $\lambda_1 \geq 0, \dots, \lambda_m \geq 0$, $\mu_1 \in R, \dots, \mu_s \in R$ such that $\lambda_i g_i(\bar{x}) = 0$ for $i = 1, 2, \dots, m$, and

$$0 = \nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{s} \mu_j \nabla h_j(\bar{x}).$$

From Theorem 1.9 we can derive the basic Lagrange multiplier rule for convex programs stated in Theorem 1.6. Indeed, suppose that the assumptions of Theorem 1.6 are satisfied and $\bar{x}$ is a solution of (P). Consider separately the following two cases: (i) The vectors $\{a_j : j = 1, \dots, s\}$ are linearly independent; (ii) The vectors $\{a_j : j = 1, \dots, s\}$ are linearly dependent. In the first case, Theorem 1.9 shows that there exist real numbers $\lambda_0 \geq 0$, $\lambda_1 \geq 0, \dots, \lambda_m \geq 0$, $\mu_1, \dots, \mu_s$, not all zero, such that (1.26) and (1.27) are satisfied. Condition (1.24) forces $\lambda_0 > 0$. Hence there exist Lagrange multipliers satisfying the Kuhn-Tucker conditions. In the second case, when $a_j = 0$ for $j = 1, \dots, s$, we can obtain the desired result; when $a_j \neq 0$ for some $j = 1, \ldots, s$, we choose a maximal linearly independent subsystem, say $\{a_1, \ldots, a_k\}$, of the vector system $\{a_1, \ldots, a_s\}$. Then we consider the problem

$$\min\{f(x) \, : \, g_1(x) \leq 0, \ldots, g_m(x) \leq 0, h_1(x) = 0, \ldots, h_k(x) = 0\}. \tag{1.30}$$

It is easy to show that the constraint set of this problem coincides with Δ. Hence $\bar{x}$ is a solution of (1.30). Applying Theorem 1.9 to problem (1.30) and using condition (1.24) we can find a set of Lagrange multipliers satisfying the Kuhn-Tucker conditions.

1.4 Linear Programs and Nonlinear Programs

Definition 1.15. A subset $\Delta \subset R^n$ is called a *polyhedral convex set* if Δ can be represented as the intersection of finitely many closed half spaces of R^n; that is, there exist nonzero vectors $a_1, \ldots, a_m \in R^n$ and real numbers $\beta_1, \ldots, \beta_m$ such that

$$\Delta = \{x \in R^n : \langle a_i, x \rangle \geq \beta_i \text{ for } i = 1, \ldots, m\}. \tag{1.31}$$

In other words, Δ is the solution set of a system of finitely many linear inequalities. (We admit that the intersection any *empty family* of closed half spaces of R^n is R^n. Hence $\Delta = R^n$ is also a polyhedral convex set.) A point $x \in \Delta$ is called an *extreme point* of Δ if there is no way to express x in the form $x = ty + (1-t)z$ where $y \in \Delta$, $z \in \Delta$, $y \neq z$, and $t \in (0, 1)$. The set of all the extreme points of Δ is denoted by extrΔ.

Let A be the $m \times n$-matrix with the elements a_{ij} ($i = 1, \ldots, m$, $j = 1, \ldots, n$), where a_{ij} stands for the j-th component of a_i. Set $b = (\beta_1, \ldots, \beta_m) \in R^m$. Then (1.31) can be rewritten as

$$\Delta = \{x \in R^n \, : \, Ax \geq b\}.$$

Here and subsequently, for any two vectors $y = (y_1, \ldots, y_m) \in R^m$ and $z = (z_1, \ldots, z_m) \in R^m$, we write $y \geq z$ if $y_i \geq z_i$ for all $i = 1, \ldots, m$. We shall write $y > z$ if $y_i > z_i$ for all $i = 1, \ldots, m$. Since

$$\{x \in R^n \, : \, Ax = b\} = \{x \in R^n \, : \, Ax \geq b, \ (-A)x \geq -b\},$$

it follows that $\{x \in R^n \, : \, Ax = b\}$ is a polyhedral convex set.

Definition 1.16. Problem (P) is called a *linear program* (a linear programming problem) if f is an affine function and Δ is a polyhedral convex set. Otherwise, (P) is said to be a *nonlinear program.*

There are three typical forms for describing linear programs

$$\begin{aligned}
&\min\{f(x) = \langle c, x\rangle \ : \ x \in R^n, \ Ax \geq b\},\\
&\min\{f(x) = \langle c, x\rangle \ : \ x \in R^n, \ Ax = b, \quad x \geq 0\},\\
&\min\{f(x) = \langle c, x\rangle \ : \ x \in R^n, \ Ax \geq b, \ Cx = d\}
\end{aligned}$$

which are called the *standard form*, the *canonical form* and the *general form*, respectively. Here $A \in R^{m\times n}$, $C \in R^{s\times n}$ are given matrices, $c \in R^n$, $b \in R^m$ and $d \in R^s$ are given vectors.

Example 1.4. Consider the following linear program of the standard form:

$$\min\left\{x_1 + \frac{1}{2}x_2 \ : \ x = (x_1, x_2), \ x_1 + x_2 \geq 1, \ x_1 \geq 0, \ x_2 \geq 0\right\}.$$

It is easy to check that $\mathrm{Sol}(P) = \{(0,1)\}$. Note that the constraint set

$$\Delta = \{x \in R^2 \ : \ x_1 + x_2 \geq 1, \ \ x_1 \geq 0, \ \ x_2 \geq 0\}$$

has two extreme points, namely $\mathrm{extr}\Delta = \{(1,0), \ (0,1)\}$. One of these points is the solution of our problem.

Definition 1.17. The *dual problems* of linear programs of the standard, canonical and general forms, respectively, are the following linear programs:

$$\begin{aligned}
&\max\{\langle b, y\rangle \ : \ y \in R^m, \ A^Ty = c, \ y \geq 0\},\\
&\max\{\langle b, y\rangle \ : \ y \in R^m, \ A^Ty \leq c\},\\
&\max\{\langle b, y\rangle + \langle d, z\rangle \ : \ (y,z) \in R^m \times R^s, \ A^Ty + C^Tz = c, \ y \geq 0\}.
\end{aligned}$$

Of course, linear programs are convex mathematical programming problems. Hence they enjoy all the properties of the class of convex programs. Besides, linear programs have many other special properties.

Theorem 1.10. (See Dantzig (1963)) *Let (P) be a linear program in one of the three typical forms. The following properties hold true:*

(i) *If the constraint set is nonempty and if $v(P) > -\infty$, then $\mathrm{Sol}(P)$ is a nonempty polyhedral convex set.*

(ii) *If both the sets* extrΔ and Sol(P) *are nonempty, then the intersection* extr$\Delta \cap$ Sol(P) *is also nonempty.*

(iii) *If* rank$A = n$ *and the set* $\Delta := \{x \in R^n \, : \, Ax = b, \; x \geq 0\}$ *is nonempty, then* Δ *must have an extreme point.*

(iv) *The optimal value* $v(P)$ *of* (P) *and the optimal value* $v(P')$ *of the dual problem* (P') *of* (P) *are equal, provided that the constraint set of at least one of these problems is nonempty.*

Note that the five problems considered in Remarks 1.2, 1.4 and Examples 1.1-1.3 are all nonlinear.

We now consider one important class of nonlinear programs, which contains the class of linear programs as a special subclass.

1.5 Quadratic Programs

Definition 1.18. We say that $f : R^n \to R$ is a *linear-quadratic function* if there exist a matrix $D \in R^{n \times n}$, a vector $c \in R^n$ and a real number α such that

$$\begin{aligned} f(x) &= \frac{1}{2} x^T D x + c^T x + \alpha \\ &= \frac{1}{2} \langle x, Dx \rangle + \langle c, x \rangle + \alpha \end{aligned} \tag{1.32}$$

for all $x \in R^n$.

If

$$D = \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \dots & \dots & \dots \\ d_{n1} & \dots & d_{nn} \end{pmatrix}, \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

then (1.32) means that

$$f(x) = \frac{1}{2} \Big(\sum_{j=1}^{n} \sum_{i=1}^{n} d_{ij} x_i x_j \Big) + \sum_{i=1}^{n} c_i x_i + \alpha.$$

Since $x^T D x = \frac{1}{2} x^T (D + D^T) x$ for every $x \in R^n$, representation (1.32) remains valid if we replace D by the symmetric matrix $\frac{1}{2}(D + D^T)$. For this reason, we will assume that the square matrix in the representation of a linear-quadratic function is symmetric. The space of the symmetric $n \times n$-matrices will be denoted by $R_S^{n \times n}$.

Definition 1.19. Problem (P) is called a *linear-quadratic mathematical programming problem* (or a *quadratic program*, for brevity) if f is a linear-quadratic function and Δ is a polyhedral convex set.

In (1.32), if D is the zero matrix then f is an affine function. Thus the class of linear programs is a subclass of the class of quadratic programs. In general, quadratic programs are *nonconvex* mathematical programming problems.

Example 1.5. The following quadratic program is nonconvex:

$$\min\{f(x) = x_1^2 - x_2^2 \,:\, x = (x_1, x_2),\ 1 \le x_1 \le 3,\ 1 \le x_2 \le 3\}.$$

It is obvious that f is a nonconvex function. One can verify that $\mathrm{Sol}(P) = \{(1,3)\}$ and $v(P) = -8$.

It is clear that if we delete the constant α in the representation (1.32) of f then we do not change the solution set of the problem $\min\{f(x) \,:\, x \in \Delta\}$, where $\Delta \subset R^n$ is a polyhedral convex set. Therefore, instead of (1.32) we will usually use the simplified form $f(x) = \frac{1}{2}x^T Dx + c^T x$ of the objective function.

Modifying the terminology used for linear programs, we call the following forms of quadratic programs

$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ Ax \ge b\right\},$$
$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ Ax \ge b,\quad x \ge 0\right\},$$
$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ Ax \ge b,\ Cx = d\right\}$$

the *standard form*, the *canonical form* and the *general form*, respectively. (The meaning of A, C, b and d is the same as in the description of the typical forms of linear programs.) Note that the representation of the constraint set of canonical quadratic programs is slightly different from that of canonical linear programs. The above definition of canonical quadratic programs is adopted because quadratic programs of this type have a very tight connection with *linear complementarity problems* (see, for instance, Murty (1976) and Cottle et al. (1992)). In Chapter 5 we will clarify this point. The relation between the general quadratic programs and *affine variational inequalities* will be studied in the same chapter.

Definition 1.20. A matrix $D \in R^{n\times n}$ is said to be *positive definite* (resp., *negative definite*) if $v^T Dv > 0$ (resp., $v^T Dv < 0$) for every $v \in R^n \setminus \{0\}$. If $v^T Dv \geq 0$ (resp., $v^T Dv \leq 0$) for every $v \in R^n$ then D is said to be *positive semidefinite* (resp., *negative semidefinite*).

Proposition 1.2. *Let* $f(x) = \frac{1}{2}x^T Dx + c^T x + \alpha$ *where* $D \in R_S^{n\times n}$, $c \in R^n$ *and* $\alpha \in R$. *If* D *is a positive semidefinite matrix, then* f *is a convex function.*

Proof. Since $x \mapsto c^T x + \alpha$ is a convex function and the sum of two convex functions is a convex function, it suffices to show that $f_1(x) := x^T Dx$ is a convex function. As D is a positive semidefinite matrix, for every $u \in R^n$ and $v \in R^n$ we have

$$0 \leq (u-v)^T D(u-v) = u^T Du - 2v^T Du + v^T Dv.$$

This implies that

$$v^T Dv \leq u^T Du - 2v^T D(u-v). \tag{1.33}$$

Given any $x \in R^n$, $y \in R^n$ and $t \in (0,1)$, we set $z = tx + (1-t)y$. Taking account of (1.33) we have

$$\begin{aligned} z^T Dz &\leq y^T Dy - 2z^T D(y-z), \\ z^T Dz &\leq x^T Dx - 2z^T D(x-z). \end{aligned}$$

Since $y - z = t(y-x)$ and $x - z = (1-t)(x-y)$, from the last two inequalities we deduce that

$$(1-t)z^T Dz + tz^T Dz \leq (1-t)y^T Dy + tx^T Dx,$$

hence

$$f_1(tx + (1-t)y) = f_1(z) \leq tf_1(x) + (1-t)f(y).$$

Thus f_1 is a convex function. □

If D is negative semidefinite, then the function f given by (1.32) is *concave*, i.e.

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y)$$

for every $x \in R^n$, $y \in R^n$ and $t \in (0,1)$. In the case where matrix D is neither assumed to be positive semidefinite nor assumed to be negative semidefinite, we say that $f(x) = \frac{1}{2}x^T Dx + c^T x$, where $c \in R^n$, is an *indefinite* linear-quadratic function. Quadratic programming problems with indefinite linear-quadratic objective functions are called *indefinite* quadratic programs.

Remark 1.6. It is clear that if f is given by (1.32), where $D \in R_S^{n\times n}$, then $\nabla^2 f(x) = D$ for every $x \in R^n$. Therefore, the fact stated in Proposition 1.2 is a direct consequence of the following theorem (see Rockafellar (1970), Theorem 4.5): *"If $f : R^n \to R$ is a C^2-function and if the Hessian matrix $\nabla^2 f(x)$ is positive semidefinite for every $x \in R^n$, then f is a convex function."*

By using Proposition 1.2 one can verify whether a given quadratic program is convex or not.

Let us consider a simple example of convex quadratic programs.

Example 1.6. Given k points $a_1, a_2, \ldots, a_k$ in R^n, we want to find a point $x \in R^n$ at which the sum

$$f(x) := \|x - a_1\|^2 + \cdots + \|x - a_k\|^2$$

attains its minimal value. Observe that

$$\begin{aligned} f(x) &= \textstyle\sum_{i=1}^k (x - a_i)^T (x - a_i) \\ &= kx^T x - 2\left(\textstyle\sum_{i=1}^k a_i\right)^T x + \textstyle\sum_{i=1}^k a_i^T a_i \end{aligned}$$

is a convex linear-quadratic function. By Theorem 1.5, $\bar{x}$ is a solution of our problem if and only if $\nabla f(\bar{x}) = 0$. Since

$$\nabla f(\bar{x}) = 2k\bar{x} - 2\sum_{i=1}^k a_i,$$

one can write the condition $0 = \nabla f(\bar{x})$ equivalently as

$$\bar{x} = \frac{1}{k}\sum_{i=1}^k a_i.$$

Thus $\bar{x} = \frac{1}{k}\sum_{i=1}^k a_i$ is the unique solution of our problem. That special point $\bar{x}$ is called the *barycenter* of the system $\{a_1, a_2, \ldots, a_k\}$.

Observe that there is a simple algorithm for constructing the barycenter. Namely, first we define $z_1 = \frac{1}{2}a_1 + \frac{1}{2}a_2$. Then we put

$$z_i = \frac{i}{i+1}z_{i-1} + \frac{1}{i+1}a_{i+1} \quad \text{for every } \ i \geq 2.$$

By induction it is not difficult to show that $\bar{x} := z_{k-1}$ is the barycenter of the system $\{a_1, a_2, \ldots, a_k\}$. For performing a sequential construction of the barycenter of a system of points in R^2, it is convenient to use the following equivalent vector form of the formula defining z_i:

$$\overrightarrow{z_{i-1}z_i} = \frac{1}{i+1}\overrightarrow{z_{i-1}a_i} \qquad (\text{for every } i \geq 2).$$

The following geometrical example leads to a (nonconvex) quadratic program of the general form.

Example 1.7. Let $\Delta = \{x \in R^n : Ax \geq b, \ Cx = d\}$, where $A \in R^{m\times n}$, $C \in R^{s\times n}$, $b \in R^m$ and $d \in R^s$. (The equality $Cx = d$ can be absent in that formula. Likewise, the inequality $Ax \geq b$ can be absent too.) Let α_i $(i = 1, \ldots, n)$, $\widetilde{\alpha}_i$ $(i = 1, \ldots, n)$, β and $\widetilde{\beta}$ be a family of $2n + 2$ real numbers satisfying the conditions

$$\sum_{i=1}^{n} \alpha_i^2 = 1, \quad \sum_{i=1}^{n} \widetilde{\alpha}_i^2 = 1. \tag{1.34}$$

Note that

$$M = \{x \in R^n : \sum_{i=1}^{n} \alpha_i x_i + \beta = 0\}$$

and

$$\widetilde{M} = \{x \in R^n : \sum_{i=1}^{n} \widetilde{\alpha}_i x_i + \widetilde{\beta} = 0\}$$

are two *hyperplanes* in R^n. The task is to find $x \in \Delta$ such that the function

$$f(x) = (\text{dist}(x, M))^2 - (\text{dist}(x, \widetilde{M}))^2,$$

where $\text{dist}(x, \Omega) = \inf\{\|x - z\| : z \in \Omega\}$ is the *distance* from x to a subset $\Omega \subset R^n$, achieves its minimum. We have

$$\text{dist}(x, M) = |\sum_{i=1}^{n} \alpha_i x_i + \beta|. \tag{1.35}$$

In order to prove this formula we consider the following convex program

$$\min\{\varphi(z) = \|x - z\|^2 \,:\, z \in M\}. \tag{1.36}$$

By Theorem 1.6, $\bar{z} = (\bar{z}_1, \ldots, \bar{z}_n) \in M$ is a solution of (1.36) if and only if there exists $\mu \in R$ such that

$$0 \in \partial\varphi(\bar{z}) + \mu(\alpha_1, \ldots, \alpha_n).$$

Since $\partial\varphi(\bar{z}) = \{\nabla\varphi(\bar{z})\} = \{-2(x - \bar{z})\}$, this inclusion is valid if and only if

$$2(x - \bar{z}) = \mu(\alpha_1, \ldots, \alpha_n).$$

This implies $\bar{z} = x - \dfrac{\mu}{2}\alpha$, where $\alpha := (\alpha_1, \ldots \alpha_n)$. As $\bar{z} \in M$, we must have

$$0 = \langle \alpha, \bar{z} \rangle + \beta = \langle \alpha, x \rangle - \frac{\mu}{2}\langle \alpha, \alpha \rangle + \beta.$$

Taking account of (1.34), we obtain $\mu = 2(\langle \alpha, x \rangle + \beta)$. Therefore

$$\begin{aligned}(\mathrm{dist}(x, M))^2 = \|x - \bar{z}\|^2 &= \|x - (x - \frac{\mu}{2}\alpha)\|^2 \\ &= \left(\frac{\mu}{2}\right)^2 \langle \alpha, \alpha \rangle \\ &= (\langle \alpha, x \rangle + \beta)^2,\end{aligned}$$

hence (1.35) holds. Similarly,

$$\mathrm{dist}(x, \widetilde{M}) = |\sum_{i=1}^{n} \widetilde{\alpha}_i x_i + \widetilde{\beta}|.$$

Consequently,

$$\begin{aligned} f(x) &= \left(\sum_{i=1}^n \alpha_i x_i + \beta\right)^2 - \left(\sum_{i=1}^n \widetilde{\alpha}_i x_i + \widetilde{\beta}\right)^2 \\ &= \sum_{j=1}^n \sum_{i=1}^n (\alpha_i \alpha_j - \widetilde{\alpha}_i \widetilde{\alpha}_j) x_i x_j \\ &\qquad + 2\sum_{i=1}^n (\beta\alpha_i - \widetilde{\beta}\widetilde{\alpha}_i) x_i + (\beta^2 - \widetilde{\beta}^2). \end{aligned}$$

From this we conclude that $f(x)$ is a linear-quadratic function; so the optimization problem under consideration is a quadratic program of the general form.

It is easy to verify that if we choose

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ -3 \\ -3 \end{pmatrix},$$

$$\alpha = (1, 0), \quad \beta = 0, \quad \widetilde{\alpha} = (0, 1), \quad \widetilde{\beta} = 0,$$

then the preceding problem, where the equation $Cx = d$ is absent, becomes the one discussed in Example 1.5. In this case, we have $M = \{x = (x_1, x_2) : x_1 = 0,\ x_2 \in R\}$, $\widetilde{M} = \{x = (x_1, x_2) : x_1 \in R,\ x_2 = 0\}$, $\text{dist}(x, M) = |x_1|$ and $\text{dist}(x, \widetilde{M}) = |x_2|$.

1.6 Commentaries

Mathematical Programming is one important branch of Optimization Theory. Other branches with many interesting problems and results are known under the names The Calculus of Variations and Optimal Control Theory. Of course, the informal division of Optimization Theory into such branches is only for convenience. In fact, research problems and methods of the three branches are actively interacted. The classical work of Ioffe and Tihomirov (1979) is an excellent textbook addressing all the three branches of Optimization Theory.

Convex Analysis and Convex Programming theory can be studied by using the books of Rockafellar (1970) and of Ioffe and Tihomirov (1979). The book of Mangasarian (1969) gives a nice introductory course to Mathematical Programming.

Linear Programming can be learned by using the books of Dantzig (1963), Murty (1976), and many other nice books.

Nonsmooth Analysis and Nonsmooth Optimization can be learned by the books of Clarke (1983), Rockafellar and Wets (1998), and many other excellent books. In addition to these books, one can study Mordukhovich (1988, 1993, 1994) to be familiar with a powerful approach to Nonsmooth Analysis and Nonsmooth Optimization which has been developed intensively in recent years.

Some theoretical results on (Nonconvex) Quadratic Programming are available in the books of Murty (1976), Bank et al. (1982), Cottle et al. (1992), and other books. The next three chapters of this book are intended to cover the basic facts on (Nonconvex) Quadratic Programming, such as the solution existence, necessary and sufficient optimality conditions, and structure of the solution sets. Eight other chapters (Chapters 10-17) establish various results on stability and sensitivity of parametric quadratic programs.

The geometric problem described in Example 1.3 is called Fermat's problem or Steiner's problem. It was proposed by Fermat to

Torricelli. Torricelli's solution was published in 1659 by his pupil Viviani (see Weisstein (1999), p. 623).

Chapter 2

Existence Theorems for Quadratic Programs

In this chapter we shall discuss the Frank-Wolfe Theorem and the Eaves Theorem, which are two fundamental existence theorems for quadratic programming problems.

2.1 The Frank-Wolfe Theorem

Consider a quadratic program of the standard form

$$\begin{cases} \text{Minimize } f(x) := \frac{1}{2}x^T Dx + c^T x \\ \text{subject to } x \in R^n, \ Ax \geq b, \end{cases} \tag{2.1}$$

where $D \in R_S^{n\times n}$, $A \in R^{m\times n}$, $c \in R^n$ and $b \in R^m$. For the constraint set and the optimal value of (2.1) we shall use the following abbreviations:

$$\Delta(A,b) = \{x \in R^n \ : \ Ax \geq b\},$$
$$\bar{\theta} = \inf\{f(x) \ : \ x \in \Delta(A,b)\}.$$

If $\Delta(A,b) = \emptyset$ then $\bar{\theta} = +\infty$ by convention. If $\Delta(A,b) \neq \emptyset$ then there are two situations: (i) $\bar{\theta} \in R$, (ii) $\bar{\theta} = -\infty$. If (ii) occurs then, surely, (2.1) has no solutions. It is natural to ask: *Whether the problem always has solutions when* (i) *occurs?*

Note that optimization problems with non-quadratic objective functions may have no solutions even in the case the optimal value is finite. For example, the problem $\min\left\{\frac{1}{x} \ : \ x \in R, \ x \geq 1\right\}$ has no solutions, while the optimal value $\bar{\theta} = \inf\left\{\frac{1}{x} : x \in R,\ x \geq 1\right\} = 0$ is finite.

The following result was published by Frank and Wolfe in 1956.

Theorem 2.1. (The Frank-Wolfe Theorem; See Frank and Wolfe (1956), p. 108) *If* $\bar{\theta} = \inf\{f(x) : x \in \Delta(A,b)\}$ *is a finite real number then problem* (2.1) *has a solution.*

Proof. We shall follow the analytical proof proposed by Blum and Oettli (1972). The assumption $\bar{\theta} \in R$ implies that $\Delta(A,b) \neq \emptyset$. Select a point $x^0 \in \Delta(A,b)$. Let $\rho > 0$ be given arbitrarily. Define

$$\Delta_\rho = \Delta(A,b) \cap \bar{B}(x^0, \rho).$$

Note that Δ_ρ is a convex, nonempty, compact set. Consider the following problem

$$\min\{f(x) : x \in \Delta_\rho\}. \tag{2.2}$$

By the Weierstrass Theorem, there exists some $y \in \Delta_\rho$ such that $f(y) = q_\rho := \min\{f(x) : x \in \Delta_\rho\}$. Since the solution set of (2.2) is nonempty and compact, there exists $y_\rho \in \Delta_\rho$ such that

$$\|y_\rho - x^0\| = \min\{\|y - x^0\| : y \in \Delta_\rho,\ f(y) = q_\rho\}.$$

We claim that there exists $\hat{\rho} > 0$ such that

$$\|y_\rho - x^0\| < \rho \quad \text{for all } \rho \geq \hat{\rho}. \tag{2.3}$$

Indeed, if the claim were false then we would find an increasing sequence $\rho_k \to +\infty$ such that for every k there exists $y_{\rho_k} \in \Delta_{\rho_k}$ such that

$$f(y_{\rho_k}) = q_{\rho_k}, \quad \|y_{\rho_k} - x^0\| = \rho_k. \tag{2.4}$$

For simplicity of notation, we write y^k instead of y_{ρ_k}. Since $y^k \in \Delta(A,b)$, we must have $A_i y^k \geq b_i$ for $i = 1, \ldots, m$, where A_i denotes the i-th row of A and b_i denotes the i-th component of b. For $i = 1$, since the sequence $\{A_1 y^k\}$ is bounded below, one can choose a subsequence $\{k'\} \subset \{k\}$ such that $\lim_{k' \to \infty} A_1 y^{k'}$ exists. (It may happen that $\lim_{k' \to \infty} A_1 y^{k'} = +\infty$.) Without restriction of generality we can assume that $\{k'\} \equiv \{k\}$, that is the sequence $\{A_1 y^k\}$ itself is convergent. Similarly, for $i = 2$ there exists a subsequence $\{k'\} \subset \{k\}$ such that $\lim_{k' \to \infty} A_2 y^{k'}$ exists. Without loss of generality we can assume that $\{k'\} \equiv \{k\}$. Continue the process until $i = m$ to find a subsequence $\{k'\} \subset \{k\}$ such that all the limits

$$\lim_{k' \to \infty} A_i y^{k'} \quad (i = 1, \ldots, m)$$

exist. For simplicity of notation, we will assume that $\{k'\} \equiv \{k\}$. Let $I = \{1, \ldots, m\}$, $I_0 = \{i \in I : \lim_{k \to \infty} A_i y^k = b_i\}$ and

$$I_1 = I \setminus I_0 = \{i \in I : \lim_{k \to \infty} A_i y^k > b_i\}.$$

Of course, there exists $\varepsilon > 0$ such that

$$\lim_{k \to \infty} A_i y^k \geq b_i + \varepsilon \quad \text{for every } i \in I_1.$$

By (2.4), $\|(y^k - x^0)/\rho_k\| = 1$ for every k. Since the unit sphere in R^n is a compact set, there is no loss of generality in assuming that the sequence

$$\left\{ \frac{y^k - x^0}{\rho_k} \right\}$$

converges to some $\bar{v} \in R^n$ as $k \to \infty$. Clearly, $\|\bar{v}\| = 1$. As $\rho_k \to +\infty$, for every $i \in I_0$ we have

$$\begin{aligned} 0 &= \lim_{k \to \infty} (A_i y^k - b_i) \\ &= \lim_{k \to \infty} \left(\frac{A_i y^k - b_i}{\rho_k} \right) \\ &= \lim_{k \to \infty} \left(A_i \frac{y^k - x^0}{\rho_k} \right) + \lim_{k \to \infty} \left(\frac{A_i x^0 - b_i}{\rho_k} \right) = A_i \bar{v}. \end{aligned}$$

Similarly, for every $i \in I_1$ we have

$$\begin{aligned} 0 &\leq \liminf_{k \to \infty} \left(\frac{A_i y^k - b_i}{\rho_k} \right) \\ &= \liminf_{k \to \infty} \left(\frac{A_i y^k - A_i x^0}{\rho_k} + \frac{A_i x^0 - b_i}{\rho_k} \right) \\ &= \lim_{k \to \infty} \left(A_i \frac{y^k - x^0}{\rho_k} \right) + \lim_{k \to \infty} \left(\frac{A_i x^0 - b_i}{\rho_k} \right) = A_i \bar{v}. \end{aligned}$$

Therefore

$$A_i \bar{v} = 0 \quad \text{for every } i \in I_0, \qquad A_i \bar{v} \geq 0 \quad \text{for every } i \in I_1. \tag{2.5}$$

From this we can conclude that $\bar{v}$ is a direction of recession of the polyhedral convex set $\Delta(A, b)$. Recall (Rockafellar (1970), p. 61) that a nonzero vector $v \in R^n$ is said to be a *direction of recession* of a nonempty convex set $\Omega \subset R^n$ if

$$x + tv \in \Omega \quad \text{for every } t \geq 0 \text{ and } x \in \Omega.$$

Recall also that the set composed by $0 \in R^n$ and all the directions $v \in R^n$ satisfying the last condition, is called the *recession cone* of Ω. In our case, from (2.5) we deduce immediately that

$$y + t\bar{v} \in \Delta(A, b) \quad \text{for every } t \geq 0 \text{ and } y \in \Delta(A, b). \tag{2.6}$$

Since

$$\begin{aligned} f(y^k) = f(y_{\rho_k}) &= q_{\rho_k} \\ &= \min\{f(x) \,:\, x \in \Delta_{\rho_k}\} \\ &= \min\{f(x) \,:\, x \in \Delta(A,b) \cap \bar{B}(x_0, \rho_k)\} \end{aligned}$$

and the increasing sequence $\{\rho_k\}$ converges to $+\infty$, we see that the sequence $\{f(y^k)\}$ is non-increasing and $f(y^k) \to \bar{\theta}$. Consequently, for k sufficiently large, we have

$$\bar{\theta} - 1 \leq f(y^k) \leq \bar{\theta} + 1.$$

Using the formula of f we can rewrite these inequalities as follows

$$\begin{aligned} \bar{\theta} - 1 \ &\leq \frac{1}{2}(y^k - x^0)^T D(y^k - x^0) + c^T(y^k - x^0) \\ &\qquad + (x^0)^T D(y^k - x^0) + \frac{1}{2}(x^0)^T D x^0 + c^T x^0 \leq \bar{\theta} + 1. \end{aligned}$$

Dividing these expressions by ρ_k^2 and taking the limits as $k \to \infty$, we get $0 \leq \frac{1}{2}\bar{v}^T D\bar{v} \leq 0$. Hence

$$\bar{v}^T D\bar{v} = 0. \tag{2.7}$$

By (2.6),

$$y^k + t\bar{v} \in \Delta(A, b) \quad \text{for every } t \geq 0 \text{ and } k \in N,$$

where N stands for the set of the positive integers. On account of (2.7), we have

$$\begin{aligned} f(y^k + t\bar{v}) &= \frac{1}{2}(y^k + t\bar{v})^T D(y^k + t\bar{v}) + c^T(y^k + t\bar{v}) \\ &= \frac{1}{2}(y^k)^T D y^k + c^T y^k + t((y^k)^T D\bar{v} + c^T\bar{v}). \end{aligned}$$

Note that

$$(y^k)^T D\bar{v} + c^T\bar{v} \geq 0 \quad \text{for every } k \in N. \tag{2.8}$$

Indeed, if (2.8) were false then we would have $f(y^k + t\bar{v}) \to -\infty$ as $t \to +\infty$, which contradicts the assumption $\bar{\theta} \in R$.

Since $\langle \bar{v}, \bar{v} \rangle = 1$ and $\dfrac{y^k - x^0}{\rho_k} \to \bar{v}$, there exists $k_1 \in N$ such that $\left\langle \dfrac{y^k - x^0}{\rho_k}, \bar{v} \right\rangle > 0$ for all $k \geq k_1$. For any fixed index $k \geq k_1$, we have $\langle y^k - x^0, \bar{v} \rangle > 0$. Therefore

$$\|y^k - x^0 - t\bar{v}\|^2 = \|y^k - x^0\|^2 - 2t\langle y^k - x^0, \bar{v}\rangle + t^2\|\bar{v}\|^2 < \|y^k - x^0\|^2 \tag{2.9}$$

for $t > 0$ small enough. By (2.5),

$$A_i(y^k - t\bar{v}) = A_i y^k \geq b_i \quad \text{for all } i \in I_0.$$

Since $\lim_{k\to\infty} A_i y^k \geq b_i + \varepsilon$ for every $i \in I_1$, there exists $k_2 \in N$, $k_2 \geq k_1$, such that $A_i y^k \geq b_i + \dfrac{\varepsilon}{2}$ for every $k \geq k_2$ and $i \in I_1$. Fix an index $k \geq k_2$ and choose $\delta_k > 0$ as small as $tA_i\bar{v} \leq \dfrac{\varepsilon}{2}$ for every $i \in I_1$ and $t \in (0, \delta_k)$. (Of course, this choice is made only in the case $I_1 \neq \emptyset$.) Then we have

$$A_i(y^k - t\bar{v}) \geq b_i + \frac{\varepsilon}{2} - tA_i\bar{v} \geq b_i$$

for all $i \in I_1$ and $t \in (0, \delta_k)$. From what has already been proved, it may be concluded that

$$y^k - t\bar{v} \in \Delta(A, b) \quad \text{for all } t \in (0, \delta_k).$$

Combining this with (2.9) we see that $y^k - t\bar{v} \in \Delta(A, b)$ and

$$\|(y^k - t\bar{v}) - x^0\| = \|y^k - x^0 - t\bar{v}\| < \|y^k - x^0\| = \rho_k \tag{2.10}$$

for all $t \in (0, \delta_k)$ small enough. By (2.7) and (2.8), we have

$$f(y^k - t\bar{v}) = f(y^k) - t((y^k)^T D\bar{v} + c^T\bar{v}) \leq f(y^k).$$

So $y^k - t\bar{v}$ is a solution of the problem

$$\min\{f(x) : x \in \Delta_{\rho_k}\}. \tag{2.11}$$

From the inequality $\|(y^k - t\bar{v}) - x^0\| < \|y^k - x^0\|$ in (2.10) it follows that y^k cannot be a solution of (2.11) with the minimal distance to x^0, a contradiction.

We have shown that there exists $\hat{\rho} > 0$ such that (2.3) holds. We proceed to show that

$$\textit{there exists } \rho \geq \hat{\rho} \textit{ such that } q_\rho = \bar{\theta}. \tag{2.12}$$

As $q_\rho = \min\{f(x) : x \in \Delta_\rho\}$, it is easily seen that the conclusion of the theorem follows from (2.12). In order to obtain (2.12), we assume on the contrary that

$$q_\rho > \bar{\theta} \quad \text{for all } \rho \geq \hat{\rho}. \tag{2.13}$$

Note that $q_\rho \geq q_{\rho'}$ whenever $\rho' \geq \rho$. Note also that $q_\rho \to \bar{\theta}$ as $\rho \to +\infty$. Hence from (2.13) it follows that there exist $\rho_i \in (\hat{\rho}, +\infty)$ $(i = 1, 2)$ such that $\rho_1 < \rho_2$ and $q_{\rho_1} > q_{\rho_2}$. Since $\rho_2 > \hat{\rho}$, by (2.3) we have

$$\|y_{\rho_2} - x^0\| < \rho_2.$$

Since $q_{\rho_1} > q_{\rho_2}$, we must have $\rho_1 < \|y_{\rho_2} - x^0\|$. (Indeed, if $\rho_1 \geq \|y_{\rho_2} - x^0\|$ then $y_{\rho_2} \in \Delta_{\rho_1}$ and $f(y_{\rho_2}) = q_{\rho_2} < q_{\rho_1} = f(y_{\rho_1})$. This contradicts the choice of y_{ρ_1}.) Setting $\rho_3 = \|y_{\rho_2} - x^0\|$ we have $\rho_1 < \rho_3 < \rho_2$. Since $\rho_3 > \hat{\rho}$ and $\rho_2 > \hat{\rho}$, from (2.3) it follows that

$$\|y_{\rho_3} - x^0\| < \rho_3 = \|y_{\rho_2} - x^0\| < \rho_2. \tag{2.14}$$

Since $\rho_2 > \rho_3$, we have

$$q_{\rho_3} = f(y_{\rho_3}) \geq f(y_{\rho_2}) = q_{\rho_2}.$$

If $f(y_{\rho_3}) = f(y_{\rho_2})$ then from (2.14) we see that y_{ρ_3} is a feasible vector of the problem

$$\min\{f(x) : x \in \Delta_{\rho_2}\} \tag{2.15}$$

at which the objective function attains its optimal value $q_{\rho_2} = f(y_{\rho_2})$. Hence y_{ρ_3} is a solution of (2.15). By (2.14),

$$\|y_{\rho_3} - x^0\| < \|y_{\rho_2} - x^0\|.$$

This implies that y_{ρ_2} cannot be a solution of (2.15) with the minimal distance to x^0, a contradiction. So we must have $f(y_{\rho_3}) > f(y_{\rho_2})$. Since $\|y_{\rho_2} - x^0\| = \rho_3$, we deduce that y_{ρ_2} is a feasible vector of the problem $\min\{f(x) : x \in \Delta_{\rho_3}\}$. Then the inequality $f(y_{\rho_3}) > f(y_{\rho_2})$ contradicts the fact that y_{ρ_3} is a solution of this optimization problem. We have established property (2.12). The proof is complete. □

In Theorem 2.1, it is assumed that f is a linear-quadratic function and Δ is a polyhedral convex set. From Definition 1.15 it follows immediately that for any polyhedral convex set $\Delta \subset R^n$ there exists an integer $m \in N$, a matrix $A \in R^{m\times n}$ and a vector $b \in R^m$ such that $\Delta = \{x \in R^n : Ax \geq b\}$. This means that the Frank-Wolfe Theorem can be stated as follows: *"If a linear-quadratic function is bounded from below on a nonempty polyhedral convex set, then the problem of minimizing this function on the set must have a solution."*

If f is a linear-quadratic function but Δ is not assumed to be a polyhedral convex set, then the conclusion of Theorem 2.1 may not hold.

Example 2.1. Let $f(x) = x_1$ for every $x = (x_1, x_2) \in R^2$. Let $\Delta = \{x = (x_1, x_2) \in R^2 : x_1x_2 \geq 1,\ x_1 \geq 0,\ x_2 \geq 0\}$. We have $\bar{\theta} := \inf\{f(x) : x \in \Delta\} = 0$, but the problem $\min\{f(x) : x \in \Delta\}$ has no solutions.

If Δ is a polyhedral convex set but f is not assumed to be a linear-quadratic function, then the conclusion of Theorem 2.1 may not hold. In the following example, f is a polynomial function of degree 4 of the variables x_1 and x_2.

Example 2.2. (See Frank and Wolfe (1956), p. 109) Let $f(x) = x_1^2 + (1 - x_1x_2)^2$ for every $x = (x_1, x_2) \in R^2$. Let $\Delta = \{x = (x_1, x_2) \in R^2 : x_1 \geq 0,\ x_2 \geq 0\}$. Observe that $f(x) \geq 0$ for every $x \in R^2$. Choosing $x^k := \left(\frac{1}{k}, 1 + k\right)$, $k \in N$, we have

$$f(x^k) = \frac{2}{k^2} \to 0 \quad \text{as } k \to \infty.$$

This implies that

$$\bar{\theta} := \inf\{f(x) : x \in \Delta\} = 0 = \inf\{f(x) : x \in R^2\}.$$

It is a simple matter to show that both the problems $\min\{f(x) : x \in \Delta\}$ and $\min\{f(x) : x \in R^2\}$ have no solutions.

In Frank and Wolfe (1956), the authors informed that Irving Kaplansky has pointed out that the problem of minimizing a polynomial function of degree greater than 2 on a nonempty polyhedral convex set may not have solutions even in the case the function is bounded from below on the set.

Given a linear-quadratic function and a polyhedral convex set, verifying whether the function is bounded from below on the set is a rather difficult task. In the next section we will discuss another fundamental existence theorem for quadratic programming which gives us a tool for dealing with the task.

2.2 The Eaves Theorem

The following result was published by Eaves in 1971.

Theorem 2.2. (The Eaves Theorem; See Eaves (1971), Theorem 3 and Corollary 4, p. 702) *Problem* (2.1) *has solutions if and only if the following three conditions are satisfied:*

(i) $\Delta(A, b)$ *is nonempty;*

(ii) *If* $v \in R^n$ *and* $Av \geq 0$ *then* $v^T Dv \geq 0$*;*

(iii) *If* $v \in R^n$ *and* $x \in R^n$ *are such that* $Av \geq 0$, $v^T Dv = 0$ *and* $Ax \geq b$*, then* $(Dx + c)^T v \geq 0$.

Proof. *Necessity:* Suppose that (2.1) has a solution $\bar{x}$. Since $\bar{x} \in \Delta(A, b)$, condition (i) is satisfied. Given any $v \in R^n$ with $Av \geq 0$, since $A(\bar{x} + tv) = A\bar{x} + tAv \geq b$ for every $t \geq 0$, we have $\bar{x} + tv \in \Delta(A, b)$ for every $t \geq 0$. Hence $f(\bar{x} + tv) \geq f(\bar{x})$ for every $t \geq 0$. It follows that $\frac{1}{2}t^2 v^T Dv + t(D\bar{x} + c)^T v \geq 0$ for every $t \geq 0$, hence $v^T Dv \geq 0$. This shows that condition (ii) is satisfied. We now suppose that there are given any $v \in R^n$ and $x \in R^n$ with the properties that $Av \geq 0$, $v^T Dv = 0$ and $Ax \geq b$. Since $x + tv \in \Delta(A, b)$ for every $t \geq 0$ and $\bar{x}$ is a solution of (2.1), we have $f(x + tv) \geq f(\bar{x})$ for every $t \geq 0$. From this and the condition $v^T Dv = 0$ we deduce that $t(Dx + c)^T v + \frac{1}{2}x^T Dx + c^T x \geq f(\bar{x})$ for every $t \geq 0$. This implies that $(Dx + c)^T v \geq 0$. We have thus shown that condition (iii) is satisfied.

Sufficiency: Assume that conditions (i), (ii) and (iii) are satisfied. Define $\bar{\theta} = \inf\{f(x) : x \in R^n, \ Ax \geq b\}$. As $\Delta(A, b) \neq \emptyset$, we have $\bar{\theta} \neq +\infty$. If $\bar{\theta} \in R$ then the assertion of the theorem follows from the Frank-Wolfe Theorem. Hence we only need to show that the situation $\bar{\theta} = -\infty$ cannot occur. To obtain a contradiction, suppose that $\bar{\theta} = -\infty$. We can now proceed analogously to the proof of Theorem 2.1.

Fix a point $x^0 \in \Delta(A,b)$. For every $\rho > 0$, define $\Delta_\rho = \Delta(A,b) \cap \bar{B}(x^0, \rho)$ and consider the minimization problem $\min\{f(x) : x \in \Delta_\rho\}$. Denote by q_ρ the optimal value of this problem. Let $y_\rho \in \Delta_\rho$ be such that $f(y_\rho) = q_\rho$ and

$$\|y_\rho - x^0\| = \min\{\|y - x^0\| \, : \, y \in \Delta_\rho, \; f(y) = q_\rho\}.$$

We claim that there exists $\hat{\rho} > 0$ such that $\|y_\rho - x^0\| < \rho \quad$ for all $\rho \geq \hat{\rho}$. Suppose the claim were false. Then we would find an increasing sequence $\rho_k \to +\infty$ such that for every k there exists $y_{\rho_k} \in \Delta_{\rho_k}$ such that

$$f(y_{\rho_k}) = q_{\rho_k}, \quad \|y_{\rho_k} - x^0\| = \rho_k.$$

For simplicity of notation, we write y^k instead of y_{ρ_k}. Since $y^k \in \Delta(A,b)$, we must have $A_i y^k \geq b_i$ for $i = 1, \ldots, m$. Analysis similar to that in the proof of Theorem 2.1 shows that there exists a subsequence $\{k'\} \subset \{k\}$ such that all the limits

$$\lim_{k' \to \infty} A_i y^{k'} \quad (i = 1, \ldots, m)$$

exist. Without restriction of generality we can assume that $\{k'\} \equiv \{k\}$. Let $I = \{1, \ldots, m\}$, $I_0 = \{i \in I \, : \, \lim_{k\to\infty} A_i y^k = b_i\}$ and

$$I_1 = I \setminus I_0 = \{i \in I \, : \, \lim_{k \to \infty} A_i y^k > b_i\}.$$

Let $\varepsilon > 0$ be such that

$$\lim_{k \to \infty} A_i y^k \geq b_i + \varepsilon \quad \text{for every } i \in I_1.$$

Since $\|(y^k - x^0)/\rho_k\| = 1$ for every k, there is no loss of generality in assuming that the sequence

$$\left\{ \frac{y^k - x^0}{\rho_k} \right\}$$

converges to some $\bar{v} \in R^n$, $\|\bar{v}\| = 1$, as $k \to \infty$. Since $\rho_k \to +\infty$, for every $i \in I_0$ we have

$$\begin{aligned} 0 &= \lim_{k\to\infty} (A_i y^k - b_i) \\ &= \lim_{k\to\infty} \frac{A_i y^k - b_i}{\rho_k} \\ &= \lim_{k\to\infty} \left(A_i \frac{y^k - x^0}{\rho_k} \right) + \lim_{k\to\infty} \left(\frac{A_i x^0 - b_i}{\rho_k} \right) = A_i \bar{v}. \end{aligned}$$

Similarly, for every $i \in I_1$ we have

$$\begin{aligned} 0 &\le \liminf_{k\to\infty} \frac{A_i y^k - b_i}{\rho_k} \\ &= \liminf_{k\to\infty} \left(\frac{A_i y^k - A_i x^0}{\rho_k} + \frac{A_i x^0 - b_i}{\rho_k} \right) \\ &= \lim_{k\to\infty} \left(A_i \frac{y^k - x^0}{\rho_k} \right) + \lim_{k\to\infty} \frac{A_i x^0 - b_i}{\rho_k} = A_i \bar{v}. \end{aligned}$$

Therefore

$$A_i\bar{v} = 0 \quad \text{for every } i \in I_0, \qquad A_i\bar{v} \ge 0 \quad \text{for every } i \in I_1.$$

From this we deduce that

$$y + t\bar{v} \in \Delta(A,b) \quad \text{for every } t \ge 0 \text{ and } y \in \Delta(A,b). \tag{2.16}$$

Since

$$\begin{aligned} f(y^k) = f(y_{\rho_k}) &= q_{\rho_k} \\ &= \min\{f(x) \,:\, x \in \Delta_{\rho_k}\} \\ &= \min\{f(x) \,:\, x \in \Delta(A,b) \cap \bar{B}(x_0, \rho_k)\} \end{aligned}$$

and the increasing sequence $\{\rho_k\}$ converges to $+\infty$, we see that the sequence $\{f(y^k)\}$ is non-increasing and $f(y^k) \to \bar{\theta} = -\infty$. Hence $f(y^k) < 0$ for all k sufficiently large. Using the formula of f we can rewrite the last inequality as follows

$$\begin{aligned} &\frac{1}{2}(y^k - x^0)^T D(y^k - x^0) + c^T(y^k - x^0) \\ &\qquad + (x^0)^T D(y^k - x^0) + \frac{1}{2}(x^0)^T D x^0 + c^T x^0 < 0. \end{aligned}$$

Dividing this inequality by ρ_k^2 and taking the limits as $k \to \infty$, we get $\bar{v}^T D\bar{v} \le 0$. Since $A\bar{v} \ge 0$, from condition (ii) it follows that $\bar{v}^T D\bar{v} \ge 0$. Hence

$$\bar{v}^T D\bar{v} = 0. \tag{2.17}$$

By (2.16),

$$y^k + t\bar{v} \in \Delta(A,b) \quad \text{for every } t \ge 0 \text{ and } k \in N.$$

By virtue of (2.17), we have

$$\begin{aligned} f(y^k + t\bar{v}) &= \frac{1}{2}(y^k + t\bar{v})^T D(y^k + t\bar{v}) + c^T(y^k + t\bar{v}) \\ &= \frac{1}{2}(y^k)^T D y^k + c^T y^k + t((y^k)^T D\bar{v} + c^T\bar{v}). \end{aligned}$$

Since $y^k \in \Delta(A,b)$, $A\bar{v} \geq 0$ and $\bar{v}^T D\bar{v} = 0$, by condition (iii) we have

$$(y^k)^T D\bar{v} + c^T\bar{v} = (Dy^k + c)^T\bar{v} \geq 0 \quad \text{for every } k \in N. \tag{2.18}$$

Since $\langle \bar{v}, \bar{v}\rangle = 1$ and $\dfrac{y^k - x^0}{\rho_k} \to \bar{v}$, there exists $k_1 \in N$ such that $\left\langle \dfrac{y^k - x^0}{\rho_k}, \bar{v} \right\rangle > 0$ for all $k \geq k_1$. For any fixed index $k \geq k_1$, we have $\langle y^k - x^0, \bar{v}\rangle > 0$. Therefore

$$\|y^k - x^0 - t\bar{v}\|^2 = \|y^k - x^0\|^2 - 2t\langle y^k - x^0, \bar{v}\rangle + t^2\|\bar{v}\|^2 < \|y^k - x^0\|^2 \tag{2.19}$$

for $t > 0$ small enough. We have

$$A_i(y^k - t\bar{v}) = A_i y^k \geq b_i \quad \text{for all } i \in I_0.$$

Since $\lim_{k\to\infty} A_i y^k \geq b_i + \varepsilon$ for every $i \in I_1$, there exists $k_2 \in N$, $k_2 \geq k_1$, such that $A_i y^k \geq b_i + \dfrac{\varepsilon}{2}$ for every $k \geq k_2$ and $i \in I_1$. Fix an index $k \geq k_2$ and choose $\delta_k > 0$ as small as $tA_i\bar{v} \leq \dfrac{\varepsilon}{2}$ for every $i \in I_1$ and $t \in (0, \delta_k)$. Then we have

$$A_i(y^k - t\bar{v}) \geq b_i + \frac{\varepsilon}{2} - tA_i\bar{v} \geq b_i$$

for all $i \in I_1$ and $t \in (0, \delta_k)$. From what has already been proved, we deduce that

$$y^k - t\bar{v} \in \Delta(A,b) \quad \text{for all } t \in (0, \delta_k).$$

Combining this with (2.19) we see that $y^k - t\bar{v} \in \Delta(A,b)$ and

$$\|(y^k - t\bar{v}) - x^0\| = \|y^k - x^0 - t\bar{v}\| < \|y^k - x^0\| = \rho_k \tag{2.20}$$

for all $t \in (0, \delta_k)$ small enough. By (2.17) and (2.18), we have

$$f(y^k - t\bar{v}) = f(y^k) - t((y^k)^T D\bar{v} + c^T\bar{v}) \leq f(y^k).$$

So $y^k - t\bar{v}$ is a solution of the problem

$$\min\{f(x) \,:\, x \in \Delta_{\rho_k}\}. \tag{2.21}$$

From the inequality $\|(y^k - t\bar{v}) - x^0\| < \|y^k - x^0\|$ in (2.20) it follows that y^k cannot be a solution of (2.21) with the minimal distance to x^0, a contradiction.

We have shown that there exists $\hat{\rho} > 0$ such that

$$\|y_\rho - x^0\| < \rho \quad \text{for all } \rho \geq \hat{\rho}. \tag{2.22}$$

Note that $q_\rho \geq q_{\rho'}$ whenever $\rho' \geq \rho$. Note also that $q_\rho \to \bar{\theta} = -\infty$ as $\rho \to +\infty$. Hence there must exist $\rho_i \in (\hat{\rho}, +\infty)$ $(i = 1, 2)$ such that $\rho_1 < \rho_2$ and $q_{\rho_1} > q_{\rho_2}$. Since $\rho_2 > \hat{\rho}$, by (2.22) we have

$$\|y_{\rho_2} - x^0\| < \rho_2.$$

Since $q_{\rho_1} > q_{\rho_2}$, we must have $\rho_1 < \|y_{\rho_2} - x^0\|$. (Indeed, if $\rho_1 \geq \|y_{\rho_2} - x^0\|$ then $y_{\rho_2} \in \Delta_{\rho_1}$ and $f(y_{\rho_2}) = q_{\rho_2} < q_{\rho_1} = f(y_{\rho_1})$. This contradicts the choice of y_{ρ_1}.) Setting $\rho_3 = \|y_{\rho_2} - x^0\|$ we have $\rho_1 < \rho_3 < \rho_2$. Since $\rho_3 > \hat{\rho}$ and $\rho_2 > \hat{\rho}$, from (2.22) it follows that

$$\|y_{\rho_3} - x^0\| < \rho_3 = \|y_{\rho_2} - x^0\| < \rho_2. \tag{2.23}$$

Since $\rho_2 > \rho_3$, we have

$$q_{\rho_3} = f(y_{\rho_3}) \geq f(y_{\rho_2}) = q_{\rho_2}.$$

If $f(y_{\rho_3}) = f(y_{\rho_2})$ then from (2.23) we see that y_{ρ_3} is a feasible vector of the problem

$$\min\{f(x) \,:\, x \in \Delta_{\rho_2}\} \tag{2.24}$$

at which the objective function attains its optimal value $q_{\rho_2} = f(y_{\rho_2})$. Hence y_{ρ_3} is a solution of (2.24). By (2.23),

$$\|y_{\rho_3} - x^0\| < \|y_{\rho_2} - x^0\|.$$

This implies that y_{ρ_2} cannot be a solution of (2.24) with the minimal distance to x^0, a contradiction. So we must have $f(y_{\rho_3}) > f(y_{\rho_2})$. Since $\|y_{\rho_2} - x^0\| = \rho_3$, we deduce that y_{ρ_2} is a feasible vector of the problem $\min\{f(x) \,:\, x \in \Delta_{\rho_3}\}$. Then the inequality $f(y_{\rho_3}) > f(y_{\rho_2})$ shows that y_{ρ_3} cannot be a solution of this optimization problem, a contradiction. The proof is complete. □

Here are several important consequences of the Eaves Theorem.

Corollary 2.1. *Assume that D is a positive semidefinite matrix. Then problem* (2.1) *has solutions if and only if $\Delta(A, b)$ is nonempty and the following condition is satisfied:*

$$\big(v \in R^n,\, x \in R^n,\, Av \geq 0,\, v^T D v = 0,\, Ax \geq b\big) \Rightarrow (Dx + c)^T v \geq 0. \tag{2.25}$$

Proof. Note that condition (ii) in Theorem 2.2 is satisfied because, by our assumption, $v^T Dv \geq 0$ for every $v \in R^n$. Therefore the conclusion follows from Theorem 2.2. $\square$

Corollary 2.2. *Assume that D is a negative semidefinite matrix. Then problem* (2.1) *has solutions if and only if $\Delta(A, b)$ is nonempty and the following conditions are satisfied:*

(i) $(v \in R^n,\ Av \geq 0) \Rightarrow v^T Dv = 0$;

(ii) $(v \in R^n,\ x \in R^n,\ Av \geq 0,\ Ax \geq b) \Rightarrow (Dx + c)^T v \geq 0$.

Proof. Since $v^T Dv \leq 0$ for every $v \in R^n$ by our assumption, we see that condition (ii) in Theorem 2.2 is can be rewritten as condition (i) in this corollary. Besides, since $Av \geq 0$ implies $v^T Dv = 0$, condition (iii) in Theorem 2.2 can be rewritten as condition (ii) in this corollary. Therefore the conclusion follows from Theorem 2.2. $\square$

Corollary 2.3. *If D is a positive definite matrix, then problem* (2.1) *has solutions if and only if $\Delta(A, b)$ is nonempty.*

Proof. Since D is a positive definite matrix, the equality $v^T Dv = 0$ implies that $v = 0$. Then the assertion follows from Corollary 2.1 because condition (2.25) is satisfied. $\square$

Corollary 2.4. *If D is a negative definite matrix, then problem* (2.1) *has solutions if and only if $\Delta(A, b)$ is nonempty and compact.*

Proof. Since D is a negative definite matrix, conditions (i) and (ii) in Corollary 2.2 are fulfilled if and only if the set $L := \{v \in R^n : Av \geq 0\}$ contains just one element $v = 0$. Since L is the recession cone of the polyhedral convex set $\Delta(A, b) = \{x \in R^n : Ax \geq b\}$ (see Rockafellar (1970), p. 62), the condition $L = \{0\}$ is equivalent to the compactness of $\Delta(A, b)$ (see Rockafellar (1970), Theorem 8.4). Hence the assertion follows from Corollary 2.2. $\square$

Theorem 2.2 allows one to verify the existence of solutions of a quadratic program of the form (2.1) through analyzing its data set $\{D,\ A,\ c,\ b\}$. If any one from the three conditions (i), (ii) and (iii) in the theorem is violated, then the problem cannot have solutions.

Formally, the Eaves Theorem formulated above allows one to deal only with quadratic programs of the standard form. It is a simple matter to derive existence results for quadratic programs of the canonical and the general forms from Theorem 2.2.

Corollary 2.5. *Let $D \in R_S^{n\times n}$, $A \in R^{m\times n}$, $c \in R^n$ and $b \in R^m$. The quadratic program*

$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ Ax \geq b,\ x \geq 0\right\} \tag{2.26}$$

has solutions if and only if the following three conditions are satisfied:

(i) *The constraint set $\{x \in R^n : Ax \geq b,\ x \geq 0\}$ is nonempty;*

(ii) *If $v \in R^n$, $Av \geq 0$ and $v \geq 0$, then $v^T Dv \geq 0$;*

(iii) *If $v \in R^n$ and $x \in R^n$ are such that $Av \geq 0$, $v \geq 0$, $v^T Dv = 0$, $Ax \geq b$ and $x \geq 0$, then $(Dx + c)^T v \geq 0$.*

Proof. Define $\widetilde{A} \in R^{(m+n)\times n}$ and $\widetilde{b} \in R^{m+n}$ by setting

$$\widetilde{A} = \begin{pmatrix} A \\ E \end{pmatrix}, \quad \widetilde{b} = \begin{pmatrix} b \\ 0 \end{pmatrix},$$

where E denotes the unit matrix in $R^{n\times n}$ and 0 stands for the zero vector in R^n. It is clear that (2.26) can be rewritten in the following form

$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ \widetilde{A}x \geq \widetilde{b}\right\}.$$

Applying Theorem 2.2 to this quadratic program we obtain the desired result. □

Corollary 2.6. *Let $D \in R_S^{n\times n}$, $A \in R^{m\times n}$, $C \in R^{s\times n}$, $c \in R^n$, $b \in R^m$ and $d \in R^s$. The quadratic program*

$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ Ax \geq b,\ Cx = d\right\} \tag{2.27}$$

has solutions if and only if the following three conditions are satisfied:

(i) *The constraint set $\{x \in R^n : Ax \geq b,\ Cx = d\}$ is nonempty;*

(ii) *If $v \in R^n$, $Av \geq 0$ and $Cv = 0$, then $v^T Dv \geq 0$;*

(iii) *If $v \in R^n$ and $x \in R^n$ are such that $Av \geq 0$, $Cv = 0$, $v^T Dv = 0$, $Ax \geq b$ and $Cx = d$, then $(Dx + c)^T v \geq 0$.*

Proof. Define $\widetilde{A} \in R^{(m+2s)\times n}$ and $\widetilde{b} \in R^{m+2s}$ by setting

$$\widetilde{A} = \begin{pmatrix} A \\ C \\ -C \end{pmatrix}, \quad \widetilde{b} = \begin{pmatrix} b \\ d \\ -d \end{pmatrix}.$$

It is clear that (2.27) can be rewritten in following form

$$\min \left\{ \frac{1}{2} x^T D x + c^T x \; : \; x \in R^n, \; \widetilde{A}x \geq \widetilde{b} \right\}.$$

Applying Theorem 2.2 to this quadratic program we obtain the desired result. □

2.3 Commentaries

The Frank-Wolfe Theorem has various applications. For example, in (Cottle et al. (1992), Chapter 3) it has been used as the main tool for obtaining many existence results for linear complementarity problems. Usually, existence theorems for optimization problems and for variational problems give only *sufficient conditions* to assure that the problem under consideration has solutions. Here we see that the Frank-Wolfe Theorem and the Eaves Theorem provide some criteria (both the necessary and sufficient conditions) for the solution existence. This is possible because the quadratic programs have a relatively simple structure.

We realized the importance of the Eaves Theorem when we studied a paper by Klatte (1985) and applied the theorem to investigate the lower semicontinuity of the solution map in parametric quadratic programs (see Chapter 15 of this book), the directional differentiability and the *piecewise linear-quadratic property* of the optimal value function in quadratic programs under linear perturbations (see Chapter 16 of this book). We believe that the theorem is really very useful and important.

The proof of the Frank-Wolfe Theorem given in Section 2.1 follows exactly the scheme proposed in Blum and Oettli (1972). For the convenience of the reader, all the arguments of Blum and Oettli are described in detail. Two other proofs of the theorem can be found in Frank and Wolfe (1956) and Eaves (1971).

The proof of the Eaves Theorem given in Section 2.2 is rather different from the original proof given in Eaves (1971). The repetition of one part of the arguments used for proving the Frank-Wolfe

Theorem is intended to show the close interrelations between the two existence theorems.

Chapter 3

Necessary and Sufficient Optimality Conditions for Quadratic Programs

This chapter is devoted to a discussion on first-order optimality conditions and second-order optimality conditions for quadratic programming problems.

3.1 First-Order Optimality Conditions

In this section we will establish first-order necessary and sufficient optimality conditions for quadratic programs. Second-order necessary and sufficient optimality conditions for these problems will be obtained in the next section.

The first assertion of the following proposition states the *Fermat rule*, which is a basic first-order necessary optimality condition for mathematical programming problems, for quadratic programs. The second assertion states the so-called *first-order sufficient optimality condition* for quadratic programs and its consequence.

Theorem 3.1. *Let $\bar{x}$ be a feasible vector of the quadratic program*

$$\min\left\{f(x) = \frac{1}{2}x^T Dx + c^T x \,:\, x \in \Delta\right\}, \tag{3.1}$$

where $D \in R_S^{n\times n}$, $c \in R^n$, and $\Delta \subset R^n$ is a polyhedral convex set.

(i) *If $\bar{x}$ is a local solution of this problem, then*

$$\langle D\bar{x} + c, x - \bar{x}\rangle \geq 0 \quad \text{for every } \ x \in \Delta. \tag{3.2}$$

(ii) *If*

$$\langle D\bar{x} + c, x - \bar{x}\rangle > 0 \quad \text{for every } x \in \Delta \setminus \{\bar{x}\}, \tag{3.3}$$

then $\bar{x}$ *is a local solution of* (3.1) *and, moreover, there exist* $\varepsilon > 0$ *and* $\varrho > 0$ *such that*

$$f(x) - f(\bar{x}) \geq \varrho\|x - \bar{x}\| \quad \text{for every } x \in \Delta \cap B(\bar{x}, \varepsilon). \tag{3.4}$$

Proof. (i) Let $\bar{x} \in \Delta$ be a local solution of (3.1). Choose $\mu > 0$ so that

$$f(y) \geq f(\bar{x}), \quad \forall y \in \Delta \cap B(\bar{x}, \mu).$$

Given any $x \in \Delta \setminus \{\bar{x}\}$, we observe that there exists $\delta > 0$ such that $\bar{x} + t(x - \bar{x})$ belongs to $\Delta \cap B(\bar{x}, \mu)$ whenever $t \in (0, \delta)$. Therefore

$$\begin{aligned} 0 \leq \lim_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} &= f'(\bar{x}; x - \bar{x}) \\ &= \langle \nabla f(\bar{x}), x - \bar{x}\rangle \\ &= \langle D\bar{x} + c, x - \bar{x}\rangle. \end{aligned}$$

Property (3.2) has been established.

(ii) It suffices to show that if (3.3) holds then there exist $\varepsilon > 0$ and $\varrho > 0$ such that (3.4) is satisfied. To obtain a contradiction, suppose that (3.3) holds but for every $\varepsilon > 0$ and $\varrho > 0$ there exists $x \in \Delta \cap B(\bar{x}, \varepsilon)$ such that $f(x) - f(\bar{x}) < \varrho\|x - \bar{x}\|$. Then there exists a sequence $\{x^k\}$ in R^n such that, for every $k \in N$, we have $x^k \in \Delta \cap B(\bar{x}, \frac{1}{k})$ and $f(x^k) - f(\bar{x}) < \frac{1}{k}\|x^k - \bar{x}\|$. There is no loss of generality in assuming that the sequence of unit vectors $\{(x^k - \bar{x})/\|x^k - \bar{x}\|\}$ converges to some unit vector $\bar{v} \in R^n$. Since

$$\begin{aligned} f(x^k) - f(\bar{x}) &= \frac{1}{2}(x^k - \bar{x})^T D(x^k - \bar{x}) + c^T(x^k - \bar{x}) \\ &\quad + (x^k - \bar{x})^T D\bar{x} < \frac{1}{k}\|x^k - \bar{x}\|. \end{aligned}$$

Dividing the last inequality by $\|x^k - \bar{x}\|$ and taking the limits as $k \to \infty$, we obtain

$$(D\bar{x} + c)^T \bar{v} \leq 0. \tag{3.5}$$

Since Δ is a polyhedral convex set, there exist $m \in N$, $a_1, \ldots, a_m$ in R^n and $\beta_1, \ldots, \beta_m$ in R such that Δ has the representation (1.31). Let $I_0 = \{i : \langle a_i, \bar{x}\rangle = \beta_i\}$, $I_1 = \{i : \langle a_i, \bar{x}\rangle > \beta_i\}$. For each $i \in I_0$, we have $\langle a_i, x^k - \bar{x}\rangle = \langle a_i, x^k\rangle - \beta_i \geq 0$. Therefore $\langle a_i, \bar{v}\rangle = \lim_{k \to \infty} \langle a_i, (x^k - \bar{x})/\|x^k - \bar{x}\|\rangle \geq 0$. Obviously, there exists $\delta_1 > 0$ such that $\langle a_i, \bar{x} + t\bar{v} \rangle > \beta_i$ for every $i \in I_1$ and $t \in (0, \delta_1)$. Consequently, $\bar{x} + t\bar{v} \in \Delta$ for every $t \in (0, \delta_1)$. Substituting $x = \bar{x} + t\bar{v}$, where t is a value from $(0, \delta_1)$, into (3.3) gives $(D\bar{x} + c)^T\bar{v} > 0$, which contradicts (3.5). The proof is complete. □

For obtaining a first-order necessary optimality condition for quadratic programs in the form of a *Lagrange multiplier rule*, we shall need the following basic result concerning linear inequalities.

Theorem 3.2. (Farkas' Lemma; See Rockafellar (1970), p. 200) *Let a_0, $a_1, \ldots,$ a_k be vectors from R^n. The inequality $\langle a_0, x \rangle \le 0$ is a consequence of the system*

$$\langle a_i, x \rangle \le 0, \quad i = 1, 2, \ldots, k,$$

if and only if there exist nonnegative real numbers $\lambda_1, \ldots, \lambda_k$ such that

$$\sum_{i=1}^{k} \lambda_i a_i = a_0.$$

Theorem 3.3. (See, for instance, Cottle et al. (1992), p. 118) *If $\bar{x} \in R^n$ is a local solution of problem* (2.1) *then there exists $\lambda = (\lambda_1, \ldots, \lambda_m) \in R^m$ such that*

$$\begin{cases} D\bar{x} - A^T\lambda + c = 0, \\ A\bar{x} - b \ge 0, \quad \lambda \ge 0, \\ \lambda^T(A\bar{x} - b) = 0. \end{cases} \tag{3.6}$$

Proof. Denote by A_i the i-th row of A, and set $a_i = A_i^T$. Denote by b_i the i-th component of vector b. Set $\Delta = \Delta(A, b) = \{x \in R^n : Ax \ge b\}$. Let $\bar{x}$ be a local solution of (2.1). By Theorem 3.1(i), property (3.2) holds. Set $I = \{1, \ldots, m\}$, $I_0 = \{i \in I : \langle a_i, \bar{x} \rangle = b_i\}$ and $I_1 = \{i \in I : \langle a_i, \bar{x} \rangle > b_i\}$. For any $v \in R^n$ satisfying

$$\langle a_i, v \rangle \ge 0 \quad \text{for every } i \in I_0,$$

analysis similar to that in the proof of Theorem 3.1(ii) shows that there exists $\delta_1 > 0$ such that $\langle a_i, \bar{x} + tv \rangle \ge b_i$ for every $i \in I$ and $t \in (0, \delta_1)$. Substituting $x = \bar{x} + tv$, where t is a value from $(0, \delta_1)$, to (3.2) yields $\langle D\bar{x} + c, v \rangle \ge 0$. We have thus shown that

$$\langle -D\bar{x} - c, v \rangle \le 0$$

for any $v \in R^n$ satisfying

$$\langle -a_i, v\rangle \leq 0 \quad \text{for every } i \in I_0.$$

By Theorem 3.2, there exist nonnegative real numbers λ_i $(i \in I_0)$ such that

$$\sum_{i\in I_0} \lambda_i(-a_i) = -D\bar{x} - c. \tag{3.7}$$

Put $\lambda_i = 0$ for all $i \in I_1$ and $\lambda = (\lambda_1, \dots, \lambda_m)$. Since $a_i = A_i^T$ for every $i \in I$, from (3.7) we obtain the first equality in (3.6). Since $\bar{x} \in \Delta(A, b)$ and $\lambda_i(A_i\bar{x} - b_i) = 0$ for each $i \in I$, the other conditions in (3.6) are satisfied too. The proof is complete. □

From Theorem 3.3 we can derive the following Lagrange multiplier rules for quadratic programs of the canonical and the general forms.

Corollary 3.1. (See, for instance, Murty (1972)) *If $\bar{x}$ is a local solution of problem* (2.26), *then there exist $\lambda = (\lambda_1, \dots, \lambda_m) \in R^m$ such that*

$$\begin{cases} D\bar{x} - A^T\lambda + c \geq 0, \\ A\bar{x} - b \geq 0, \quad \bar{x} \geq 0, \quad \lambda \geq 0, \\ \bar{x}^T(D\bar{x} - A^T\lambda + c) + \lambda^T(A\bar{x} - b) = 0. \end{cases} \tag{3.8}$$

Proof. Define matrix $\tilde{A} \in R^{(m+n)\times n}$ and vector $\tilde{b} \in R^{m+n}$ as in the proof of Corollary 2.5 and note that problem (2.26) can be rewritten in the form

$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ \tilde{A}x \geq \tilde{b}\right\}.$$

Applying Theorem 3.3 to this quadratic program we deduce that there exists $\tilde{\lambda} = (\lambda_1, \dots, \lambda_{m+n}) \in R^{m+n}$ such that

$$D\bar{x} - \tilde{A}^T\tilde{\lambda} + c = 0, \quad \tilde{A}\bar{x} \geq \tilde{b}, \quad \tilde{\lambda} \geq 0, \quad \tilde{\lambda}^T(\tilde{A}\bar{x} - \tilde{b}) = 0.$$

Taking $\lambda = (\lambda_1, \dots, \lambda_m)$, we can obtain the properties stated in (3.8) from the last ones. □

Corollary 3.2. *If $\bar{x}$ is a local solution of problem* (2.27), *then there exist $\lambda = (\lambda_1, \dots, \lambda_m) \in R^m$ and $\mu = (\mu_1, \dots, \mu_s) \in R^s$ such that*

$$\begin{cases} D\bar{x} - A^T\lambda - C^T\mu + c = 0, \\ A\bar{x} - b \geq 0, \quad C\bar{x} = d, \quad \lambda \geq 0, \\ \lambda^T(A\bar{x} - b) = 0. \end{cases} \tag{3.9}$$

Proof. Define $\tilde{A} \in R^{(m+2s)\times n}$ and $\tilde{b} \in R^{m+2s}$ as in the proof of Corollary 2.6 and note that problem (2.27) can be rewritten in the form

$$\min\left\{\frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ \tilde{A}x \geq \tilde{b}\right\}.$$

Applying Theorem 3.3 to this quadratic program we see that there exists $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_{m+2n}) \in R^{m+2n}$ such that

$$D\bar{x} - \tilde{A}^T\tilde{\lambda} + c = 0, \quad \tilde{A}\bar{x} \geq \tilde{b}, \quad \tilde{\lambda} \geq 0, \quad \tilde{\lambda}^T(\tilde{A}\bar{x} - \tilde{b}) = 0.$$

Taking $\lambda = (\lambda_1, \ldots, \lambda_m)$ and $\mu = (\mu_1, \ldots, \mu_s)$, where $\mu_j = \lambda_{m+j} - \lambda_{m+s+j}$ for $j = 1, \ldots, s$, we can obtain the properties stated in (3.9) from the last ones. □

Definition 3.1. If $(\bar{x}, \lambda) \in R^n \times R^m$ is such a pair that (3.6) (resp., (3.8)) holds, then we say that $(\bar{x}, \lambda)$ is a *Karush-Kuhn-Tucker pair* (KKT pair for short) of the standard quadratic program (2.1) (resp., of the canonical quadratic program (2.26)). The point $\bar{x}$ is called a *Karush-Kuhn-Tucker point* (KKT point for short), and the real numbers $\lambda_1, \ldots, \lambda_m$ are called the *Lagrange multipliers* corresponding to $\bar{x}$. Similarly, if $(\bar{x}, \lambda, \mu) \in R^n \times R^m \times R^s$ is such a triple that (3.9) is satisfied then $\bar{x}$ is called a *Karush-Kuhn-Tucker point* of the general quadratic program (2.27), and the real numbers $\lambda_1, \ldots, \lambda_m$, $\mu_1, \ldots, \mu_s$ are called the *Lagrange multipliers* corresponding to $\bar{x}$. Sometimes the vectors $\lambda = (\lambda_1, \ldots, \lambda_m)$ and $\mu = (\mu_1, \ldots, \mu_s)$ are also called the *Lagrange multipliers* corresponding to $\bar{x}$.

In the sequel, by abuse of notation, we will abbreviate both the KKT point sets of (2.1) and (2.26) to $S(D, A, c, b)$. Likewise, both the solution sets (resp., both the local-solution sets) of (2.1) and (2.26) are abbreviated to $\text{Sol}(D, A, c, b)$ (resp., to $\text{loc}(D, A, c, b)$).

From Theorem 3.3 and Corollary 3.1 it follows that

$$\text{Sol}(D, A, c, b) \subset \text{loc}(D, A, c, b) \subset S(D, A, c, b). \tag{3.10}$$

Later on, we will encounter with examples where the three sets figured in (3.10) are different each from others.

3.2 Second-Order Optimality Conditions

This section provides a detailed exposition of second-order necessary and sufficient optimality conditions for quadratic programming problems. The main result in this direction was published by Majthay in 1971. It is stated in Theorem 3.4 below. The proof given by Majthay contains one inaccurate argument (see Contesse (1980), p. 331). The first rigorous proof of this result belongs to Contesse (1980).

Theorem 3.4. (See Majthay (1971) and Contesse (1980), Théorème 1) *The necessary and sufficient condition for a point $\bar{x} \in R^n$ to be a local solution of problem* (2.27) *is that there exists a pair of vectors*

$$(\bar{\lambda}, \bar{\mu}) = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m, \bar{\mu}_1, \ldots, \bar{\mu}_s) \in R^m \times R^s$$

such that

(i) *the system*

$$\begin{cases} D\bar{x} - A^T\bar{\lambda} - C^T\bar{\mu} + c = 0, \\ A\bar{x} - b \geq 0, \quad C\bar{x} = d, \quad \bar{\lambda} \geq 0, \\ \bar{\lambda}^T(A\bar{x} - b) = 0. \end{cases} \tag{3.11}$$

is satisfied, and

(ii) *if $v \in R^n \setminus \{0\}$ is such that $A_{I_1}v = 0$, $A_{I_2}v \geq 0$, $Cv = 0$, where*

$$I_1 = \{i \,:\, A_i\bar{x} = b_i,\ \bar{\lambda}_i > 0\}, \quad I_2 = \{i \,:\, A_i\bar{x} = b_i,\ \bar{\lambda}_i = 0\}, \tag{3.12}$$

then $v^TDv \geq 0$.

Consider problem (2.27) and set

$$\Delta = \{x \in R^n \,:\, Ax \geq b,\ Cx = d\}.$$

Denote the objective function of (2.27) by $f(x)$. Let the symbol $(\nabla f(x))^{\perp}$ stand for the linear subspace of R^n orthogonal to $\nabla f(x)$, that is

$$(\nabla f(x))^{\perp} = \{v \in R^n \,:\, \langle \nabla f(x), v \rangle = 0\}.$$

For proving Theorem 3.4 we shall need the following fact.

Lemma 3.1. *Let $\bar{x} \in R^n$, $\bar{\lambda} \in R^m$ and $\bar{\mu} \in R^s$ be such that the system* (3.11) *is satisfied. Let I_1 and I_2 be as in* (3.12). *Then*

$$\begin{aligned}&\{v \in R^n : A_{I_1}v = 0,\ A_{I_2}v \geq 0,\ Cv = 0\}\\&= \{v \in R^n : A_{I_0}v \geq 0,\ Cv = 0\} \cap (\nabla f(\bar{x}))^{\perp}\\&= T_{\Delta}(\bar{x}) \cap (\nabla f(\bar{x}))^{\perp},\end{aligned}$$

where $I_0 := I_1 \cup I_2 = \{i : A_i\bar{x} = b_i\}$, and $T_{\Delta}(\bar{x})$ denotes the tangent cone to Δ at $\bar{x}$.

Proof. Using Theorem 1.8 (b), it is a simple matter to show that

$$T_{\Delta}(\bar{x}) = \{v \in R^n : A_{I_0}v \geq 0,\ Cv = 0\}. \tag{3.13}$$

So it suffices to prove the first equality in the assertion of the lemma. Suppose that $v \in R^n$, $A_{I_1}v = 0$, $A_{I_2}v \geq 0$, $Cv = 0$. Define $I = \{1, 2, \ldots, m\}$. By (3.11) we have

$$\begin{aligned}\langle \nabla f(\bar{x}), v\rangle = (D\bar{x} + c)^T v &= (A^T\bar{\lambda} + C^T\bar{\mu})^T v\\&= \bar{\lambda}^T Av + \bar{\mu}^T \underbrace{Cv}_{=0}\\&= \bar{\lambda}_{I_1}^T \underbrace{A_{I_1}v}_{=0} + \underbrace{\bar{\lambda}_{I_2}^T}_{=0} A_{I_2}v + \underbrace{\bar{\lambda}_{I\setminus I_0}^T}_{=0} A_{I\setminus I_0}v\\&= 0.\end{aligned}$$

Hence $v \in (\nabla f(\bar{x}))^{\perp}$. It follows that

$$\begin{aligned}&\{v \in R^n : A_{I_1}v = 0,\ A_{I_2}v \geq 0,\ Cv = 0\}\\&\subset \{v \in R^n : A_{I_0}v \geq 0,\ Cv = 0\} \cap (\nabla f(\bar{x}))^{\perp}.\end{aligned}$$

To obtain the reverse inclusion, suppose that $v \in R^n$, $A_{I_0}v \geq 0$, $Cv = 0$, $v \in (\nabla f(\bar{x}))^{\perp}$. We need only to show that $A_{I_1}v = 0$. From (3.11) we deduce that

$$\begin{aligned}0 = \langle \nabla f(\bar{x}), v\rangle &= (D\bar{x} + c)^T v\\&= (A^T\bar{\lambda} + C^T\bar{\mu})^T v\\&= \bar{\lambda}^T Av\\&= \underbrace{\bar{\lambda}_{I_1}^T}_{>0} \underbrace{A_{I_1}v}_{\geq 0} + \underbrace{\bar{\lambda}_{I_2}^T}_{=0} A_{I_2}v.\end{aligned}$$

Hence $A_{I_1}v = 0$, and the proof is complete. □

Note that Theorem 3.4 can be reformulated in the following equivalent form which does not require the use of Lagrange multipliers.

Theorem 3.5. (See Cottle et al. (1992), p. 116) *The necessary and sufficient condition for a point $\bar{x} \in R^n$ to be a local solution of problem* (2.27) *is that the next two properties are valid:*

(i) $\langle \nabla f(\bar{x}), v \rangle = (D\bar{x} + c)^T v \geq 0$ *for every* $v \in T_\Delta(\bar{x}) = \{v \in R^n : A_{I_0} v \geq 0,\ Cv = 0\}$, *where* $I_0 = \{i : A_i \bar{x} = b_i\}$;

(ii) $v^T D v \geq 0$ *for every* $v \in T_\Delta(\bar{x}) \cap (\nabla f(\bar{x}))^\perp$, *where* $(\nabla f(\bar{x}))^\perp = \{v \in R^n : \langle \nabla f(\bar{x}), v \rangle = 0\}$.

The fact that the first property is equivalent to the existence of a pair $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^s$ satisfying system (3.11) can be established by using the Farkas Lemma (see Theorem 3.2) and some arguments similar to those in the proof of Lemma 3.1. The equivalence between property (ii) in Theorem 3.5 and property (ii) in Theorem 3.4, which is formulated via a Lagrange multipliers set $(\lambda, \bar{\mu}) \in R^m \times R^s$, follows from Lemma 3.1. Hence Theorem 3.5 is an equivalent form of Theorem 3.4.

Proof of Theorem 3.4.

Necessity: Suppose that $\bar{x}$ is a local solution of (2.27). Then there exists $\varepsilon > 0$ such that

$$f(x) - f(\bar{x}) \geq 0 \quad \forall x \in \Delta \cap B(\bar{x}, \varepsilon). \tag{3.14}$$

According to Corollary 3.2, there exists $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^n$ such that condition (i) is satisfied. Let I_1 and I_2 be defined as in (3.12). Suppose that property (ii) were false. Then we could find $\bar{v} \in R^n \setminus \{0\}$ such that

$$A_{I_1}\bar{v} = 0, \quad A_{I_2}\bar{v} \geq 0, \quad C\bar{v} = 0, \quad \bar{v}^T D \bar{v} < 0.$$

By Lemma 3.1, $(D\bar{x} + c)^T \bar{v} = \langle \nabla f(\bar{x}), \bar{v} \rangle = 0$. Consequently, for each $t \in (0, 1)$ we have

$$f(\bar{x} + t\bar{v}) - f(\bar{x}) = t(D\bar{x} + c)^T \bar{v} + \frac{t^2}{2}\bar{v}^T D \bar{v} = \frac{1}{2} t^2 \bar{v}^T D \bar{v} < 0.$$

As $\bar{x} + t\bar{v} \in \Delta \cap B(\bar{x}, \varepsilon)$ for all $t \in (0, 1)$ sufficiently small, the last fact contradicts (3.14). Thus (ii) must hold true.

Sufficiency: Suppose that $\bar{x} \in R^n$ is such that there exists $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^n$ such that conditions (i) and (ii) are satisfied. We shall prove that $\bar{x}$ is a local solution of (2.27). The main idea of the proof is to decompose the tangent cone $T_\Delta(\bar{x})$ into the sum of a subspace and a pointed polyhedral convex cone. Set $I = \{1, 2, \ldots, m\}$, $I_0 = \{i \in I : A_i \bar{x} = b_i\}$, and observe that $I_0 = I_1 \cup I_2$. Define

$$M = \{v \in R^n : A_{I_0} v = 0,\ Cv = 0\},$$
$$M^\perp = \{v \in R^n : \langle v, u \rangle = 0\ \forall u \in M\}.$$

Let

$$\begin{aligned} K &= \{v \in M^{\perp} : v = z - u \text{ for some } z \in T_{\Delta}(\bar{x}) \text{ and } u \in M\} \\ &= \mathrm{Pr}_{M^{\perp}}(T_{\Delta}(\bar{x})), \end{aligned}$$

where $\mathrm{Pr}_{M^{\perp}}(\cdot)$ denotes the orthogonal projection of R^n onto $M^{\perp}$. Since $K = \mathrm{Pr}_{M^{\perp}}(T_{\Delta}(\bar{x}))$ and $M \subset T_{\Delta}(\bar{x})$, it follows that

$$T_{\Delta}(\bar{x}) = M + K. \tag{3.15}$$

We have

$$K = M^{\perp} \cap T_{\Delta}(\bar{x}). \tag{3.16}$$

Indeed, if $v \in K$ then $v \in M^{\perp}$ and $v = z - u$ for some $z \in T_{\Delta}(\bar{x})$ and $u \in M$. Hence

$$A_{I_0}u = 0, \quad Cu = 0, \quad A_{I_0}z \geq 0, \quad Cz = 0.$$

Therefore

$$\begin{aligned} &A_{I_0}v = A_{I_0}z - A_{I_0}u = A_{I_0}z \geq 0, \\ &Cv = Cz - Cu = 0. \end{aligned}$$

So $v \in M^{\perp} \cap T_{\Delta}(\bar{x})$. It follows that $K \subset M^{\perp} \cap T_{\Delta}(\bar{x})$. For proving the reverse inclusion, it suffices to note that each $v \in M^{\perp} \cap T_{\Delta}(\bar{x})$ admits the representation $v = v - 0$ where $v \in T_{\Delta}(\bar{x})$, $0 \in M$.

We next show that K is a pointed polyhedral convex cone. Recall that a cone $K \subset R^n$ is said to be *pointed* if $K \cap (-K) = \{0\}$. From (3.16) it follows that K is a polyhedral convex cone. So we need only to show that K is pointed. If it were true that K is not pointed, there would be $v \in K \setminus \{0\}$ such that $-v \in K$. On one hand, from (3.13) and (3.16) it follows that

$$A_{I_0}v \geq 0, \quad Cv = 0, \quad A_{I_0}(-v) \geq 0.$$

This implies that $A_{I_0}v = 0$ and $Cv = 0$. So $v \in M$. On the other hand, since $v \in K$, by (3.16) we have $v \in M^{\perp}$. Thus $v \in M \cap M^{\perp} = \{0\}$, a contradiction.

Define $K_0 = \{v \in K : \langle \nabla f(\bar{x}), v \rangle = 0\}$. Since K is a pointed polyhedral convex cone, we see that K_0 is also a pointed polyhedral convex cone.

From (i) it follows that

$$\langle \nabla f(\bar{x}), v \rangle \geq 0 \quad \forall v \in T_{\Delta}(\bar{x}). \tag{3.17}$$

Indeed, let $v \in T_\Delta(\bar{x})$. By (i) and (3.13),

$$\begin{aligned}\langle \nabla f(\bar{x}), v\rangle = (D\bar{x}+c)^T v &= (A^T\bar{\lambda} + C^T\bar{\mu})^T v\\ &= \bar{\lambda}^T Av + \bar{\mu}^T Cv\\ &= \bar{\lambda}^T Av\\ &\geq 0.\end{aligned}$$

Since $M \subset T_\Delta(\bar{x})$ and $-v \in M$ whenever $v \in M$, it follows that

$$\langle \nabla f(\bar{x}), v\rangle = 0 \quad \forall v \in M. \tag{3.18}$$

From (3.17) and the definition of K_0 we have

$$\langle \nabla f(\bar{x}), v\rangle > 0 \quad \forall v \in K \setminus K_0. \tag{3.19}$$

Since K is a polyhedral convex cone, according to Theorem 19.1 in Rockafellar (1970), K is a *finitely generated cone.* The latter means that there exists a finite system of nonzero vectors $\{z^1, \ldots, z^q\}$, called the *generators* of K, such that

$$K = \{v = \sum_{j=1}^{q} t_j z^j \ : \ t_j \geq 0 \text{ for all } j = 1, \ldots, q\}. \tag{3.20}$$

If $K_0 \neq \{0\}$ then some of the vectors z^j $(j = 1, \ldots, q)$ must belong to K_0. To prove this, suppose, contrary to our claim, that $K_0 \neq \{0\}$ and all the generators z^j $(j = 1, \ldots, q)$ belong to $K \setminus K_0$. Let $\bar{v} = \sum_{j=1}^{q} t_j z^j$, where $t_j \geq 0$ for all j, be a nonzero vector from K_0. Since at least one of the values t_j must be nonzero, from (3.19) we deduce that

$$\langle \nabla f(\bar{x}), \bar{v}\rangle = \sum_{j=1}^{q} t_j \langle \nabla f(\bar{x}), z^j\rangle > 0.$$

This contradicts our assumption that $\bar{v} \in K_0$. If $K_0 \neq \{0\}$, there is no loss of generality in assuming that the first q_0 generators z^j $(j = 1, \ldots, q_0)$ belong to K_0, and the other generators z_j $(j = q_0 + 1, \ldots, q)$ belong to $K \setminus K_0$. Thus $\langle \nabla f(\bar{x}), z^j\rangle = 0$ for $j = 1, \ldots, q_0$, $\langle \nabla f(\bar{x}), z^j\rangle > 0$ for $j = q_0 + 1, \ldots, q$, and $q_0 \in \{1, \ldots, q\}$.

We are now in a position to prove the following claim:

Claim. *$\bar{x}$ is a local solution of* (2.27).

If $\bar{x}$ were not a local solution of (2.27), we would find a sequence $\{x^k\} \subset \Delta$ such that $x^k \to \bar{x}$, and

$$f(x^k) - f(\bar{x}) < 0 \quad \forall k \in N.$$

For each $k \in N$, on account of (3.15), we have

$$x^k - \bar{x} \in T_\Delta(\bar{x}) = M + K.$$

Combining this with (3.20) we deduce that there exist $t_j^k \geq 0$ ($j = 1, \ldots, q$) and $u^k \in M$ such that

$$x^k - \bar{x} = u^k + \sum_{j=1}^{q_0} t_j^k z^j + \sum_{j=q_0+1}^{q} t_j^k z^j. \tag{3.21}$$

Writing $v^k = \sum_{j=1}^{q_0} t_j^k z^j$ and $w^k = \sum_{j=q_0+1}^{q} t_j^k z^j$ yields

$$x^k - \bar{x} = u^k + v^k + w^k. \tag{3.22}$$

It is understood that v^k (resp., w^k) is absent in the last representation if $K_0 = \{0\}$ (resp., $K \setminus K_0 = \emptyset$). There are two possible cases:

- *Case 1:* There exists a subsequence $\{k'\} \subset \{k\}$ such that $w^{k'} = 0$ for all k'. (If $K \setminus K_0 = \emptyset$, then w^k is vacuous for all $k \in N$. Such situation is also included in this case.)

- *Case 2:* There exists a number $k_* \in N$ such that $w^k \neq 0$ for every $k \geq k_*$.

If *Case 1* occurs, then without restriction of generality we can assume that $\{k'\} \equiv \{k\}$. Since $x^k - \bar{x} = u^k + v^k$, from (3.18) we have

$$\begin{aligned}(D\bar{x} + c)^T(x^k - \bar{x}) = \langle \nabla f(\bar{x}), x^k - \bar{x}\rangle \quad &= \langle \nabla f(\bar{x}), u^k\rangle + \langle \nabla f(\bar{x}), v^k\rangle \\ &= 0.\end{aligned}$$

Therefore

$$\begin{aligned} f(x^k) - f(\bar{x}) \quad &= \frac{1}{2}(x^k - \bar{x})^T D(x^k - \bar{x}) + (D\bar{x} + c)^T(x^k - \bar{x}) \\ &= \frac{1}{2}(x^k - \bar{x})^T D(x^k - \bar{x}). \end{aligned}$$

Hence

$$(x^k - \bar{x})^T D(x^k - \bar{x}) < 0 \quad \forall k \in N. \tag{3.23}$$

Since $x^k - \bar{x} \in T_\Delta(\bar{x})$ and $\langle \nabla f(\bar{x}), x^k - \bar{x}\rangle = 0$, we have $x^k - \bar{x} \in T_\Delta(\bar{x}) \cap (\nabla f(\bar{x}))^\perp$. Consequently, from Lemma 3.1 and from assumption (ii) we obtain $(x^k - \bar{x})^T D(x^k - \bar{x}) \geq 0$, contrary to (3.23).

If *Case 2* happens, then there is no loss of generality in assuming that $w^k \neq 0$ for all $k \in N$. For each k, since t_j^k $(j = q_0 + 1, \dots, q)$ are nonnegative and not all zero, there must exists some $j(k) \in \{q_0 + 1, \dots, q\}$ such that

$$t_{j(k)}^k = \max\{t_j^k \,:\, j \in \{q_0 + 1, \dots, q\}\} > 0.$$

It is clear that there must exist an index $j_* \in \{q_0 + 1, \dots, q\}$ and a subsequence $\{k'\} \subset \{k\}$ such that $j(k') = j_*$ for every k'. Without loss of generality we can assume that $\{k'\} \equiv \{k\}$. On account of (3.21) and (3.22), we have

$$\begin{aligned} f(x^k) - f(\bar{x}) &= \frac{1}{2}(x^k - \bar{x})^T D(x^k - \bar{x}) + (D\bar{x} + c)^T (x^k - \bar{x}) \\ &= \frac{1}{2}(u^k + v^k + w^k)^T D(u^k + v^k + w^k) \\ &\qquad\qquad + (D\bar{x} + c)^T (u^k + v^k + w^k) \\ &= \underbrace{\frac{1}{2}(u^k + v^k)^T D(u^k + v^k)}_{\geq 0} + (u^k + v^k)^T D w^k \\ &\qquad + \textstyle\sum_{j=q_0+1}^q t_j^k (D\bar{x} + c)^T z^j + \frac{1}{2}(w^k)^T D w^k \\ &\geq \textstyle\sum_{j=q_0+1}^q t_j^k (u^k + v^k)^T D z^j + t_{j_*}^k (D\bar{x} + c)^T z^{j_*} \\ &\qquad + \frac{1}{2}\left(\sum_{j=q_0+1}^{q} t_j^k z^j\right)^T D w^k. \end{aligned}$$

In these transformations, we have used the inequality

$$(u^k + v^k)^T D(u^k + v^k) \geq 0,$$

which is a consequence of Lemma 3.1 and condition (ii). From what has already been proved, it follows that

$$0 > \sum_{j=q_0+1}^{q} t_j^k (u^k + v^k)^T D z^j + t_{j_*}^k (D\bar{x} + c)^T z^{j_*} + \frac{1}{2}\left(\sum_{j=q_0+1}^{q} t_j^k z^j\right)^T D w^k. \tag{3.24}$$

Dividing (3.24) by $t_{j_*}^k$, noting that $0 \leq t_j^k / t_{j_*}^k \leq 1$ for every $j = q_0 + 1, \dots, q$, letting $k \to \infty$ and using the following

FACT. *If $x^k \to \bar{x}$ then $u^k \to 0$, $v^k \to 0$, and $w^k \to 0$,* we get

$$0 \geq (D\bar{x} + c)^T z^{j_*}, \tag{3.25}$$

a contradiction. This finishes the proof our Claim.

What is left is to show that the Fact is true. For proving it, we first observe that

$$\begin{aligned}\|x^k - \bar{x}\|^2 = \langle x^k - \bar{x}, x^k - \bar{x}\rangle &= \langle u^k + v^k + w^k, u^k + v^k + w^k\rangle \\ &= \|u^k\|^2 + \|v^k + w^k\|^2.\end{aligned}$$

Since $\|x^k - \bar{x}\| \to 0$, it follows that $u^k \to 0$ and $v^k + w^k \to 0$. We have

$$v^k + w^k = \sum_{j=1}^{q} t_j^k z^j.$$

It suffices to prove that, for any $j \in \{1, \ldots, q\}$, $t_j^k \to 0$ as $k \to \infty$. On the contrary, suppose that there exists $j_1 \in \{1, \ldots, q\}$ such that the sequence $\{t_{j_1}^k\}$ does not converge to 0 as $k \to \infty$. Then there exist $\varepsilon > 0$ and a subsequence $\{k'\} \subset \{k\}$ such that $t_{j_1}^{k'} \geq \varepsilon$ for every k'. Since $\sum_{j=1}^q t_j^{k'} \geq t_{j_1}^{k'} \geq \varepsilon$ for every k', we can write

$$v^{k'} + w^{k'} = \sum_{j=1}^{q} t_j^{k'} z^j = \left(\sum_{j=1}^{q} t_j^{k'}\right) \sum_{j=1}^{q} \frac{t_j^{k'}}{\sum_{j=1}^q t_j^{k'}} z^j. \tag{3.26}$$

Replace $\{k'\}$ by a subsequence if necessary, we can assume that, for every $j \in \{1, \ldots, q\}$,

$$\frac{t_j^{k'}}{\sum_{j=1}^q t_j^{k'}} \to \bar{\tau}_j$$

for some $\bar{\tau}_j \in [0, 1]$. It is clear that $\sum_{j=1}^q \bar{\tau}_j = 1$. We must have $\sum_{j=1}^q \bar{\tau}^j z^j \neq 0$. Indeed, if it were true that $\sum_{j=1}^q \bar{\tau}_j z^j \neq 0$, there would be some $j_0 \in \{1, \ldots, q\}$ such that $\bar{\tau}_{j_0} \neq 0$. Then

$$\sum_{j \neq j_0} \bar{\tau}_j z^j = -\bar{\tau}_{j_0} z^{j_0}.$$

This implies that $-\bar{\tau}_{j_0} z^{j_0} \in K$, $\bar{\tau}_{j_0} z^{j_0} \in K$, $\bar{\tau}_{j_0} z^{j_0} \neq 0$. Hence the cone K is not pointed, a contradiction. We have thus proved that $\bar{z} := \sum_{j=1}^q \bar{\tau}_j z^j$ is a nonzero vector. If the sequence $\{\sum_{j=1}^q t_j^{k'}\}$ is bounded, then without loss of generality we can assume that it converges to some limit $\hat{\tau} \geq \varepsilon$. Letting $k' \to \infty$, from (3.26) we deduce that $0 = \hat{\tau}\bar{z}$, a contradiction. If the sequence $\{\sum_{j=1}^q t_j^{k'}\}$ is unbounded, then without loss of generality we can assume that it converges to $+\infty$. From (3.26) it follows that

$$\|v^{k'} + w^{k'}\| = \left(\sum_{j=1}^{q} t_j^{k'}\right) \left\| \sum_{j=1}^{q} \frac{t_j^{k'}}{\sum_{j=1}^{q} t_j^{k'}} z^j \right\|.$$

Letting $k' \to \infty$ we obtain $0 = +\infty \|\bar{z}\|$, an absurd. □

Definition 3.2. (See Mangasarian (1980), p. 201) A point $\bar{x} \in \Delta$ is called a *locally unique solution* of the problem $\min\{f(x) : x \in \Delta\}$, where $f : R^n \to R$ is a real function and $\Delta \subset R^n$ is a given subset, if there exists $\varepsilon > 0$ such that

$$f(x) > f(\bar{x}) \quad \forall x \in (\Delta \cap B(\bar{x}, \varepsilon)) \setminus \{\bar{x}\}.$$

Of course, if $\bar{x}$ is a locally unique solution of a minimization problem then it is a local solution of that problem. The converse is not true in general.

The following theorem describes the (second-order) necessary and sufficient condition for a point to be a locally unique solution of a quadratic program.

Theorem 3.6. (See Mangasarian (1980), Theorem 2.1, and Contesse (1980), Théorème 1) *The necessary and sufficient condition for a point $\bar{x} \in R^n$ to be a locally unique solution of problem* (2.27) *is that there exists a pair of vectors*

$$(\bar{\lambda}, \bar{\mu}) = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m, \bar{\mu}_1, \ldots, \bar{\mu}_s) \in R^m \times R^s$$

such that

(i) *The system* (3.11) *is satisfied, and*

(ii) *If $v \in R^n \setminus \{0\}$ is such that $A_{I_1} v = 0$, $A_{I_2} v \geq 0$, $Cv = 0$, where*

$$I_1 = \{i : A_i \bar{x} = b_i, \ \bar{\lambda}_i > 0\}, \quad I_2 = \{i : A_i \bar{x} = b_i, \ \bar{\lambda}_i = 0\},$$

then $v^T D v > 0$.

Proof. *Necessity:* Suppose that $\bar{x}$ is a locally unique solution of (2.27). Then there exists $\varepsilon > 0$ such that

$$f(x) - f(\bar{x}) > 0 \quad \forall x \in (\Delta \cap B(\bar{x}, \varepsilon)) \setminus \{\bar{x}\}. \tag{3.27}$$

According to Corollary 3.2, there exists $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^n$ such that condition (i) is satisfied. Suppose that property (ii) were false. Then we could find $\bar{v} \in R^n \setminus \{0\}$ such that

$$A_{I_1}\bar{v} = 0, \quad A_{I_2}\bar{v} \geq 0, \quad C\bar{v} = 0, \quad \bar{v}^T D\bar{v} \leq 0.$$

By Lemma 3.1, $(D\bar{x} + c)^T\bar{v} = \langle \nabla f(\bar{x}), \bar{v}\rangle = 0$. Consequently, for each $t \in (0,1)$ we have

$$f(\bar{x} + t\bar{v}) - f(\bar{x}) = t(D\bar{x} + c)^T\bar{v} + \frac{t^2}{2}\bar{v}^T D\bar{v} = \frac{1}{2}t^2\bar{v}^T D\bar{v} \leq 0.$$

As $\bar{x} + t\bar{v} \in \Delta \cap B(\bar{x}, \varepsilon)$ for all $t \in (0,1)$ sufficiently small, the last fact contradicts (3.27). Thus (ii) must hold true.

Sufficiency: Suppose that $\bar{x} \in R^n$ is such that there exists $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^n$ such that (i) and (ii) are satisfied. We shall show that $\bar{x}$ is a locally unique solution of (2.27). Set $I = \{1, 2, \ldots, m\}$, $I_0 = \{i \in I \; : \; A_i\bar{x} = b_i\}$. Let M, $M^\perp$, K, K_0, $z^1, \ldots, z^q$, and q_0 be defined as in the proof of Theorem 3.4. Then the properties (3.15)–(3.20) are valid.

If $\bar{x}$ were not a locally unique solution of (2.27), we would find a sequence $\{x^k\} \subset \Delta$ such that $x^k \to \bar{x}$, and

$$f(x^k) - f(\bar{x}) \leq 0 \quad \forall k \in N.$$

For each $k \in N$, on account of (3.15), we have

$$x^k - \bar{x} \in T_\Delta(\bar{x}) = M + K.$$

Combining this with (3.20) we conclude that there exist $t_j^k \geq 0$ $(j = 1, \ldots, q)$ and $u^k \in M$ such that (3.21) holds. Setting $v^k = \sum_{j=1}^{q_0} t_j^k z^j$ and $w^k = \sum_{j=q_0+1}^{q} t_j^k z^j$ we have (3.22). As before, if $K_0 = \{0\}$ (resp., $K \setminus K_0 = \emptyset$) then it is understood that v^k (resp., w^k) is absent in the representation (3.22). We consider separately the following two cases:

- *Case 1:* There exists a subsequence $\{k'\} \subset \{k\}$ such that $w^{k'} = 0$ for all k'. (If $K \setminus K_0 = \emptyset$, then w^k is vacuous for all $k \in N$. Such situation is also included in this case.)

- *Case 2:* There exists a number $k_* \in N$ such that $w^k \neq 0$ for every $k \geq k_*$.

If *Case 1* occurs, then without restriction of generality we can assume that $\{k'\} \equiv \{k\}$. Arguing similarly as in the analysis of *Case 1* in the preceding proof, we obtain

$$(x^k - \bar{x})^T D(x^k - \bar{x}) \leq 0. \tag{3.28}$$

Since $x^k - \bar{x} \in T_\Delta(\bar{x})$ and $\langle \nabla f(\bar{x}), x^k - \bar{x}\rangle = 0$, we have $x^k - \bar{x} \in T_\Delta(\bar{x}) \cap (\nabla f(\bar{x}))^\perp$. Hence from Lemma 3.1 and from assumption (ii) it follows that $(x^k - \bar{x})^T D(x^k - \bar{x}) > 0$, contrary to (3.28).

If *Case 2* happens then there is no loss of generality in assuming that $w^k \neq 0$ for all $k \in N$. Construct the sequence $\{j(k)\}$ $(k \in N)$ as in the proof of Theorem 3.4. Then there must exist an index $j_* \in \{q_0+1, \dots, q\}$ and a subsequence $\{k'\} \subset \{k\}$ such that $j(k') = j_*$ for every k'. Without loss of generality we can assume that $\{k'\} \equiv \{k\}$. Analysis similar to that in the proof of Theorem 3.4 shows that

$$0 \geq \sum_{j=q_0+1}^{q} t_j^k (u^k+v^k)^T D z^j + t_{j_*}^k (D\bar{x}+c)^T z^{j_*} + \frac{1}{2}\left(\sum_{j=q_0+1}^{q} t_j^k z^j\right)^T Dw^k. \tag{3.29}$$

Dividing (3.29) by $t_{j_*}^k$, noting that $0 \leq t_j^k / t_{j_*}^k \leq 1$ for every $j = q_0 + 1, \dots, q$, letting $k \to \infty$ and using the Fact established in the preceding proof, we get (3.25). This contradicts (3.19) because $z^{j_*} \in K \setminus K_0$. We have thus proved that $\bar{x}$ is a locally unique solution of (2.27). □

Note that Theorem 3.6 can be reformulated in the following equivalent form which does not require the use of Lagrange multipliers.

Theorem 3.7. *The necessary and sufficient condition for a point $\bar{x} \in R^n$ to be a locally unique solution of problem (2.27) is that the next two properties are valid:*

(i) $\langle \nabla f(\bar{x}), v\rangle = (D\bar{x} + c)^T v \geq 0$ *for every* $v \in T_\Delta(\bar{x}) = \{v \in R^n : A_{I_0} v \geq 0,\ Cv = 0\}$, *where* $I_0 = \{i : A_i \bar{x} = b_i\}$;

(ii) $v^T D v > 0$ *for every nonzero vector* $v \in T_\Delta(\bar{x}) \cap (\nabla f(\bar{x}))^\perp$, *where* $(\nabla f(\bar{x}))^\perp = \{v \in R^n : \langle \nabla f(\bar{x}), v\rangle = 0\}$.

As it has been noted after the formulation of Theorem 3.5, the first property is equivalent to the existence of a pair $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^s$ satisfying system (3.11). The equivalence between property (ii) in Theorem 3.7 and property (ii) in Theorem 3.6, which is formulated via a Lagrange multipliers set $(\lambda, \bar{\mu}) \in R^m \times R^s$, follows from Lemma 3.1. Hence Theorem 3.7 is an equivalent form of Theorem 3.6.

It is interesting to observe that if $\bar{x}$ is a locally unique solution of a quadratic program then a property similar to (3.4) holds.

Theorem 3.8. *If $\bar{x} \in R^n$ is a locally unique solution of problem (2.27) then there exist $\varepsilon > 0$ and $\varrho > 0$ such that*

$$f(x) - f(\bar{x}) \geq \varrho \|x - \bar{x}\|^2 \quad \text{for every } x \in \Delta \cap B(\bar{x}, \varepsilon), \tag{3.30}$$

where $\Delta = \{x \in R^n : Ax \geq b,\ Cx = d\}$ is the constraint set of (2.27).

Proof. Let $\bar{x} \in R^n$ be a locally unique solution of (2.27). By Theorem 3.6, there exists a pair of vectors

$$(\bar{\lambda}, \bar{\mu}) = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m, \bar{\mu}_1, \ldots, \bar{\mu}_s) \in R^m \times R^s$$

such that

(i)' The system (3.11) is satisfied, and

(ii)' If $v \in R^n \setminus \{0\}$ is such that $A_{I_1} v = 0$, $A_{I_2} v \geq 0$, $Cv = 0$, where

$$I_1 = \{i : A_i \bar{x} = b_i,\ \bar{\lambda}_i > 0\}, \quad I_2 = \{i : A_i \bar{x} = b_i,\ \bar{\lambda}_i = 0\},$$

then $v^T D v > 0$.

As it has been noted in the proof of Theorem 3.4, from (i)' it follows that (3.17) is valid.

To obtain a contradiction, suppose that one cannot find any pair of positive numbers (ε, ϱ) satisfying (3.30). Then, for each $k \in N$, there exists $x^k \in \Delta$ such that $\|x^k - \bar{x}\| \leq \dfrac{1}{k}$ and

$$f(x^k) - f(\bar{x}) < \frac{1}{k}\|x^k - \bar{x}\|^2. \tag{3.31}$$

The last inequality implies that $x^k \neq \bar{x}$. Without loss of generality we can assume that the sequence $\{(x^k - \bar{x})/\|x^k - \bar{x}\|\}$ converges to some $\bar{v} \in R^n$ with $\|\bar{v}\| = 1$. By (3.31), we have

$$\frac{1}{k}\|x^k - \bar{x}\|^2 > f(x^k) - f(\bar{x}) = \frac{1}{2}(x^k - \bar{x})^T D (x^k - \bar{x}) + (D\bar{x} + c)^T (x^k - \bar{x}). \tag{3.32}$$

Dividing this expression by $\|x^k - \bar{x}\|$ and letting $k \to \infty$ we get $0 \geq (D\bar{x} + c)^T \bar{v}$. Since $x^k - \bar{x} \in T_\Delta(\bar{x})$ for every $k \in N$, we must have $\bar{v} \in T_\Delta(\bar{x})$. By (3.17), $(D\bar{x}+c)^T\bar{v} \geq 0$. Thus $\langle \nabla f(\bar{x}), \bar{v}\rangle = (D\bar{x}+c)^T\bar{v} = 0$. As $x^k - \bar{x} \in T_\Delta(\bar{x})$ for every $k \in N$, according to (3.17) we have $(D\bar{x}+c)^T(x^k-\bar{x}) \geq 0$. Combining this with (3.32) yields

$$\frac{1}{k}\|x^k - \bar{x}\|^2 > \frac{1}{2}(x^k-\bar{x})^T D(x^k-\bar{x}).$$

Dividing the last inequality by $\|x^k-\bar{x}\|^2$ and letting $k \to \infty$ we obtain $0 \geq \bar{v}^T D\bar{v}$. Since $\bar{v} \in T_\Delta(\bar{x}) \cap (\nabla f(\bar{x}))^\perp$, from Lemma 3.1 and (ii)' it follows that $\bar{v}^T D\bar{v} > 0$. We have arrived at a contradiction. The proof is complete. □

3.3 Commentaries

First-order necessary and sufficient optimality conditions for (nonconvex) quadratic programs are proved in several textbooks. Meanwhile, to our knowledge, the paper of Contesse (1980) is the only place where one can find a satisfactory proof of the second-order necessary and sufficient optimality condition for quadratic programs which was noted firstly by Majthay in 1971 and which has many interesting applications (see, for instance, Cottle et al. (1992) and Chapters 4, 10, 14 of this book). The reason might be that the proof is rather long and complicated. The proof described in this chapter is essentially that one of Contesse. For the benefit of the reader, we have proposed a series of minor modifications in the presentation. The formulation given in Theorem 3.4 can be used effectively in performing practical calculations to find the local solution set, while the formulation given in Theorem 3.5 is very convenient for theoretical investigations concerning the solution sets of quadratic programs (see the next chapter).

The necessary and sufficient condition for locally unique solutions of quadratic programs described in Theorem 3.6 and Theorem 3.7 is also a good criterion for the stability of the local solutions. The result formulated in Theorem 3.6 was obtained independently by Mangasarian (1980) and Contesse (1980). The proof given in this chapter follows the scheme proposed by Contesse. Another nice proof of the "Necessity" part of Theorem 3.6 can be found in Mangasarian (1980). In Mangasarian (1980) it was noted that the "Sufficiency" part of the result stated in Theorem 3.6 follows from the general second-order sufficient optimality condition for smooth mathematical programming problems established by McCormick (see McCormick (1967), Theorem 6). Actually, the studies of Majthay (1971), Mangasarian (1980), and Contesse (1980) on second-order optimality conditions for quadratic programs have been originated from that work of McCormick.

In mathematical programming theory, it is well known that the estimation like the one in (3.4) (resp., in (3.30)) is a consequence of a strict first-order sufficient optimality condition (resp., of a strong second-order sufficient optimality condition). In this chapter, the two estimations are obtained by simple direct proofs.

Chapter 4

Properties of the Solution Sets of Quadratic Programs

This chapter investigates the structure of the solution sets of quadratic programming problems. We consider the problem

$$(P) \qquad \min\{f(x) = \frac{1}{2}x^T Dx + c^T x \,:\, x \in R^n,\ Ax \geq b,\ Cx = d\},$$

where $D \in R_S^{n\times n}$, $A \in R^{m\times n}$, $C \in R^{s\times n}$, $c \in R^n$, $b \in R^m$, $d \in R^s$. Let

$$\Delta = \{x \in R^n \,:\, Ax \geq b,\ Cx = d\},\ I = \{1, \ldots, m\},\ J = \{1, \ldots, s\}.$$

Denote by $\mathrm{Sol}(P)$, $\mathrm{loc}(P)$ and $S(P)$, respectively, the solution set, the local-solution set and the KKT point set of (P). Our aim is to study such properties of the solution sets $\mathrm{Sol}(P)$, $\mathrm{loc}(P)$ and $S(P)$ as boundedness, closedness and finiteness. Note that sometimes the elements of $S(P)$ are called the *Karush-Kuhn-Tucker solutions* of (P). *The above notations will be kept throughout this chapter.*

4.1 Characterizations of the Unboundedness of the Solution Sets

Denote by $\mathrm{Sol}(P_0)$ the solution set of the following *homogeneous* quadratic program associated with (P):

$$(P_0) \qquad \min\left\{\frac{1}{2}v^T Dv \,:\, v \in R^n,\ Av \geq 0,\ Cv = 0\right\}.$$

Definition 4.1. A half-line $\omega = \{\bar{x} + t\bar{v} \, : \, t \geq 0\}$, where $\bar{v} \in R^n \setminus \{0\}$, which is a subset of $\mathrm{Sol}(P)$ (resp., $\mathrm{loc}(P)$, $S(P)$), is called a *solution ray* (resp., a *local-solution ray*, a *KKT point ray*) of (P).

Theorem 4.1. *The set* $\mathrm{Sol}(P)$ *is unbounded if and only if* (P) *has a solution ray. A necessary and sufficient condition for* $\mathrm{Sol}(P)$ *to be unbounded is that there exist* $\bar{x} \in \mathrm{Sol}(P)$ *and* $\bar{v} \in \mathrm{Sol}(P_0) \setminus \{0\}$ *such that*

$$(D\bar{x} + c)^T\bar{v} = 0. \tag{4.1}$$

The following fact follows directly from the above theorem.

Corollary 4.1. *If the solution set* $\mathrm{Sol}(P_0)$ *is empty or it consists of just one element* 0 *then, for any* $(c, b, d) \in R^n \times R^m \times R^s$, *the solution set* $\mathrm{Sol}(P)$ *is bounded. In the case where* $\mathrm{Sol}(P_0)$ *contains a nonzero element, if*

$$(D\bar{x} + c)^T\bar{v} > 0 \quad \forall \bar{x} \in \mathrm{Sol}(P), \ \forall \bar{v} \in \mathrm{Sol}(P_0) \setminus \{0\},$$

then $\mathrm{Sol}(P)$ *is bounded.*

Proof of Theorem 4.1.

Suppose that $\mathrm{Sol}(P)$ is unbounded. Then there exists a sequence $\{x^k\}$ in $\mathrm{Sol}(P)$ such that $\|x^k\| \to +\infty$ as $k \to \infty$. Without loss of generality we can assume that $x^k \neq 0$ for all k and $\dfrac{x^k}{\|x^k\|} \to \bar{v}$ with $\|\bar{v}\| = 1$. We will show that $\bar{v} \in \mathrm{Sol}(P_0)$. Since $x^k \in \mathrm{Sol}(P)$, we have $Ax^k \geq b$ and $Cx^k = d$. This implies that $A\dfrac{x^k}{\|x^k\|} \geq \dfrac{b}{\|x^k\|}$ and $C\dfrac{x^k}{\|x^k\|} = \dfrac{d}{\|x^k\|}$. Letting $k \to \infty$ we obtain $A\bar{v} \geq 0$ and $C\bar{v} = 0$. Hence $\bar{v}$ is a feasible vector of (P_0). Since $\mathrm{Sol}(P) \neq \emptyset$, by the Eaves Theorem (see Corollary 2.6), $v^TDv \geq 0$ for every $v \in R^n$ satisfying $Av \geq 0$, $Cv = 0$. In particular, $\bar{v}^TD\bar{v} \geq 0$. Fix a point $\hat{x} \in \Delta$. Since $x^k \in \mathrm{Sol}(P)$, we have

$$\frac{1}{2}(x^k)^TDx^k + c^Tx^k \leq f(\hat{x}) \quad (\forall k \in N).$$

Dividing this inequality by $\|x^k\|^2$ and letting $k \to \infty$, we obtain $\bar{v}^TD\bar{v} \leq 0$. Hence

$$\bar{v}^TD\bar{v} = 0. \tag{4.2}$$

Let $v \in R^n$ be any feasible vector of (P_0), that is $Av \geq 0$ and $Cv = 0$. On account of a preceding remark, we have $v^TDv \geq 0$. Combining this with (4.2) we deduce that $\bar{v} \in \mathrm{Sol}(P_0) \setminus \{0\}$.

We now show that there exists $\bar{x} \in \mathrm{Sol}(P)$ satisfying (4.1). Since $Ax^k \geq b$ for every $k \in N$, arguing similarly as in the proof of Theorem 2.1 we can find a subsequence $\{k'\} \subset \{k\}$ such that for each $i \in I$ the limit $\lim_{k' \to \infty} A_i x^{k'}$ exists (it may happen that $\lim_{k' \to \infty} A_i x^{k'} = +\infty$). Obviously,

$$\lim_{k' \to \infty} A_i x^{k'} \geq b_i \quad (\forall i \in I)$$

and

$$\lim_{k' \to \infty} C_j x^{k'} = d_j \quad (\forall j \in J).$$

Without restriction of generality we can assume that $\{k'\} \equiv \{k\}$. Let

$$I_0 = \{i \in I \,:\, \lim_{k \to \infty} A_i x^k = b_i\}, \quad I_1 = \{i \in I \,:\, \lim_{k \to \infty} A_i x^k > b_i\}.$$

It is clear that there exist $\varepsilon > 0$ and $k_0 \in N$ such that

$$A_i x^k \geq b_i + \varepsilon \quad (\forall i \in I_1,\ \forall k \geq k_0).$$

We have

$$A_i \bar{v} = \lim_{k \to \infty} A_i \frac{x^k}{\|x^k\|} = \lim_{k \to \infty} \frac{b_i}{\|x^k\|} = 0 \quad (\forall i \in I_0),$$

and

$$A_i \bar{v} = \lim_{k \to \infty} A_i \frac{x^k}{\|x^k\|} \geq \lim_{k \to \infty} \frac{b_i + \varepsilon}{\|x^k\|} = 0 \quad (\forall i \in I_1).$$

Let $x^k(t) = x^k - t\bar{v}$, where $t > 0$ and $k \geq k_0$. We have

$$A_i x^k(t) = A_i x^k - tA_i \bar{v} = A_i x^k \geq b_i \quad (\forall i \in I_0),$$

$$A_i x^k(t) = A_i x^k - tA_i \bar{v} \geq b_i + \varepsilon - tA_i \bar{v} \quad (\forall i \in I_1).$$

Fix an index $k \geq k_0$. From what has been said it follows that there exists $\delta > 0$ such that, for every $t \in (0, \delta)$,

$$A_i x^k(t) \geq b_i \quad (\forall i \in I = I_0 \cup I_1).$$

It is obvious that $C_j x^k(t) = d_j$ for all $j \in J$. Hence $x^k(t) \in \Delta$ for every $t \in (0, \delta)$. Since

$$\begin{aligned} 0 &\leq f(x^k(t)) - f(x^k) \\ &= \frac{1}{2}(x^k(t) - x^k)^T D(x^k(t) - x^k) + (Dx^k + c)^T (x^k(t) - x^k), \end{aligned}$$

we have

$$-\frac{1}{2}t\bar{v}^T D\bar{v} + (Dx^k + c)^T\bar{v} \leq 0.$$

Combining this with (4.2) we get

$$(Dx^k + c)^T\bar{v} \leq 0.$$

On the other hand, applying Corollary 2.6 we can assert that $(Dx^k + c)^T\bar{v} \geq 0$. Hence $(Dx^k + c)^T\bar{v} = 0$. Taking $\bar{x} = x^k$ we see that (4.1) is satisfied.

Let us prove that if there exist $\bar{x} \in \mathrm{Sol}(P)$ and $\bar{v} \in \mathrm{Sol}(P_0) \setminus \{0\}$ such that (4.1) holds, then $\omega := \{\bar{x} + t\bar{v} \,:\, t \geq 0\}$ is a solution ray of (P). For each $t > 0$, since $\bar{x} \in \mathrm{Sol}(P)$ and $\bar{v} \in \mathrm{Sol}(P_0)$, we have

$$A(\bar{x} + t\bar{v}) = A\bar{x} + tA\bar{v} \geq b,$$
$$C(\bar{x} + t\bar{v}) = C\bar{x} + tC\bar{v} = d.$$

Hence $\bar{x} + t\bar{v} \in \Delta$. Since $\bar{v} \in \mathrm{Sol}(P_0)$ and 0 is a feasible vector of (P_0), we have $\bar{v}^T D\bar{v} \leq 0$. If $\bar{v}^T D\bar{v} < 0$ then we check at once that (P_0) have no solutions, which is impossible. Thus $\bar{v}^T D\bar{v} = 0$. Combining this with (4.1) we deduce that

$$\begin{aligned} f(\bar{x} + t\bar{v}) &= \frac{1}{2}(\bar{x} + t\bar{v})^T D(\bar{x} + t\bar{v}) + c^T(\bar{x} + t\bar{v}) \\ &= (\frac{1}{2}\bar{x}^T D\bar{x} + c^T\bar{x}) + t(D\bar{x} + c)^T\bar{v} + \frac{1}{2}t^2\bar{v}^T D\bar{v} \\ &= f(\bar{x}). \end{aligned}$$

Since $\bar{x} \in \mathrm{Sol}(P)$, we conclude that $\bar{x} + t\bar{v} \in \mathrm{Sol}(P)$ for all $t \geq 0$. We have thus shown that if $\mathrm{Sol}(P)$ is unbounded then there exists $\bar{x} \in \mathrm{Sol}(P)$ and $\bar{v} \in \mathrm{Sol}(P_0) \setminus \{0\}$ satisfying (4.1) and $\omega = \{\bar{x} + t\bar{v} \,:\, t \geq 0\}$ is a solution ray of (P).

The claim that if (P) has a solution ray then $\mathrm{Sol}(P)$ is unbounded is obvious. The proof is complete. □

Theorem 4.2. *The set* $\mathrm{loc}(P)$ *is unbounded if and only if* (P) *has a local-solution ray.*

Proof. It suffices to prove that if $\mathrm{loc}(P)$ is unbounded then (P) has a local-solution ray. Suppose that there is a sequence $\{x^k\}$ in $\mathrm{loc}(P)$ satisfying the condition $\|x^k\| \to +\infty$. Let $\alpha \subset I$ be an index set. The set

$$F_\alpha = \{x \in R^n \,:\, A_\alpha x = b_\alpha,\ A_{I\setminus\alpha}x > b_{I\setminus\alpha},\ Cx = d\}$$

(which may be empty) is called a *pseudo-face* (see, for instance, Bank et al. (1982), p. 102) of Δ corresponding to α. Recall (Rockafellar (1970), p. 162) that a *face* of a convex set $X \subset R^n$ is a convex subset F of X such that every closed line segment in X with a relative interior point in F has both endpoints in F. In agreement with this definition, the sets F of the form

$$F = \{x \in R^n \,:\, A_\alpha x = b_\alpha,\ A_{I\setminus\alpha}x \geq b_{I\setminus\alpha},\ Cx = d\}$$

are the faces of the polyhedral convex set Δ under our consideration. Thus pseudo-faces are not faces in the sense of Rockafellar (1970). However, the closures of pseudo-faces are faces in that sense. It is clear that

$$\Delta = \bigcup\{F_\alpha \,:\, \alpha \subset I\}$$

and

$$F_\alpha \cap F_{\alpha'} = \emptyset \quad \text{whenever } \alpha \neq \alpha'.$$

It is a simple matter to show that for any $\alpha \subset I$ and for any $\bar{x} \in F_\alpha$ it holds

$$T_\Delta(\bar{x}) = \{v \in R^n \,:\, A_\alpha v \geq 0,\ Cv = 0\}.$$

Thus the tangent cone $T_\Delta(\bar{x})$ does not change when $\bar{x}$ varies inside a given pseudo-face F_α.

Since the number of pseudo-faces of Δ is finite, we conclude that there exist an index set $\alpha_* \subset I$ and a subsequence $\{k'\} \subset \{k\}$ such that $x^{k'} \in F_{\alpha_*}$ for every k'. There is no loss of generality in assuming that $\{k'\} \equiv \{k\}$.

We shall apply the construction due to Contesse which helped us to prove Theorem 3.4.

Since $x^k \in F_{\alpha_*}$ for all $k \in N$, we deduce that

$$T_\Delta(x^k) = \{v \in R^n \,:\, A_{\alpha_*}v \geq 0,\ Cv = 0\} \quad (\forall k \in N).$$

Let

$$M = \{v \in R^n \,:\, A_{\alpha_*}v = 0,\ Cv = 0\}.$$

Then M is a linear subspace and $M \subset T_\Delta(x^k)$. Let $M^\perp = \{v \in R^n \,:\, \langle v, u\rangle = 0 \text{ for every } u \in M\}$ and let

$$K = T_\Delta(x^k) \cap M^\perp = \mathrm{Pr}_{M^\perp}(T_\Delta(x^k)),$$

where $\mathrm{Pr}_{M^\perp}(\cdot)$ denotes the orthogonal projection of R^n onto the subspace $M^\perp$. We have

$$K = \{v \in R^n \,:\, A_{\alpha_*}v \geq 0,\ Cv = 0,\ v \in M^\perp\}$$

and $T_\Delta(x^k) = M + K$. Let

$$K_0^k = \{v \in K \,:\, \langle \nabla f(x^k), v\rangle = 0\}.$$

From the inclusion $x^k \in \mathrm{loc}(P)$ and from Theorem 3.5 it follows that $\langle \nabla f(x^k), u\rangle = 0$ for every $u \in M$ and $\langle \nabla f(x^k), v\rangle \geq 0$ for every $v \in K$. This implies that K_0^k is a face of K.

Since the polyhedral convex cone K has only a finite number of faces, there must exist a face K_0 of K and a subsequence $\{k_l\} \subset \{k\}$ such that

$$K_0^{k_l} = K_0 \quad \forall l \in N.$$

Consider the sequence of unit vectors $\{\dfrac{x^{k_l} - x^{k_1}}{\|x^{k_l} - x^{k_1}\|}\}$. Without loss of generality we can assume that

$$\frac{x^{k_l} - x^{k_1}}{\|x^{k_l} - x^{k_1}\|} \to \bar{z}, \quad \|\bar{z}\| = 1.$$

Set $\omega = \{x^{k_1} + t\bar{z} \,:\, t \geq 0\}$.

Claim 1. $\omega \subset S(P)$.

Let $x = x^{k_1} + t\bar{z}$, $t > 0$. For every $v \in M + K$ and $l \in N$, since $M + K = T_\Delta(x^{k_l})$ and $x^{k_l} \in \mathrm{loc}(P)$, by Theorem 3.5 we have

$$(Dx^{k_l} + c)^T v = \langle \nabla f(x^{k_l}), v\rangle \geq 0.$$

Hence

$$\left(D\frac{x^{k_l} - x^{k_1}}{\|x^{k_l} - x^{k_1}\|} + \frac{c + Dx^{k_1}}{\|x^{k_l} - x^{k_1}\|}\right)^T v \geq 0.$$

Letting $k \to \infty$ we deduce that

$$(D\bar{z})^T v \geq 0 \quad (\forall v \in M + K). \tag{4.3}$$

Since $M + K = T_\Delta(x^{k_1})$ and $x^{k_1} \in \mathrm{loc}(P)$, by Theorem 3.5 we have

$$(Dx^{k_1} + c)^T v = \langle \nabla f(x^{k_1}), v\rangle \geq 0 \quad (\forall v \in M + K). \tag{4.4}$$

Combining (4.4) with (4.3) gives

$$\langle \nabla f(x), v\rangle = (Dx + c)^T v = (Dx^{k_1} + c)^T v + t(D\bar{z})^T v \geq 0 \tag{4.5}$$

for all $v \in M + K$. We have $x \in F_{\alpha_*}$. Indeed, since $x^{k_l} \in F_{\alpha_*}$ for every $l \in N$, it follows that

$$A_i x = A_i(x^{k_1} + t\bar{z}) = A_i x^{k_1} + t \lim_{l\to\infty} \frac{A_i x^{k_l} - A_i x^{k_1}}{\|x^{k_l} - x^{k_1}\|} = b_i \quad (\forall i \in \alpha_*).$$

For every $i \in I$, we have

$$A_i \frac{x^{k_l} - x^{k_1}}{\|x^{k_l} - x^{k_1}\|} \geq \frac{b_i - A_i x^{k_1}}{\|x^{k_l} - x^{k_1}\|}.$$

Letting $l \to \infty$ yields

$$A_i \bar{z} \geq 0 \quad (\forall i \in I).$$

Consequently,

$$A_i x = A_i(x^{k_1} + t\bar{z}) = A_i x^{k_1} + tA_i \bar{z} > b_i \quad (\forall i \in I \setminus \alpha_*).$$

The equality $Cx = d$ can be established without any difficulty. From what has already been proved, we deduce that $x \in F_{\alpha_*}$. This implies that $T_\Delta(x) = M + K$. Hence from (4.5) it follows that

$$\langle \nabla f(x), v \rangle = (Dx + c)^T v \geq 0 \quad (\forall v \in T_\Delta(x)).$$

This shows that $x \in S(P)$. (Recall that property (i) in Theorem 3.5 is equivalent to the existence of a pair $(\bar{\lambda}, \bar{\mu}) \in R^m \times R^s$ satisfying system (3.11).)

CLAIM 2. $\omega \subset \text{loc}(P)$.

Let $x = x^{k_1} + t\bar{z}$, $t > 0$. By Claim 1, $x \in S(P)$, that is

$$\langle \nabla f(x), v \rangle = (Dx + c)^T v \geq 0 \quad (\forall v \in T_\Delta(x)). \tag{4.6}$$

We want to show that

$$K \cap (\nabla f(x))^\perp = K_0. \tag{4.7}$$

For each $l \in N$, we have

$$K_0 = \{v \in K : \langle \nabla f(x^{k_l}), v \rangle = 0\}.$$

So, for every $v \in K_0$, it holds $(Dx^{k_l} + c)^T v = \langle \nabla f(x^{k_l}), v \rangle = 0$ for all $l \in N$. Therefore

$$\left(D \frac{x^{k_l} - x^{k_1}}{\|x^{k_l} - x^{k_1}\|} + \frac{c + Dx^{k_1}}{\|x^{k_l} - x^{k_1}\|} \right)^T v = 0.$$

Letting $k \to \infty$ we deduce that

$$(D\bar{z})^T v = 0 \quad (\forall v \in K_0).$$

Hence

$$\langle \nabla f(x), v\rangle = \langle \nabla f(x^{k_1}), v\rangle + t(D\bar{z})^T v = 0 \quad (\forall v \in K_0).$$

This shows that $K_0 \subset K \cap (\nabla f(x))^{\perp}$. To prove the reverse inclusion, let us fix any $v \in K \cap (\nabla f(x))^{\perp}$. On account of (4.3) and (4.4), we have

$$0 = \langle \nabla f(x), v\rangle = \langle \nabla f(x^{k_1}), v\rangle + t(D\bar{z})^T v,$$

$$(D\bar{z})^T v \geq 0, \quad \langle \nabla f(x^{k_1}), v\rangle \geq 0.$$

This clearly forces $\langle \nabla f(x^{k_1}), v\rangle = 0$. So $v \in K_0$ whenever $v \in K \cap (\nabla f(x))^{\perp}$. The equality (4.7) has been established. We now show that

$$v^T D v \geq 0 \quad \forall v \in T_{\Delta}(x) \cap (\nabla f(x))^{\perp}. \tag{4.8}$$

In the proof of Claim 1, it has been shown that $T_{\Delta}(x) = M + K = T_{\Delta}(x^{k_1})$. Besides, from (4.6) we deduce that $M \subset (\nabla f(x))^{\perp}$. Hence, from (4.7) and the construction of the sequence $\{x^{k_l}\}$ it follows that

$$T_{\Delta}(x) \cap (\nabla f(x))^{\perp} = M + K_0 = T_{\Delta}(x^{k_1}) \cap (\nabla f(x^{k_1}))^{\perp}. \tag{4.9}$$

Since $x^{k_1} \in \mathrm{loc}(P)$, Theorem 3.5 shows that

$$v^T D v \geq 0 \quad \forall v \in T_{\Delta}(x^{k_1}) \cap (\nabla f(x^{k_1})^{\perp}.$$

Combining this with (4.9) we get (4.8). From (4.6), (4.8) and Theorem 3.5, we deduce that $x \in \mathrm{loc}(P)$. This completes the proof of Claim 2 and the proof of our theorem. □

Theorem 4.3. *The set $S(P)$ is unbounded if and only if (P) has a KKT point ray.*

Proof. By definition, $x \in S(P)$ if and only if there exists $(\lambda, \mu) \in R^m \times R^s$ such that

$$\begin{cases} Dx - A^T\lambda - C^T\mu + c = 0, \\ Ax \geq b, \quad Cx = d, \\ \lambda \geq 0, \quad \lambda^T(Ax - b) = 0. \end{cases} \tag{4.10}$$

Given a point $x \in S(P)$, we set $I_0 = \{i \in I : A_i x = b_i\}$, $I_1 = I \setminus I_0 = \{i \in I : A_i x > b_i\}$. From the last equality in (4.10) we get

$$\lambda_i = 0 \quad \forall i \in I_1.$$

Hence (x, λ, μ) satisfies the following system

$$\begin{cases} Dx - A^T\lambda - C^T\mu + c = 0, \\ A_{I_0}x = b_{I_0}, \quad \lambda_{I_0} \geq 0, \\ A_{I_1}x \geq b_{I_1}, \quad \lambda_{I_1} = 0, \\ Cx = d. \end{cases} \tag{4.11}$$

Fix any $I_0 \subset I$ and denote by Q_{I_0} the set of all (x, λ, μ) satisfying (4.11). It is obvious that Q_{I_0} is a polyhedral convex set. From what has been said it follows that

$$S(P) = \bigcup \{\mathrm{Pr}_{R^n}(Q_{I_0}) \, : \, I_0 \subset I\}, \tag{4.12}$$

where $\mathrm{Pr}_{R^n}(x, \lambda, \mu) := x$. Since $\mathrm{Pr}_{R^n}(\cdot) : R^n \times R^m \times R^s \to R^n$ is a linear operator, $\mathrm{Pr}_{R^n}(Q_{I_0})$ is a polyhedral convex set for every $I_0 \subset I$. Indeed, as Q_{I_0} is a polyhedral convex set, it is *finitely generated*, i.e., there exist vectors $z^1, \ldots, z^k, w^1, \ldots, w^l$ in $R^n \times R^m \times R^s$ such that

$$Q_{I_0} = \{z = \textstyle\sum_{i=1}^k t_i z^i + \sum_{j=1}^l \theta_j w^j \, : \, t_i \geq 0 \text{ for all } i, \\ \theta_j \geq 0 \text{ for all } j, \text{ and } \sum_{i=1}^k t_i = 1\}$$

(see Rockafellar (1970), Theorem 19.1). Then, by the linearity of the operator $\mathrm{Pr}_{R^n}(\cdot)$, we have

$$\mathrm{Pr}_{R^n}(Q_{I_0}) = \{x = \textstyle\sum_{i=1}^k t_i x^i + \sum_{j=1}^l \theta_j v^j \, : \, t_i \geq 0 \text{ for all } i, \\ \theta_j \geq 0 \text{ for all } j, \text{ and } \sum_{i=1}^k t_i = 1\},$$

where $x^i = \mathrm{Pr}_{R^n}(z^i)$ for all i and $v^i = \mathrm{Pr}_{R^n}(w^j)$ for all j. This shows that the set $\mathrm{Pr}_{R^n}(Q_{I_0})$ is finitely generated, hence it is a polyhedral convex set (see Rockafellar (1970), Theorem 19.1).

If $S(P)$ is unbounded then from (4.12) it follows that there exists an index set $I_0 \subset I$ such that $\Omega_{I_0} := \mathrm{Pr}_{R^n}(Q_{I_0})$ is an unbounded set. Since Ω_{I_0} is a polyhedral convex set, it is an unbounded closed convex set. By Theorem 8.4 in Rockafellar (1970), Ω_{I_0} admits a direction of recession; that is there exists $\bar{v} \in R^n \setminus \{0\}$ such that

$$x + t\bar{v} \in \Omega_{I_0} \quad \forall x \in \Omega_{I_0}, \ \forall t \geq 0. \tag{4.13}$$

Taking any $\bar{x} \in \Omega_{I_0}$ we deduce from (4.12) and (4.13) that $\bar{x} + t\bar{v} \in S(P)$ for all $t \geq 0$. Thus we have proved that (P) admits a KKT point ray. Conversely, it is obvious that if (P) admits a KKT point ray then $S(P)$ is unbounded. $\square$

Remark 4.1. Formula (4.12) shows that $S(P)$ is a union of finitely many polyhedral convex sets.

Let us derive another formula for the KKT point set of (P). For each index set $\alpha \subset I$, denote by F_α the pseudo-face of Δ corresponding to α; that is

$$F_\alpha = \{x \in R^n : A_\alpha x = b_\alpha,\ A_{I\setminus\alpha} x > b_{I\setminus\alpha},\ Cx = d\}.$$

Since $\Delta = \cup\{F_\alpha : \alpha \subset I\}$, we deduce that

$$S(P) = \bigcup\{S(P) \cap F_\alpha : \alpha \subset I\}. \tag{4.14}$$

Lemma 4.1. *For every $\alpha \subset I$, $S(P) \cap F_\alpha$ is a convex set.*

Proof. Let $\alpha \subset I$ be any index set. From the definition of $S(P)$ it follows that $x \in S(P) \cap F_\alpha$ if and only if there exist $(\lambda, \mu) \in R^m \times R^s$ such that

$$\begin{cases} Dx - A^T\lambda - C^T\mu + c = 0, \\ A_\alpha x = b_\alpha,\ \ \lambda_\alpha \geq 0, \\ A_{I\setminus\alpha} x > b_{I\setminus\alpha},\ \ \lambda_{I\setminus\alpha} = 0, \\ Cx = d. \end{cases} \tag{4.15}$$

Let Z_α denote the set of all the points $(x, \lambda, \mu) \in R^n \times R^m \times R^s$ satisfying the system (4.15). It is clear that Z_α is a convex set. From what has already been said it follows that $S(P) \cap F_\alpha = \mathrm{Pr}_{R^n}(Z_\alpha)$, where $\mathrm{Pr}_{R^n}(x, \lambda, \mu) := x$. Since $\mathrm{Pr}_{R^n}(\cdot)$ is a linear operator, we conclude that $S(P) \cap F_\alpha$ is a convex set. $\square$

Note that, in general, the convex sets $S(P) \cap F_\alpha$, $\alpha \subset I$, in the representation (4.14) may not be closed.

We know that $\mathrm{Sol}(D, A, c, b) \subset \mathrm{loc}(D, A, c, b) \subset S(D, A, c, b)$ (see (3.10)). We have characterized the unboundedness of these solution sets. The following questions arise:

QUESTION 1: Is it true that $\mathrm{Sol}(P)$ is unbounded whenever $\mathrm{loc}(P)$ is unbounded?

QUESTION 2: Is it true that $\mathrm{loc}(P)$ is unbounded whenever $S(P)$ is unbounded?

The following example gives a negative answer to Question 1.

Example 4.1. Consider the problem

$$(P_1) \quad \min\{f(x) = -x_2^2 + 2x_2 : x = (x_1, x_2),\ x_1 \geq 0,\ x_2 \geq 0\}.$$

Denote by Δ the feasible region of (P_1). We have

$$\begin{aligned}&\mathrm{Sol}(P_1) = \emptyset, \quad \mathrm{loc}(P_1) = \{x \in R^2 \,:\, x_1 \geq 0, x_2 = 0\},\\ &S(P_1) = \{x \in R^2 \,:\, x_1 \geq 0, x_2 = 0 \text{ or } x_2 = 1\}.\end{aligned}$$

Thus $\mathrm{loc}(P_1)$ is unbounded, but $\mathrm{Sol}(P_1) = \emptyset$. In order to establish the above results, one can argue as follows. Since $I = \{1,2\}$, the constraint set of (P_1) is composed by four pseudo-faces:

$$\begin{aligned}&F_{\{1,2\}} = \{(0,0)\},\\ &F_{\{1\}} = \{x \in R^2 \,:\, x_1 = 0, x_2 > 0\}\\ &F_{\{2\}} = \{x \in R^2 \,:\, x_2 = 0, x_1 > 0\},\\ &F_{\emptyset} = \{x \in R^2 \,:\, x_1 > 0, x_2 > 0\}.\end{aligned}$$

Since $\nabla f(x) = (0, -2(x_2 - 1))$, by solving four KKT systems of the form (4.15) where C, d are vacuous,

$$D = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

we obtain

$$\begin{aligned}&S(P_1) \cap F_{\{1,2\}} = \{(0,0)\},\\ &S(P_1) \cap F_{\{1\}} = \{(0,1)\},\\ &S(P_1) \cap F_{\{2\}} = \{x \in R^2 \,:\, x_2 = 0, x_1 > 0\},\\ &S(P_1) \cap F_{\emptyset} = \{x \in R^2 \,:\, x_1 > 0, x_2 = 1\}.\end{aligned}$$

From formula (4.14) it follows that

$$S(P_1) = \{x \in R^2 \,:\, x_1 \geq 0, \ x_2 = 0 \ \text{ or } \ x_2 = 1\}.$$

Since $\lim_{x_2 \to +\infty} f(0, x_2) = -\infty$ and, for each $x_2 \geq 0$, $x = (0, x_2)$ is a feasible vector for (P_1), we conclude that $\mathrm{Sol}(P_1) = \emptyset$. For any $x = (x_1, 1) \in S(P_1) \cap F_{\emptyset}$, we have $T_{\Delta}(x) = R^2$, $(\nabla f(x))^{\perp} = R^2$. Then the condition

$$v^T D v \geq 0 \quad \forall v \in T_{\Delta}(x) \cap (\nabla f(x))^{\perp}, \tag{4.16}$$

which is equivalent to the condition

$$-2v_2^2 \geq 0 \quad \forall v_2 \in R,$$

cannot be satisfied. By Theorem 3.5, $x \notin \mathrm{loc}(P_1)$. Now, for any $x = (x_1, 0) \in S(P_1) \cap F_{\{2\}}$, we have

$$T_{\Delta}(x) = \{v = (v_1, v_2) \in R^2 \,:\, v_2 \geq 0\},$$

$$(\nabla f(x))^{\perp} = \{v = (v_1, v_2) \in R^2 \, : \, v_2 = 0\}.$$

Condition (4.16), which is now equivalent to the requirement

$$-2v_2^2 \geq 0 \quad \text{for } v_2 = 0,$$

is satisfied. Applying Theorem 3.5 we can assert that $x \in \text{loc}(P_1)$. In the same manner we can see that the unique point $x = (0, 1)$ of the set $S(P_1) \cap F_{\{1\}}$ does not belong to $\text{loc}(P_1)$, while the unique point $x = (0, 0)$ of the set $S(P_1) \cap F_{\{1,2\}}$ belongs to $\text{loc}(P_1)$. Thus we have shown that $\text{loc}(P_1) = \{x \in R^2 \, : \, x_1 \geq 0, \;\; x_2 = 0\}$.

The following example gives a negative answer to Question 2.

Example 4.2. Consider the problem

$$(P_2) \qquad \min\{f(x) = -x_2^2 \, : \, x = (x_1, x_2), \; x_1 \geq 0, \; x_2 \geq 0\}.$$

Analysis similar to that in the preceding example shows that

$$\text{Sol}(P_2) = \text{loc}(P_2) = \emptyset, \quad S(P_2) = \{x = (x_1, x_2) \, : \, x_1 \geq 0, \; x_2 = 0\}.$$

4.2 Closedness of the Solution Sets

Since $S(P)$ is a union of finitely many polyhedral convex set (see (4.12)), it is a closed set. The set $\text{Sol}(P)$ is also closed. Indeed, we have

$$\text{Sol}(P) = \{x \in \Delta \, : \, f(x) = v(P)\},$$

where $v(P) = \inf\{f(x) \, : \, x \in \Delta\}$. If $v(P)$ is finite then from the closedness of Δ, the continuity of f, and the above formula, it follows that $\text{Sol}(P)$ is closed. If $v(P) = +\infty$ then $\Delta = \emptyset$, hence $\text{Sol}(P) = \emptyset$. If $v(P) = -\infty$ then it is obvious that $\text{Sol}(P) = \emptyset$. Thus we conclude that $\text{Sol}(P)$ is always a closed set.

The following question arises:

QUESTION 3: Is it true that $\text{loc}(P)$ is always a closed set?

The following example gives a negative answer to Question 3.

Example 4.3. Consider the problem

$$\min\{f(x) = -x_2^2 + x_1 x_2 \, : \, x = (x_1, x_2), \; x_1 \geq 0, \; x_2 \geq 0\}. \qquad (P_3)$$

Analysis similar to that in Example 4.1 shows that

$$\text{Sol}(P_3) = \emptyset, \quad \text{loc}(P_3) = \{x = (x_1, x_2) \, : \, x_1 > 0, \; x_2 = 0\},$$
$$S(P_3) = \{x = (x_1, x_2) \, : \, x_1 \geq 0, \; x_2 = 0\}.$$

Our next aim is to study the situation where $\text{Sol}(P)$ (resp., $\text{loc}(P)$, $S(P)$) is a bounded set having infinitely many elements.

4.3 A Property of the Bounded Infinite Solution Sets

Definition 4.2. A line segment $\omega_\delta = \{\bar{x} + t\bar{v} \,:\, t \in [0, \delta)\}$, where $\bar{v} \in R^n \setminus \{0\}$ and $\delta > 0$, which is a subset of $\mathrm{Sol}(P)$ (resp., $\mathrm{loc}(P)$, $S(P)$), is called a *solution interval* (resp., a *local-solution interval*, a *KKT point interval*) of (P).

Proposition 4.1. *If the set* $\mathrm{Sol}(P)$ *is bounded and infinite, then* (P) *has a solution interval.*

Proof. For each index set $\alpha \subset I$, denote by F_α the pseudo-face of Δ corresponding to α. As $\mathrm{Sol}(P) \subset \Delta$ has infinitely many elements and $\Delta = \cup\{F_\alpha \,:\, \alpha \subset I\}$, there must exists some $\alpha_* \subset I$ such that the intersection $\mathrm{Sol}(P) \cap F_{\alpha_*}$ has infinitely many elements. For each $x \in \mathrm{Sol}(P) \cap F_{\alpha_*}$ we have $T_\Delta(x) = \{v \in R^n \,:\, A_{\alpha_*}v \geq 0,\ Cv = 0\}$ and, by Theorem 3.5, $\langle \nabla f(x), v\rangle \geq 0$ for every $v \in T_\Delta(x)$. Hence $T := T_\Delta(x)$ is a constant polyhedral convex cone which does not depend on the position of x in the pseudo-face F_{α_*} of Δ, and

$$T_0^x := \{v \in T \,:\, \langle \nabla f(x), v\rangle = 0\}$$

is a face of T. Since the number of faces of T is finite, from what has already been said it follows that there must exist two different points x and y of $\mathrm{Sol}(P) \cap F_{\alpha_*}$ such that $T_0^x = T_0^y$. Set $T_0 = T_0^x$. By Lemma 4.1, the intersection $S(P) \cap F_{\alpha_*}$ is convex. Since $x \in \mathrm{Sol}(P)$, $y \in \mathrm{Sol}(P)$ and $\mathrm{Sol}(P) \subset S(P)$, it follows that $z_t := (1-t)x + ty$ belongs to $S(P) \cap F_{\alpha_*}$ for every $t \in [0, 1]$. By the remark following Theorem 3.5, for every $t \in (0, 1)$, we have

$$\langle \nabla f(z_t), v\rangle \geq 0 \quad \forall v \in T.$$

Therefore

$$\begin{aligned}
&T_\Delta(z_t) \cap (\nabla f(z_t))^\perp \\
&= \{v \in T \,:\, \langle \nabla f(z_t), v\rangle = 0\} \\
&= \{v \in T \,:\, (1-t)\underbrace{\langle \nabla f(x), v\rangle}_{\geq 0} + t\underbrace{\langle \nabla f(y), v\rangle}_{\geq 0} = 0\} \\
&= \{v \in T \,:\, \langle \nabla f(x), v\rangle = 0,\ \langle \nabla f(y), v\rangle = 0\} \\
&= T_0.
\end{aligned}$$

Since $x \in \mathrm{Sol}(P)$ and $T_\Delta(x) \cap (\nabla f(x))^\perp = T_0$, from Theorem 3.5 it follows that

$$v^T D v \geq 0 \quad \forall v \in T_0.$$

We have shown that, for every $t \in (0,1)$, $z_t \in S(P)$ and

$$T_\Delta(z_t) \cap (\nabla f(z_t))^\perp = T_0.$$

Combining these facts and applying Theorem 3.5 we deduce that $z_t \in \text{loc}(P)$. Consider the function $\varphi : [0,1] \to R$ defined by setting $\varphi(t) = f(z_t)$ for all $t \in [0,1]$. It is clear that φ is a continuous function which is differentiable at each $t \in (0,1)$. Since z_t is a local solution of (P), φ attains a local minimum at every $t \in (0,1)$. Hence $\varphi'(t) = 0$ for every $t \in (0,1)$. Consequently, $\varphi(t)$ is a constant function. Since $x \in \text{Sol}(P)$, we see that $z_t \in \text{Sol}(P)$ for all $t \in [0,1]$. Thus $[x,y) := \{(1-t)x + ty \,:\, t \in [0,1)\}$ is a solution interval of (P). □

Proposition 4.2. (See Phu and Yen (2001), Theorem 3) *If the set* $\text{loc}(P)$ *is bounded and infinite, then* (P) *has a local-solution interval.*

Proof. For each index set $\alpha \subset I$, denote by F_α the pseudo-face of Δ corresponding to α. As $\text{loc}(P) \subset \Delta$ is an infinite set and $\Delta = \cup\{F_\alpha \,:\, \alpha \subset I\}$, there must exists some $\alpha_* \subset I$ such that the intersection $\text{loc}(P) \cap F_{\alpha_*}$ has infinitely many elements. For each $x \in \text{loc}(P) \cap F_{\alpha_*}$ we have $T_\Delta(x) = \{v \in R^n \,:\, A_{\alpha *}v \geq 0,\ Cv = 0\}$ and, by Theorem 3.5, $\langle \nabla f(x), v\rangle \geq 0$ for every $v \in T_\Delta(x)$. Hence $T := T_\Delta(x)$ is a constant polyhedral convex cone which does not depend on the position of x in the pseudo-face F_{α_*} of Δ, and

$$T_0^x := \{v \in T \,:\, \langle \nabla f(x), v\rangle = 0\}$$

is a face of T. Since the number of faces of T is finite, it follows that there must exist two different point x and y of $\text{loc}(P) \cap F_{\alpha_*}$ such that $T_0^x = T_0^y$. Set $T_0 = T_0^x$. By Lemma 4.1, the intersection $S(P) \cap F_{\alpha_*}$ is convex. Since $x \in \text{loc}(P)$, $y \in \text{loc}(P)$ and $\text{loc}(P) \subset S(P)$, it follows that $z_t := (1-t)x + ty$ belongs to $S(P) \cap F_{\alpha_*}$ for every $t \in [0,1]$. According to the remark following Theorem 3.5, for every $t \in (0,1)$ we have

$$\langle \nabla f(z_t), v\rangle \geq 0 \quad \forall v \in T.$$

As in the proof of Proposition 4.1, we have

$$T_\Delta(z_t) \cap (\nabla f(z_t))^\perp = T_0.$$

Since $x \in \text{loc}(P)$ and $T_\Delta(x) \cap (\nabla f(x))^\perp = T_0$, from Theorem 3.5 it follows that

$$v^T D v \geq 0 \quad \forall v \in T_0.$$

Therefore, for every $t \in (0,1)$, $z_t \in S(P)$ and

$$T_\Delta(z_t) \cap (\nabla f(z_t))^\perp = T_0.$$

On account of these facts and of Theorem 3.5, we conclude that z_t is a local solution of (P). Thus $[x, y)$ is a local-solution interval of (P). □

Proposition 4.3. *If the set $S(P)$ is bounded and infinite, then (P) has a KKT point interval.*

Proof. For each index set $\alpha \subset I$, denote by F_α the pseudo-face of Δ corresponding to α. As $S(P) \subset \Delta$ is an infinite set and $\Delta = \cup\{F_\alpha : \alpha \subset I\}$, there must exists some $\alpha_* \subset I$ such that the intersection $S(P) \cap F_{\alpha*}$ has infinitely many elements. Hence there must exist two different point x and y of $S(P) \cap F_{\alpha_*}$. By Lemma 4.1, the intersection $S(P) \cap F_{\alpha*}$ is convex. This implies that $[x, y)$ is a KKT point interval of (P). □

4.4 Finiteness of the Solution Sets

Theorem 4.4. *The following assertions are valid:*

(i) *If D is a positive definite matrix and Δ is nonempty, then (P) has a unique solution and it holds* $\mathrm{Sol}(P) = \mathrm{loc}(P) = S(P)$.

(ii) *If D is a negative definite matrix then each local solution of (P) is an extreme point of Δ. In this case,* $\mathrm{Sol}(P) \subset \mathrm{loc}(P) \subset \mathrm{extr}\Delta$. *Hence, if D is a negative definite matrix then the number of solutions of (P) (resp., the number of local solutions of (P)) is always less than or equal to the number of extreme points of Δ. Besides, if* $\mathrm{Sol}(P)$ *is nonempty then Δ is a compact polyhedral convex set.*

(iii) *If D is a positive semidefinite matrix then* $\mathrm{Sol}(P)$ *is a closed convex set and it holds* $\mathrm{Sol}(P) = \mathrm{loc}(P) = S(P)$. *Hence, if D is a positive semidefinite matrix then* $\mathrm{Sol}(P)$ *is finite if and only if it is a singleton or it is empty.*

Proof. (i) Suppose that the symmetric matrix D is positive definite and the set $\Delta := \{x \in R^n : Ax \geq b, Cx = d\}$ is nonempty. Setting

$$\varrho := \inf\{v^T Dv : v \in R^n, \|v\| = 1\} > 0$$

we deduce that $x^TDx \geq \varrho\|x\|^2$ for every $x \in R^n$. Fix any $x^0 \in \Delta$. Note that

$$\begin{aligned} f(x) - f(x_0) &= \frac{1}{2}(x - x^0)^T D(x - x^0) + (Dx^0 + c)^T(x - x^0) \\ &\geq \frac{1}{2}\varrho\|x - x^0\|^2 - \|Dx^0 + c\|\|x - x^0\|. \end{aligned}$$

The last expression tends to $+\infty$ as $\|x - x^0\| \to +\infty$. Hence there exists $\gamma > 0$ such that

$$f(x) - f(x^0) \geq 1 \quad \forall x \in \Delta \setminus \bar{B}(x^0, \gamma). \tag{4.17}$$

From (4.17) it follows that (P) cannot have solutions in $\Delta\setminus\bar{B}(x^0,\gamma)$. Since $\Delta \cap \bar{B}(x^0, \gamma) \neq \emptyset$ the problem $\min\{f(x) : x \in \Delta \cap \bar{B}(x^0, \gamma)\}$ possesses a solution $\bar{x}$. By (4.17), $\bar{x} \in \mathrm{Sol}(P)$. Assume, contrary to our claim, that there are two different solutions $\bar{x}$ and $\bar{y}$ of (P). Since $\bar{y} - \bar{x} \in T_\Delta(\bar{x})$ and $\bar{x} \in \mathrm{Sol}(P)$, by Theorem 3.1 we have $(D\bar{x} + c)^T(\bar{y} - \bar{x}) \geq 0$. As $\bar{y} \neq \bar{x}$ and D is positive definite, we have $(\bar{y} - \bar{x})^T D(\bar{y} - \bar{x}) > 0$. It follows that

$$0 = f(\bar{y}) - f(\bar{x}) = \frac{1}{2}(\bar{y} - \bar{x})^T D(\bar{y} - \bar{x}) + (D\bar{x} + c)^T(\bar{y} - \bar{x}) > 0,$$

a contradiction. The equalities $\mathrm{Sol}(P) = \mathrm{loc}(P) = S(P)$ follow from the fact that, under our assumptions, f is a convex function (see Proposition 1.2).

(ii) Let D be a negative definite matrix and $\bar{x} \in \mathrm{loc}(P)$. If $\bar{x} \notin \mathrm{extr}\Delta$ then there exist $x \in \Delta$, $y \in \Delta$, $x \neq y$, and $t \in (0,1)$ such that $\bar{x} = (1-t)x + ty$. Since $\bar{x} \in \Delta$, $x - \bar{x} \in T_\Delta(\bar{x})$, $y - \bar{x} \in T_\Delta(\bar{x})$, and $y - \bar{x} = -\dfrac{1-t}{t}(x - \bar{x})$, applying Theorem 3.5 we get $\langle \nabla f(\bar{x}), x - \bar{x}\rangle \geq 0$ and

$$-\frac{1-t}{t}\langle \nabla f(\bar{x}), x - \bar{x}\rangle = \langle \nabla f(\bar{x}), y - \bar{x}\rangle \geq 0.$$

Therefore $\langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0$. This equality and the assumption $\bar{x} \in \mathrm{loc}(P)$ allows us to apply Theorem 3.5 to obtain $(x-\bar{x})^T D(x - \bar{x}) \geq 0$. This contradicts the fact that matrix D is negative definite. We have thus proved that $\mathrm{loc}(P) \subset \mathrm{extr}\Delta$. Consequently, $\mathrm{Sol}(P) \subset \mathrm{loc}(P) \subset \mathrm{extr}\Delta$. We now suppose that $\mathrm{Sol}(P) \neq \emptyset$. By Corollary 2.6,

$$(v \in R^n,\ Av \geq 0,\ Cv = 0) \Longrightarrow v^T Dv \geq 0.$$

Combining this with the negative definiteness of D we conclude that

$$\{v \in R^n : Av \geq 0,\ Cv = 0\} = \{0\}.$$

Hence Δ has no directions of recession. By Theorem 8.4 in Rockafellar (1970), Δ is a compact set.

(iii) Let D be a positive semidefinite matrix. By Proposition 1.2, f is a convex function. Hence $\mathrm{Sol}(P)$ is a closed convex set and it holds $\mathrm{Sol}(P) = \mathrm{loc}(P) = S(P)$. □

Example 4.4. Consider problem (P) of the following form

$$\min\{f(x) = -x_1^2 - x_2^2 + 1 : x = (x_1, x_2),\ x_1 \geq -1,\\ -x_1 \geq -1,\ x_2 \geq -1,\ -x_2 \geq -1\}.$$

The matrix $D = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ corresponding to this problem is negative definite. It is easily seen that

$$\mathrm{Sol}(P) = \mathrm{loc}(P) = \{(1,1), (1,-1), (-1,1), (-1,-1)\},$$

$$S(P) = \mathrm{Sol}(P) \cup \{(0,0), (1,0), (0,1), (-1,0), (0,-1)\}.$$

Theorem 4.5. *If* $\mathrm{Sol}(P)$ *(resp.,* $S(P)$, $\mathrm{loc}(P)$*) is a finite set, then each pseudo-face of* Δ *cannot contain more than one element of* $\mathrm{Sol}(P)$ *(resp.,* $S(P)$, $\mathrm{loc}(P)$*). Hence, if* $\mathrm{Sol}(P)$ *(resp.,* $S(P)$, $\mathrm{loc}(P)$*) is a finite set, then* $\mathrm{Sol}(P)$ *(resp.,* $S(P)$, $\mathrm{loc}(P)$*) cannot have more than* 2^m *elements, where* m *is the number of inequality constraints of* (P).

Before proving this theorem let us establish the following two auxiliary results.

Proposition 4.4. *Assume that* $x \in \mathrm{loc}(P) \cap F_\alpha$, $y \in S(P) \cap F_\alpha$, $y \neq x$, *where* F_α *is a pseudo-face of* Δ. *Then there exists* $\delta > 0$ *such that, for every* $t \in (0, \delta)$, $a(t) := (1-t)x + ty$ *is a local solution of* (P).

Proof. Let $x \in \mathrm{loc}(P) \cap F_\alpha$, $y \in S(P) \cap F_\alpha$, $y \neq x$, where $\alpha \subset I = \{1, \ldots, m\}$ and $F_\alpha = \{x \in R^n : A_\alpha x = b_\alpha,\ A_{I\setminus\alpha} x > b_{I\setminus\alpha},\ Cx = d\}$. Since $x \in F_\alpha$, we have $T_\Delta(x) = \{v \in R^n : A_\alpha v \geq 0,\ Cv = 0\}$. Let

$$M = \{v \in R^n : A_\alpha v = 0,\ Cv = 0\}.$$

Then M is a linear subspace and $M \subset T_\Delta(x)$. Let $M^\perp = \{v \in R^n : \langle v, u\rangle = 0$ for every $u \in M\}$ and let

$$K = T_\Delta(x) \cap M^\perp = \mathrm{Pr}_{M^\perp}(T_\Delta(x)),$$

where $\mathrm{Pr}_{M^\perp}(\cdot)$ denotes the orthogonal projection of R^n onto the subspace $M^\perp$. We have

$$K = \{v \in R^n \,:\, A_\alpha v \geq 0,\ Cv = 0,\ v \in M^\perp\}$$

and $T_\Delta(x) = M + K$. Since K is a pointed polyhedral convex cone (see the proof of Theorem 3.4), according to Theorem 19.1 in Rockafellar (1970), there exists a finite system of generators $\{z^1, \ldots, z^q\}$ of K. By convention, if $K = \{0\}$ then the system is vacuous. In the case where that system is not vacuous, we have

$$K = \{v = \sum_{j=1}^{q} t_j z^j \,:\, t_j \geq 0 \text{ for all } j = 1, \ldots, q\}.$$

Let $Q = \{1, \ldots, q\}$,

$$Q_0 = \{j \in Q \,:\, \langle \nabla f(x), z^j \rangle = 0\},\ Q_1 = \{j \in Q \,:\, \langle \nabla f(x), z^j \rangle > 0\}. \tag{4.18}$$

Since $x \in \mathrm{loc}(P)$, we must have

$$\langle \nabla f(x), v \rangle \geq 0 \quad \forall v \in T_\Delta(x).$$

From this we deduce that $Q = Q_0 \cup Q_1$. For every $a \in F_\alpha$, let $K_0^a = \{v \in K \,:\, \langle \nabla f(a), v \rangle = 0\}$. For every $t \in [0, 1]$, we set $a(t) = (1-t)x + ty$. Since $x \in F_\alpha$ and $y \in F_\alpha$, it follows that $a(t) \in F_\alpha$ for every $t \in [0, 1]$. Consequently,

$$T_\Delta(a(t)) = T_\Delta(x) = \{v \in R^n \,:\, A_\alpha v \geq 0,\ Cv = 0\} = M + K. \tag{4.19}$$

It follows from (4.18) that there exists $\delta \in (0, 1)$ such that

$$\langle \nabla f(a(t)), z^j \rangle > 0 \quad \forall t \in (0, \delta),\ \forall j \in Q_1. \tag{4.20}$$

For any $t \in (0, \delta)$, by Lemma 4.1 we have $a(t) \in S(P)$. Therefore $\langle \nabla f(a(t)), v \rangle \geq 0$ for every $v \in T_\Delta(a(t))$. Combining this with (4.20) we deduce that

$$K_0^{a(t)} \subset K_0^x = \{v = \sum_{j \in Q_0} t_j z^j \,:\, t_j \geq 0 \text{ for all } j \in Q_0\}. \tag{4.21}$$

We claim that $a(t) \in \mathrm{loc}(P)$ for every $t \in (0, \delta)$. Indeed, since $\langle \nabla f(a(t)), v \rangle \geq 0$ for every $v \in T_\Delta(a(t))$, we get

$$\langle \nabla f(a(t)), v \rangle = 0 \quad \forall v \in M,\ \forall t \in (0, \delta).$$

By (4.19) and (4.21),

$$T_\Delta(a(t)) \cap (\nabla f(a(t))^\perp = M + K_0^{a(t)} \subset M + K_0^x. \tag{4.22}$$

As $x \in \text{loc}(P)$, by Theorem 3.5 we have

$$v^T Dv \geq 0 \quad \forall v \in T_\Delta(x) \cap (\nabla f(x))^\perp = M + K_0^x.$$

Combining this with (4.22) yields

$$v^T Dv \geq 0 \quad \forall v \in T_\Delta(a(t)) \cap (\nabla f(a(t)))^\perp = M + K_0^{a(t)}.$$

Since $a(t) \in S(P)$, from the last fact and Theorem 3.5 we conclude that $a(t) \in \text{loc}(P)$. □

Proposition 4.5. *Assume that x and y are two different Karush-Kuhn-Tucker points of (P) belonging to the same pseudo-face of Δ. Then the function $\varphi(t) := f((1-t)x + ty)$ is constant on $[0,1]$.*

Proof. Let $\alpha \subset I$, $x \in S(P) \cap F_\alpha$, $y \in S(P) \cap F_\alpha$, $x \neq y$. For every $t \in [0,1]$, we define $a(t) = (1-t)x + ty$. Since $a(t) \in F_\alpha$, it follows that $T_\Delta(a(t)) = \{v \in R^n : A_\alpha v \geq 0,\ Cv = 0\}$. By Lemma 4.1, $a(t) \in S(P)$. Hence $\langle \nabla f(a(t)), v\rangle \geq 0$ for every $v \in T_\Delta(a(t))$. Combining these facts we see that $\langle \nabla f(a(t)), v\rangle = 0$ for every $v \in M := \{v \in R^n : A_\alpha v = 0,\ Cv = 0\}$. It is easy to check that $y - x \in M$. So we have $\langle \nabla f(a(t)), y - x\rangle = 0$. From this and the obvious relation

$$\nabla \varphi(t) = \langle \nabla f(a(t)), y - x\rangle$$

we deduce that the function φ is constant on $[0,1]$, as desired. □

Proof of Theorem 4.5.

We first consider the case where $\text{Sol}(P)$ is a finite set. Suppose, contrary to our claim, that there exists a pseudo-face F_α of Δ containing two different elements x, y of $\text{Sol}(P)$. By Proposition 4.5, the function $\varphi(t) := f((1-t)x + ty)$ is constant on $[0,1]$. From this and the inclusion $x \in \text{Sol}(P)$ we conclude that whole the segment $[x,y]$ is contained in $\text{Sol}(P)$. This contradicts the finiteness of $\text{Sol}(P)$.

The fact that if $S(P)$ is finite then each pseudo-face of Δ cannot contain more than one element of $S(P)$ follows immediately from Lemma 4.1.

The fact that if $\mathrm{loc}(P)$ is finite then each pseudo-face of Δ cannot contain more than one element of $\mathrm{loc}(P)$ is a direct consequence of Proposition 4.4. □

Actually, in the course of the preceding proof we have established the following useful fact.

Proposition 4.6. *If the intersection* $\mathrm{Sol}(P) \cap F_\alpha$ *of the solution set of* (P) *with a pseudo-face of* Δ *is nonempty, then* $S(P) \cap F_\alpha = \mathrm{Sol}(P) \cap F_\alpha$.

Combining the last proposition with Lemma 4.1 we obtain the following statement.

Proposition 4.7. *The intersection* $\mathrm{Sol}(P)\cap F_\alpha$ *of the solution set of* (P) *with a pseudo-face of* Δ *is always a convex set (may be empty).*

In connection with Propositions 4.4 and 4.7, the following two open questions seem to be interesting:

QUESTION 4: Let $x \in \mathrm{loc}(P) \cap F_\alpha$, $y \in S(P) \cap \overline{F}_\alpha$, $y \neq x$, where F_α is a pseudo-face of Δ and $\overline{F}_\alpha$ denotes the topological closure of F_α. Is it true that there must exist some $\delta > 0$ such that, for every $t \in (0, \delta)$, $a(t) := (1 - t)x + ty$ is a local solution of (P)?

QUESTION 5: Is it true that the intersection $\mathrm{loc}(P) \cap F_\alpha$ of the local-solution set of (P) with a pseudo-face of Δ is always a convex set?

4.5 Commentaries

The notion of solution ray has proved to be very efficient for studying the structure of the solution set of linear complementarity problems (see, for instance, Cottle et al. (1992)) and affine variational inequalities (see, for instance, Gowda and Pang (1994a)).

This chapter shows that the notions of solution ray and solution interval interval are also useful for studying the structure of the solution sets of (nonconvex) quadratic programs.

Lemma 4.1, Propositions 4.2 and 4.3, and Theorems 4.3 and 4.4 are well known facts. Other results might be new.

Chapter 5

Affine Variational Inequalities

In this chapter, the notions of affine variational inequality and linear complementarity problem are discussed in a broader context of variational inequalities and complementarity problems. Besides, a characterization of the solutions of affine variational inequalities via Lagrange multipliers and a basic formula for representing the solution sets will be given.

5.1 Variational Inequalities

Variational inequality problems arise in a natural way in the framework of optimization problems.

Let $f : R^n \to R$ be a C^1-function and $\Delta \subset R^n$ a nonempty, closed, convex set.

Proposition 5.1. *If $\bar{x}$ is a local solution of the optimization problem*

$$\min\{f(x) : x \in \Delta\} \tag{5.1}$$

then

$$\langle \nabla f(\bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall y \in \Delta. \tag{5.2}$$

Proof. Similar to the proof of Theorem 3.1 (i). □

Setting

$$\phi(x) = \nabla f(x) = \begin{pmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x_n} \end{pmatrix} \quad \forall x \in R^n, \tag{5.3}$$

we see that (5.2) can be rewritten as

$$\langle \phi(\bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall y \in \Delta. \tag{5.4}$$

Definition 5.1. If $\Delta \subset R^n$ is a nonempty, closed, convex subset and $\phi : \Delta \to R^n$ is a given operator (mapping) then the problem of finding some $\bar{x} \in \Delta$ satisfying (5.4) is called a *variational inequality problem* or, simply, a *variational inequality* (VI, for brevity). It is denoted by $\mathrm{VI}(\phi, \Delta)$. The *solution set* $\mathrm{Sol}(\mathrm{VI}(\phi, \Delta))$ of $\mathrm{VI}(\phi, \Delta)$ is the set of all $\bar{x} \in \Delta$ satisfying (5.4).

It is easy to check that $\bar{x} \in \mathrm{Sol}(\mathrm{VI}(\phi, \Delta))$ if and only if the inclusion

$$0 \in \phi(\bar{x}) + N_\Delta(\bar{x}),$$

where $N_\Delta(\bar{x})$ denotes the normal cone to Δ at $\bar{x}$ (see Definition 1.9), is satisfied.

Proposition 5.1 shows how smooth optimization problems can lead to variational inequalities. A natural question arises: *Given a variational inequality* $\mathrm{VI}(\phi, \Delta)$ *with a continuous operator* $\phi : R^n \to R^n$*, can one find a* C^1*-function* $f : R^n \to R$ *such that* $\mathrm{VI}(\phi, \Delta)$ *can be obtained from optimization problem* (5.1) *by the above-described procedure or not*? If such a function f exists, we must have

$$\phi(x) = \nabla f(x) \quad \forall x \in \Delta. \tag{5.5}$$

One can observe that if f is a C^2-function then the operator $\phi : R^n \to R^n$ defined by (5.3) has a *symmetric* Jacobian matrix. Recall that if a vector-valued function $\phi : R^n \to R^n$ has smooth components $\phi_1, \ldots, \phi_n$ then the *Jacobian matrix* of ϕ at x is defined by the formula

$$J\phi(x) = \begin{pmatrix} \dfrac{\partial \phi_1(x)}{\partial x_1} & \dfrac{\partial \phi_1(x)}{\partial x_2} & \cdots & \dfrac{\partial \phi_1(x)}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial \phi_n(x)}{\partial x_1} & \dfrac{\partial \phi_n(x)}{\partial x_2} & \cdots & \dfrac{\partial \phi_n(x)}{\partial x_n} \end{pmatrix}.$$

Since f is assumed to be a C^2-function, from (5.3) we deduce that

$$\frac{\partial \phi_i(x)}{\partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial \phi_j(x)}{\partial x_i}$$

for all i, j. This shows that $J\phi(x)$ is a symmetric matrix.

Proposition 5.2. (See, for instance, Nagurney (1993)) *Let* $\Delta \subset R^n$ *be a nonempty, closed, convex set. If* $\phi : R^n \to R^n$ *is such a vector-valued function with smooth components that* $\dfrac{\partial \phi_i(x)}{\partial x_j} = \dfrac{\partial \phi_j(x)}{\partial x_i}$ *for all* i *and* j *(a smooth symmetric operator), then there exists a* C^2*-function* $f : R^n \to R$ *such that the relation* (5.5) *is satisfied. This means that the variational inequality problem* $\mathrm{VI}(\phi, \Delta)$ *can be regarded as the first-order necessary optimality condition of the optimization problem* (5.1).

So, we have seen that C^2*-smooth optimization problems correspond to variational inequalities with smooth symmetric operators.* However, when one studies the VI model, one can consider also VI problems with asymmetric discontinuous operators. Thus the VI model is a mathematical subject which is treated independently from its original interpretation as the first-order necessary optimality condition of a smooth optimization problem.

The following simple statement shows that, unlike the solutions of mathematical programming problems, solutions of VI problems have a local character. From this point of view, VI problems should be regarded as *generalized equations* (see, for instance, Robinson (1979, 1981)), but not as something similar to optimization problems.

Proposition 5.3. *Let* $\bar{x} \in \Delta$*. If there exists* $\varepsilon > 0$ *such that*

$$\langle \phi(\bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall y \in \Delta \cap \bar{B}(\bar{x}, \varepsilon), \tag{5.6}$$

then $\bar{x} \in \mathrm{Sol}(\mathrm{VI}(\phi, \Delta))$.

Proof. Suppose that $\varepsilon > 0$ satisfies (5.6). Obviously, for each $y \in \Delta$ there exists $t = t(y) \in (0, 1)$ such that $y(t) := \bar{x} + t(y - \bar{x})$ belongs to $\Delta \cap \bar{B}(\bar{x}, \varepsilon)$. By (5.6), $0 \leq \langle \phi(\bar{x}), y(t) - \bar{x} \rangle = t\langle \phi(\bar{x}), y - \bar{x} \rangle$. This implies that $\langle \phi(\bar{x}), y - \bar{x} \rangle \geq 0$ for every $y \in \Delta$. Hence $\bar{x} \in \mathrm{Sol}(\mathrm{VI}(\phi, \Delta))$. $\square$

Problem $\mathrm{VI}(\phi, \Delta)$ depends on two data: the set Δ and the operator ϕ. Structure of the solution set $\mathrm{Sol}(\mathrm{VI}(\phi, \Delta))$ is decided by the properties of the set and the operator. In variational inequality theory, the following topics are fundamental: solution existence and uniqueness, stability and sensitivity of the solution sets with respect to perturbations of the problem data, algorithms for finding all the solutions or one part of the solution set.

The following Hartman-Stampacchia Theorem is a fundamental existence theorem for VI problems. It is proved by using the Brouwer fixed point theorem.

Theorem 5.1. (See Hartman and Stampacchia (1966), Kinderlehrer and Stampacchia (1980), Theorem 3.1 in Chapter 1) *If $\Delta \subset R^n$ is nonempty, compact, convex and $\phi : \Delta \to R^n$ is continuous, then problem* $\mathrm{VI}(\phi, \Delta)$ *has a solution.*

Under suitable *coercivity conditions*, one can have existence theorems for problems on noncompact convex sets. For example, the following result is valid.

Theorem 5.2. (See Kinderlehrer and Stampacchia (1980), p. 14) *Let $\Delta \subset R^n$ be a nonempty, closed, convex set and $\phi : \Delta \to R^n$ a continuous operator. If there exists $x^0 \in \Delta$ such that*

$$\langle \phi(y) - \phi(x^0), y - x^0 \rangle \|y - x^0\| \to +\infty \text{ as } \|y\| \to +\infty, \ y \in \Delta, \tag{5.7}$$

then problem $\mathrm{VI}(\phi, \Delta)$ *has a solution.*

The exact meaning of (5.7) is as follows: Given any $\gamma > 0$ one can find $\rho > 0$ such that

$$\frac{\langle \phi(y) - \phi(x^0), y - x^0 \rangle}{\|y - x^0\|} \geq \gamma \quad \text{for every } y \in \Delta \quad \text{satisfying } \|y\| > \rho.$$

It is obvious that if Δ is compact then, for any $x^0 \in \Delta$, (5.7) is valid. If there exists $x^0 \in \Delta$ such that (5.7) holds then one says that the *coercivity condition* is satisfied. Coercivity conditions play an important role in the study of variational inequalities on noncompact constraint sets. Note that (5.7) is only one of the most well-known forms of coercivity conditions.

If there exists $x^0 \in \Delta$ and $\alpha > 0$ such that

$$\langle \phi(y) - \phi(x^0), y - x^0 \rangle \geq \alpha \|y - x^0\|^2 \quad \forall y \in \Delta \tag{5.8}$$

then, surely, (5.7) holds. It is clear that there exists $\alpha > 0$ such that

$$\langle \phi(y) - \phi(x), y - x \rangle \geq \alpha \|y - x\|^2 \quad \forall x \in \Delta, \ \forall y \in \Delta, \tag{5.9}$$

then (5.8) is satisfied.

Definition 5.2. If there exists $\alpha > 0$ such that (5.9) holds then ϕ is said to be *strongly monotone* on Δ. If the following weaker conditions

$$\langle \phi(y) - \phi(x), y - x \rangle > 0 \quad \forall x \in \Delta, \ \ \forall y \in \Delta, \ \ x \neq y, \tag{5.10}$$

and

$$\langle \phi(y) - \phi(x), y - x \rangle \geq 0 \quad \forall x \in \Delta, \ \ \forall y \in \Delta, \tag{5.11}$$

hold, then ϕ is said to be *strictly monotone* on Δ and *monotone* on Δ, respectively.

Example 5.1. Let $\Delta \subset R^n$ is a nonempty, closed, convex set. Let $D \in R^{n \times n}$ and $c \in R^n$. If matrix D is positive definite then the operator $\phi : \Delta \to R^n$ defined by $\phi(x) = Dx + c$, $x \in \Delta$, is strongly monotone on Δ. In this case, it is easily verified that $\alpha > 0$ required for the fulfilment of (5.9) can be defined by setting

$$\alpha = \inf\{v^T D v \, : \, v \in R^n, \ \ \|v\| = 1\}.$$

Likewise, if D is positive semidefinite then the formula $\phi(x) = Dx + c$, $x \in \Delta$, defines a monotone operator.

Proposition 5.4. *The following statements are valid:*

(i) *If ϕ is strictly monotone on Δ then problem* VI(ϕ, Δ) *cannot have more than one solution;*

(ii) *If ϕ is continuous and monotone on Δ then the solution set of problem* VI(ϕ, Δ) *is closed and convex (possibly empty).*

For proving the second statement in the preceding proposition we shall need the following useful fact about monotone VI problems.

Lemma 5.1. (The Minty Lemma; Kinderlehrer and Stampacchia (1980), Lemma 1.5 in Chapter 3) *If $\Delta \subset R^n$ is a closed, convex set and $\phi : \Delta \to R^n$ is a continuous, monotone operator, then* $\bar{x} \in \text{Sol}(\text{VI}((\phi, \Delta))$ *if and only if $\bar{x} \in \Delta$ and*

$$\langle \phi(y), y - \bar{x} \rangle \geq 0 \quad \forall y \in \Delta. \tag{5.12}$$

Proof. *Necessity:* Let $\bar{x} \in \text{Sol}(\text{VI}((\phi, \Delta))$. By the monotonicity of ϕ, we have

$$\langle \phi(y) - \phi(\bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall y \in \Delta.$$

Combining this with (5.4) yields

$$\langle \phi(y), y - \bar{x} \rangle \geq \langle \phi(\bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall y \in \Delta.$$

Property (5.12) has been established.

Sufficiency: Suppose that $\bar{x} \in \Delta$ and (5.12) is satisfied. Fix any $y \in \Delta$. By the convexity of Δ, $y(t) := \bar{x} + t(y - \bar{x})$ belongs to Δ for every $t \in (0, 1)$. Substituting $y = y(t)$ into (5.12) gives

$$0 \leq \langle \phi(y(t)), y(t) - \bar{x} \rangle = \langle \phi(\bar{x} + t(y - \bar{x}), t(y - \bar{x}) \rangle.$$

This implies that

$$\langle \phi(\bar{x} + t(y - \bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall t \in (0, 1).$$

Letting $t \to 0$, by the continuity of ϕ we obtain $\langle \phi(\bar{x}), y - \bar{x} \rangle \geq 0$. Since the last inequality holds for every $y \in \Delta$, we conclude that $\bar{x} \in \mathrm{Sol}(\mathrm{VI}((\phi, \Delta))$. □

Proof of Proposition 5.4.

(i) Suppose, contrary to our claim, that ϕ is strictly monotone on Δ but problem $\mathrm{VI}(\phi, \Delta)$ has two different solutions $\bar{x}$ and $\bar{y}$. Then $\langle \phi(\bar{x}), \bar{y} - \bar{x} \rangle \geq 0$ and $\langle \phi(\bar{y}), \bar{x} - \bar{y} \rangle \geq 0$. Combining these inequalities we get $\langle \phi(\bar{x}) - \phi(\bar{y}), \bar{y} - \bar{x} \rangle \geq 0$. The last inequality contradicts the fact that $\langle \phi(\bar{y}) - \phi(\bar{x}), \bar{y} - \bar{x} \rangle > 0$.

(ii) Assume that ϕ is continuous and monotone on Δ. For every $y \in \Delta$, denote by $\Omega(y)$ the set of all $\bar{x} \in \Delta$ satisfying the inequality $\langle \phi(y), y - \bar{x} \rangle \geq 0$. It is clear that $\Omega(y)$ is closed and convex. From Lemma 5.1 it follows that

$$\mathrm{Sol}(\mathrm{VI}((\phi, \Delta)) = \bigcap_{y \in \Delta} \Omega(y).$$

Hence $\mathrm{Sol}(\mathrm{VI}((\phi, \Delta))$ is closed and convex (possibly empty). □

From Theorem 5.2 and Proposition 5.4(i) it follows that if Δ is nonempty and $\phi : \Delta \to R^n$ is a continuous, strongly monotone operator then problem $\mathrm{VI}(\phi, \Delta)$ has a unique solution.

In the next section, we will consider variational inequality problems in the case where the constraint set Δ is a cone.

5.2 Complementarity Problems

The following fact paves a way to the notion of (nonlinear) complementarity problem.

Proposition 5.5. *If Δ is a closed convex cone, then problem* $\mathrm{VI}(\phi, \Delta)$ *can be rewritten equivalently as follows*

$$\bar{x} \in \Delta, \quad \phi(\bar{x}) \in \Delta^+, \quad \langle \phi(\bar{x}), \bar{x} \rangle = 0, \tag{5.13}$$

where $\Delta^+ = \{\xi \in R^n : \langle \xi, v \rangle \geq 0 \ \ \forall v \in \Delta\}$ *denotes the positive dual cone of* Δ.

Proof. Let $\bar{x}$ be a solution of (5.4). For any $v \in \Delta$, since Δ is a convex cone, we have $\bar{x} + v \in \Delta$. From (5.4) we deduce that

$$0 \leq \langle \phi(\bar{x}), (\bar{x} + v) - \bar{x} \rangle = \langle \phi(\bar{x}), v \rangle.$$

So $\phi(\bar{x}) \in \Delta^+$. Furthermore, since $\frac{1}{2}\bar{x} \in \Delta$ and $2\bar{x} \in \Delta$, by (5.4) we have

$$0 \leq \langle \phi(\bar{x}), \frac{1}{2}\bar{x} - \bar{x} \rangle = -\frac{1}{2}\langle \phi(\bar{x}), \bar{x} \rangle$$

and

$$0 \leq \langle \phi(\bar{x}), 2\bar{x} - \bar{x} \rangle = \langle \phi(\bar{x}), \bar{x} \rangle.$$

Hence $\langle \phi(\bar{x}), \bar{x} \rangle = 0$. We have proved that (5.13) is satisfied.

Now, let $\bar{x}$ be such that (5.13) holds. For every $y \in \Delta$, since $\langle \phi(\bar{x}), \bar{x} \rangle = 0$ and $\phi(\bar{x}) \in \Delta^+$, we have

$$\langle \phi(\bar{x}), y - \bar{x} \rangle = \langle \phi(\bar{x}), y \rangle \geq 0.$$

This shows that $\bar{x} \in \mathrm{Sol}(\mathrm{VI}((\phi, \Delta))$. □

Definition 5.3. Problem (5.13) where $\Delta \subset R^n$ is a closed convex cone and $\phi : R^n \to R^n$, is denoted by $\mathrm{NCP}(\phi, \Delta)$ and is called the *(nonlinear) complementarity problem* defined by ϕ and Δ.

Since complementarity problems are variational inequality problems of a special type, existence theorems for VI problems can be applied to them.

5.3 Affine Variational Inequalities

By Theorem 3.1, if $\bar{x}$ is a local solution of the quadratic program

$$\min \left\{ f(x) = \frac{1}{2}x^T M x + q^T x \; : \; x \in \Delta \right\}, \tag{5.14}$$

where $M \in R_S^{n\times n}$, $q \in R^n$, and $\Delta \subset R^n$ is a polyhedral convex set, then $\langle M\bar{x} + q, y - \bar{x}\rangle \geq 0$ for every $y \in \Delta$. This implies that $\bar{x}$ is a solution of the problem $\mathrm{VI}(\phi, \Delta)$ where $\phi(x) = Mx + q$ is an *affine* operator having the constant symmetric Jacobian matrix M.

Definition 5.4. Let $M \in R^{n\times n}$, $q \in R^n$. Let $\Delta \subset R^n$ be a polyhedral convex set. The variational inequality problem

$$\text{Find } \bar{x} \in \Delta \text{ such that } \langle M\bar{x} + q, y - \bar{x}\rangle \geq 0 \quad \forall y \in \Delta \tag{5.15}$$

is called the *affine variational inequality* (AVI, for brevity) problem defined by the data set $\{M, q, \Delta\}$ and is denoted by $\mathrm{AVI}(M, q, \Delta)$. The solution set of this problem is abbreviated to $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$.

The remarks given at the beginning of this section show that quadratic programs lead to *symmetric* AVI problems. Later on, in the study of AVI problems we will not restrict ourselves only to the case of the symmetric problems.

The following theorem shows that solutions of an AVI problem can be characterized by using some Lagrange multipliers.

Theorem 5.3. (See, for instance, Gowda and Pang (1994b), p. 834) *Vector $\bar{x} \in R^n$ is a solution of* (5.15) *where Δ is given by the formula*

$$\Delta = \{x \in R^n \,:\, Ax \geq b\} \tag{5.16}$$

with $A \in R^{m\times n}$, $b \in R^m$, if and only if there exists $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m) \in R^m$ such that

$$\begin{cases} M\bar{x} - A^T\bar{\lambda} + q = 0, \\ A\bar{x} \geq b, \quad \bar{\lambda} \geq 0, \\ \bar{\lambda}^T(A\bar{x} - b) = 0. \end{cases} \tag{5.17}$$

Proof. The necessity part of this proof is very similar to the proof of Theorem 3.3. As in the preceding chapters, we denote by A_i the i-th row of A and by b_i the i-th component of vector b. We set $a_i = A_i^T$ for every $i = 1, \ldots, m$. Let $\bar{x} \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$. Define $I = \{1, \ldots, m\}$, $I_0 = \{i \in I \,:\, \langle a_i, \bar{x}\rangle = b_i\}$ and $I_1 = \{i \in I \,:\, \langle a_i, \bar{x}\rangle > b_i\}$. For any $v \in R^n$ satisfying

$$\langle a_i, v\rangle \geq 0 \quad \text{for every } i \in I_0,$$

it is easily seen that there exists $\delta_1 > 0$ such that $\langle a_i, \bar{x} + tv\rangle \geq b_i$ for every $i \in I$ and $t \in (0, \delta_1)$. Substituting $y = \bar{x} + tv$, where $t \in (0, \delta_1)$, into (5.15) gives $\langle M\bar{x} + q, v\rangle \geq 0$. Thus

$$\langle -M\bar{x} - q, v\rangle \leq 0$$

for any $v \in R^n$ satisfying

$$\langle -a_i, v\rangle \le 0 \quad \text{for every } i \in I_0.$$

By the Farkas Lemma (see Theorem 3.2), there exist non-negative real numbers $\bar{\lambda}_i$ $(i \in I_0)$ such that

$$\sum_{i\in I_0} \bar{\lambda}_i(-a_i) = -M\bar{x} - q. \tag{5.18}$$

Put $\bar{\lambda}_i = 0$ for all $i \in I_1$ and $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m)$. Since $a_i = A_i^T$ for every $i \in I$, from (5.18) we obtain the first equality in (5.17). Since $\bar{x} \in \Delta(A, b)$ and $\bar{\lambda}_i(A_i\bar{x} - b_i) = 0$ for each $i \in I$, the other conditions in (5.17) are also satisfied.

In order to prove the sufficiency part, suppose that there exists $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \lambda_m) \in R^m$ such that (5.17) holds. Then, for any $y \in \Delta$, one has

$$\begin{aligned}\langle M\bar{x} + q, y - \bar{x}\rangle = \langle A^T\bar{\lambda}, y - \bar{x}\rangle &= \langle\bar{\lambda}, (Ay - b) - (A\bar{x} - b)\rangle \\ &= \bar{\lambda}^T(Ay - b) + \bar{\lambda}^T(A\bar{x} - b) \\ &= \bar{\lambda}^T(Ay - b) \ge 0.\end{aligned}$$

This shows that $\bar{x}$ is a solution of (5.15). The proof is complete. □

One can derive from Theorem 5.3 the following two corollaries, one of which is applicable to the situation where Δ has the representation

$$\Delta = \{x \in R^n : Ax \ge b, \ x \ge 0\} \tag{5.19}$$

and the other is applicable to the situation where Δ has the representation

$$\Delta = \{x \in R^n : Ax \ge b, \ Cx = d\}. \tag{5.20}$$

Here $A \in R^{m\times n}$, $b \in R^m$, $C \in R^{s\times n}$, and $d \in R^s$.

Corollary 5.1. *Vector $\bar{x} \in R^n$ is a solution of (5.15) where Δ is given by (5.19) if and only if there exists $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m) \in R^m$ such that*

$$\begin{cases} M\bar{x} - A^T\bar{\lambda} + q \ge 0, \\ A\bar{x} \ge b, \quad \bar{x} \ge 0, \quad \bar{\lambda} \ge 0, \\ \bar{x}^T(M\bar{x} - A^T\bar{\lambda} + q) + \bar{\lambda}^T(A\bar{x} - b) = 0. \end{cases} \tag{5.21}$$

Proof. Define matrix $\tilde{A} \in R^{(m+n)\times n}$ and vector $\tilde{b} \in R^{m+n}$ as in the proof of Corollary 2.5. Then problem (5.15), where Δ is given by (5.19), is equivalent to the problem

$$\begin{cases} \text{Find } \ \bar{x} \in \tilde{\Delta} := \{x \in R^n \ : \ \tilde{A}x \geq \tilde{b}\} \ \text{ such that} \\ \langle M\bar{x} + q, y - \bar{x}\rangle \geq 0 \quad \forall y \in \tilde{\Delta}. \end{cases}$$

Applying Theorem 5.3 to this AVI problem we deduce that $\bar{x}$ is a solution of the latter if and only if there exists $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_{m+n}) \in R^{m+n}$ such that

$$M\bar{x} - \tilde{A}^T\tilde{\lambda} + q = 0, \quad \tilde{A}\bar{x} \geq \tilde{b}, \quad \tilde{\lambda} \geq 0, \quad \tilde{\lambda}^T(\tilde{A}\bar{x} - \tilde{b}) = 0.$$

Taking $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m)$ where $\bar{\lambda}_i = \lambda_i$ for every $i \in \{1, \ldots, m\}$, we can obtain the desired properties in (5.21) from the last ones. □

Corollary 5.2. *Vector $\bar{x} \in R^n$ is a solution of* (5.15) *where Δ is given by* (5.20) *if and only if there exist $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m) \in R^m$ and $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_s) \in R^s$ such that*

$$\begin{cases} M\bar{x} - A^T\bar{\lambda} - C^T\bar{\mu} + q = 0, \\ A\bar{x} \geq b, \quad C\bar{x} = d, \quad \bar{\lambda} \geq 0, \\ \bar{\lambda}^T(A\bar{x} - b) = 0. \end{cases} \tag{5.22}$$

Proof. Define $\tilde{A} \in R^{(m+2s)\times n}$ and $\tilde{b} \in R^{m+2s}$ as in the proof of Corollary 2.6. Then problem (5.15), where Δ is given by (5.20), is equivalent to the problem

$$\begin{cases} \text{Find } \bar{x} \in \tilde{\Delta} := \{x \in R^n \ : \ \tilde{A}x \geq \tilde{b}\} \text{ such that} \\ \langle M\bar{x} + q, y - \bar{x}\rangle \geq 0 \quad \forall y \in \tilde{\Delta}. \end{cases}$$

Applying Theorem 5.3 to this AVI problem we deduce that $\bar{x}$ is a solution of the latter if and only if there exists $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_{m+2s}) \in R^{m+2s}$ such that

$$M\bar{x} - \tilde{A}^T\tilde{\lambda} + q = 0, \quad \tilde{A}\bar{x} \geq \tilde{b}, \quad \tilde{\lambda} \geq 0, \quad \tilde{\lambda}^T(\tilde{A}\bar{x} - \tilde{b}) = 0.$$

Taking $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m)$ and $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_s)$ where $\bar{\lambda}_i = \lambda_i$ for every $i \in \{1, \ldots, m\}$ and $\bar{\mu}_j = \lambda_{m+j} - \lambda_{m+s+j}$ for every $j \in \{1, \ldots, s\}$, we can obtain the properties stated in (5.22) from the last ones. □

Unlike the solution set and the local solution set of a nonconvex quadratic program, the solution set of an AVI problem has a rather simple structure.

Theorem 5.4. *The solution set of any affine variational inequality problem is the union of finitely many polyhedral convex sets.*

Proof. This proof follows the idea of the proof of formula (4.12). Consider a general AVI problem in the form (5.15). Since Δ is a polyhedral convex set, there exists $m \in N$, $A \in R^{m\times n}$, $b \in R^m$ such that $\Delta = \{x \in R^n : Ax \geq b\}$. According to Theorem 5.3, $x \in \text{Sol}(\text{AVI}(M, q, \Delta))$ if and only if there exists $\lambda = (\lambda_1, \ldots, \lambda_m) \in R^m$ such that

$$\begin{cases} Mx - A^T\lambda + q = 0, \\ Ax \geq b, \quad \lambda \geq 0, \\ \lambda^T(Ax - b) = 0. \end{cases} \tag{5.23}$$

Let $I = \{1, \ldots, m\}$. Given a point $x \in \text{Sol}(\text{AVI}(M, q, \Delta))$, we set $I_0 = \{i \in I : A_i x = b_i\}$, $I_1 = I \setminus I_0 = \{i \in I : A_i x > b_i\}$. From the last equality in (5.23) we get

$$\lambda_i = 0 \quad \forall i \in I_1.$$

Hence (x, λ) satisfies the system

$$\begin{cases} Mx - A^T\lambda + q = 0, \\ A_{I_0}x = b_{I_0}, \quad \lambda_{I_0} \geq 0, \\ A_{I_1}x \geq b_{I_1}, \quad \lambda_{I_1} = 0. \end{cases} \tag{5.24}$$

Fix any subset $I_0 \subset I$ and denote by Q_{I_0} the set of all (x, λ) satisfying (5.24). It is obvious that Q_{I_0} is a polyhedral convex set. From what has been said it follows that

$$\text{Sol}(\text{AVI}(M, q, \Delta)) = \bigcup \{\text{Pr}_{R^n}(Q_{I_0}) : I_0 \subset I\}, \tag{5.25}$$

where $\text{Pr}_{R^n}(x, \lambda) := x$. Since $\text{Pr}_{R^n}(\cdot) : R^n \times R^m \to R^n$ is a linear operator, for every $I_0 \subset I$, $\text{Pr}_{R^n}(Q_{I_0})$ is a polyhedral convex set. From (5.25) it follows that $\text{Sol}(\text{AVI}(M, q, \Delta))$ is the union of finitely many polyhedral convex sets. □

Definition 5.5. A half-line $\omega = \{\bar{x} + t\bar{v} : t \geq 0\}$, where $\bar{v} \in R^n \setminus \{0\}$, which is a subset of $\text{Sol}(\text{AVI}(M, q, \Delta))$, is called a *solution ray* of problem (5.15).

Definition 5.6. A line segment $\omega_\delta = \{\bar{x} + t\bar{v} : t \in [0, \delta)\}$, where $\bar{v} \in R^n \setminus \{0\}$ and $\delta > 0$, which is a subset of $\text{Sol}(\text{AVI}(M, q, \Delta))$, is called a *solution interval* of problem (5.15).

Corollary 5.3. *The following statements hold:*

(i) *The solution set of any affine variational inequality is a closed set (possibly empty);*

(ii) *If the solution set of an affine variational inequality is unbounded, then the problem has a solution ray;*

(iii) *If the solution set of an affine variational inequality is infinite, then the problem has a solution interval.*

Proof. Statement (i) follows directly from formula (5.25) because, for any $I_0 \subset I$, the set $\mathrm{Pr}_{R^n}(Q_{I_0})$, being polyhedral convex, is closed. If $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ is unbounded then from (5.25) it follows that there exists an index set $I_0 \subset I$ such that

$$\Omega_{I_0} := \mathrm{Pr}_{R^n}(Q_{I_0}) \tag{5.26}$$

is an unbounded set. Since Ω_{I_0} is a polyhedral convex set, it is an unbounded closed convex set. By Theorem 8.4 in Rockafellar (1970), Ω_{I_0} admits a direction of recession; that is there exists $\bar{v} \in R^n \setminus \{0\}$ such that

$$x + t\bar{v} \in \Omega_{I_0} \forall x \in \Omega_{I_0},\ \forall t \geq 0. \tag{5.27}$$

Taking any $\bar{x} \in \Omega_{I_0}$ we deduce from (5.25) and (5.27) that $\bar{x} + t\bar{v} \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ for all $t \geq 0$. Thus we have proved that problem (5.15) has a solution ray. If $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ is infinite then from (5.25) we deduce that there is an index set $I_0 \subset I$ such that the polyhedral convex set Ω_{I_0} defined by (5.26) is infinite. Then there must exist two different points $x \in \Omega_{I_0}$ and $y \in \Omega_{I_0}$. It is clear that the set $[x, y) := \{x + t(y - x) : t \in [0, 1)\}$ is a solution interval of (5.15). $\square$

Using Theorem 5.4 one can obtain a complete characterization for the unboundedness property of the solution set of an AVI problem. Let us consider problem (5.15) where Δ is given by (5.16) and introduce the following notations:

$$\begin{aligned} \delta(A) &= \{v \in R^n : Av \geq 0\}, \\ \delta(A)^+ &= \{z \in R^n : z^T v \geq 0 \quad \forall v \in \delta(A)\}, \\ \ell(M) &= \{z \in R^n : z^T M z = 0\}. \end{aligned}$$

Note that $\delta(A)$ and $\{v \in R^n : Av \in \delta(A)^+\}$ are polyhedral convex cones, while $\ell(M)$ is, in general, a nonconvex closed cone. Note also that $\delta(A) = 0^+\Delta$ and $\delta(A)^+ = (0^+\Delta)^+$.

Theorem 5.5 (cf. Gowda and Pang (1994a)). *The solution set of* (5.15) *is unbounded if and only if there exists a pair* $(v, u^0) \in R^n \times R^n$, $v \neq 0$, $u^0 \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$, *such that*

(i) $v \in \delta(A), \quad Mv \in \delta(A)^+, \quad v \in \ell(M)$;

(ii) $(Mu^0 + q)^T v = 0$;

(iii) $\langle Mv, y - u^0 \rangle \geq 0 \quad \forall y \in \Delta$.

Proof. *Sufficiency:* Suppose that there is a pair $(v, u^0) \in R^n \times R^n$, $v \neq 0$, $u^0 \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$, such that (i)-(iii) are fulfilled. Let $x_t = u^0 + tv$, $t > 0$. Given any $y \in \Delta$, we deduce from (i)-(iii) that

$$\begin{aligned} \langle Mx_t + q, y - x_t \rangle &= \langle M(u^0 + tv) + q, y - (u^0 + tv) \rangle \\ &= \underbrace{\langle Mu^0 + q, y - u^0 \rangle}_{\geq 0} - t \underbrace{\langle Mu^0 + q, v \rangle}_{=0} \\ &\quad + t \underbrace{\langle Mv, y - u^0 \rangle}_{\geq 0} - t^2 \underbrace{\langle Mv, v \rangle}_{=0} \\ &\geq 0. \end{aligned}$$

This implies that $x_t \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ for every $t > 0$. Hence the solution set is unbounded.

Necessity: Suppose that the set $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ is unbounded. By (5.25), there exists $I_0 \subset I$ such that the set Ω_{I_0} defined by (5.26) is unbounded. Applying Theorem 8.4 from Rockafellar (1970), we can assert that there exist $v \in R^n$, $v \neq 0$, and $u^0 \in \Omega_{I_0}$ such that

$$u^0 + tv \in \Omega_{I_0} \subset \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) \quad \forall t \geq 0. \tag{5.28}$$

Since $A(u^0 + tv) \geq b$ for every $t > 0$, we can deduce that $Av \geq 0$. This means that $v \in \delta(A)$. By (5.28), we have

$$\langle M(u^0 + tv) + q, y - (u^0 + tv) \rangle \geq 0 \quad \forall y \in \Delta. \tag{5.29}$$

Fixing any $y \in \Delta$, we deduce from (5.29) that

$$\left\langle \frac{1}{t} Mu^0 + Mv + \frac{1}{t} q, \frac{1}{t} y - \frac{1}{t} u_0 - v \right\rangle \geq 0 \quad \forall t > 0.$$

Therefore

$$\langle Mv, -v \rangle \geq 0. \tag{5.30}$$

Substituting $y = u^0 + t^2 v$, where $t > 1$, into (5.29) and dividing the inequality by $t(t^2 - t)$, we obtain

$$\left\langle \frac{1}{t} M u^0 + Mv + \frac{1}{t} q, v \right\rangle \geq 0 \quad \forall t > 1.$$

Letting $t \to +\infty$ yields $\langle Mv, v \rangle \geq 0$. Combining this with (5.30) we get

$$\langle Mv, v \rangle = 0. \tag{5.31}$$

This shows that $v \in \ell(M)$. Substituting $y = u^0$ into (5.29) and taking account of (5.31) we have $\langle Mu^0 + q, v \rangle \leq 0$. Substituting $y = u^0 + t^2 v$, where $t > 1$, into (5.29) and using (5.31) we can deduce that $\langle Mu^0 + q, v \rangle \geq 0$. This and the preceding inequality shows that (ii) is satisfied. By (5.29), (5.31) and (ii), for every $y \in \Delta$ we have

$$\begin{aligned} 0 &\leq \langle Mu^0 + q + tMv, y - u^0 - tv \rangle \\ &= \langle Mu^0 + q, y - u^0 \rangle + t \langle Mv, y - u^0 \rangle \end{aligned}$$

for all $t > 0$. This implies that the inequality $\langle Mv, y - u^0 \rangle < 0$ must be false. So we have

$$\langle Mv, y - u^0 \rangle \geq 0 \quad \forall y \in \Delta. \tag{5.32}$$

Substituting $y = u^0 + w$, where $w \in \delta(A)$, into the inequality in (5.32) we deduce that $\langle Mv, w \rangle \geq 0$ for every $w \in \delta(A)$. This means that $Mv \in \delta(A)^+$. We have thus shown that all the three inclusions in (i) are valid. The proof is complete. □

Several simple sufficient conditions for (5.15) to have a compact solution set can be obtained directly from the preceding theorem.

Corollary 5.4. *Problem* (5.15) *has a compact solution set (possibly empty) if one of the following conditions is satisfied:*

(γ_1) *the cone* $\ell(M)$ *consists of only one element* 0*;*

(γ_2) *the intersection of the cones* $\ell(M)$ *and* $\{v \in R^n : Mv \in \delta(A)^+\}$ *consists of only one element* 0*;*

(γ_3) *the intersection of the cones* $\ell(M)$, $\{v \in R^n : Mv \in \delta(A)^+\}$ *and* $\delta(A)$, *consists of only one element* 0.

Examples given in the next section will show how the above sufficient conditions can be used in practice.

5.4 Linear Complementarity Problems

We now consider a special case of the model (5.13) which plays a very important role in theory of finite-dimensional variational inequalities and complementarity problems (see, for instance, Harker and Pang (1990) and Cottle et al. (1992)).

Definition 5.7. Problem (5.13) with $\Delta = R^n_+$ and $\phi(x) = Mx + q$ where $M \in R^{n\times n}$ and $q \in R^n$, is denoted by LCP(M, q) and is called the *linear complementarity problem* defined by M and q. The solution set of this problem is denoted by Sol(M, q).

We can write LCP(M, q) as follows

$$\bar{x} \geq 0, \quad M\bar{x} + q \geq 0, \quad \bar{x}^T(M\bar{x} + q) = 0. \tag{5.33}$$

Thus LCP problem is a special case of the NCP problem where $\Delta = R^n_+$ and ϕ is an affine operator.

If $\bar{x}$ is a local solution of quadratic program (3.1) where $\Delta = R^n_+$, then $\bar{x} \in R^n_+$ and, by Theorem 3.1,

$$\langle D\bar{x} + c, y - \bar{x}\rangle \geq 0 \quad \forall y \in R^n_+.$$

This amounts to saying that $\bar{x}$ is a solution of the linear complementarity problem LCP(D, c) defined D and c.

By Corollary 3.1, if $\bar{x}$ is a local solution of the quadratic program (2.26) then there exists $\lambda = (\lambda_1, \ldots, \lambda_m) \in R^m$ such that (3.8) holds. Setting

$$M = \begin{bmatrix} D & -A^T \\ A & 0 \end{bmatrix}, \quad q = \begin{pmatrix} c \\ -b \end{pmatrix}, \quad \bar{z} = \begin{pmatrix} \bar{x} \\ \lambda \end{pmatrix}, \tag{5.34}$$

we have $M \in R^{(n+m)\times(n+m)}$, $q \in R^{n+m}$, $\bar{z} \in R^{n+m}$. It is easily verified that (3.8) is equivalent to the system

$$\bar{z} \geq 0, \quad M\bar{z} + q \geq 0, \quad \bar{z}^T(M\bar{z} + q) = 0.$$

Thus (3.8) can be interpreted as a LCP problem.

Definition 5.8. If Δ is a polyhedral convex cone and there exist $M \in R^{n\times n}$, $q \in R^n$, such that $\phi(x) = Mx + q$ for every $x \in \Delta$, then (5.13) is said to be a *generalized linear complementarity problem.* It is denoted by GLCP(M, q, Δ).

From the above definition we see that generalized linear complementarity problems are the AVI problems of a special type. Comparing Definition 5.8 with Definition 5.7 we see at once that the structure of GLCP problems is very similar to that of LCP problems. This explains why many results concerning LCP problems can be extended to GLCP problems.

It is easily seen that if in (5.15) one chooses $\Delta = R^n_+$ then one obtains the linear complementarity problem LCP(M, q). Hence linear complementarity problems are the AVI problems of a special type. In this book, as a rule, we try first to prove theorems (on the solution existence, on the solution stability, etc.) for AVI problems then apply them to LCP problems.

Theorem 5.5 can be specialized for LCP problems as follows.

Proposition 5.6. (See Yen and Hung (2001), Theorem 2) *The solution set of* (5.33) *is unbounded if and only if there exists a pair* $(v, u^0) \in R^n \times R^n$, $v \neq 0$, $u^0 \in \mathrm{Sol}(M, q)$, *such that*

(i) $v \geq 0, \quad Mv \geq 0, \quad v \in \ell(M)$;

(ii) $(Mu^0 + q)^T v = 0$;

(iii) $\langle Mv, u^0 \rangle = 0$.

Corollary 5.4 is specialized for LCP problems as follows.

Corollary 5.5. *Problem* (5.33) *has a compact solution set (possibly empty) if one of the following conditions is satisfied:*

(γ_1) *the cone* $\ell(M)$ *consists of only one element* 0;

(γ_2) *the intersection of the cones* $\ell(M)$ *and* $\{v \in R^n : Mv \geq 0\}$ *consists of only one element* 0;

(γ_3) *the intersection of the cones* $\ell(M)$, $\{v \in R^n : Mv \geq 0\}$ *and* R^n_+, *consists of only one element* 0.

Example 5.2. (See Yen and Hung (2001)) Consider problem (5.33) with

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \in R^2, \quad n = 2.$$

A direct computation shows that the intersection of the cones $\ell(M)$, $\{v : Mv \geq 0\}$ and $\{v : v \geq 0\}$, consists of only 0. By Corollary 5.5, $\mathrm{Sol}(M, q)$ is a compact set.

Observe that M in the above example is a nondegenerate matrix, so from the theory in Chapter 3 of Cottle et al. (1982), it follows that $\mathrm{Sol}(M, q)$ is a finite set. By definition, $M = (a_{ij})$ is said to be a *nondegenerate matrix* if, for any nonempty subset $\alpha \subset \{1, \ldots, n\}$, the determinant of the principal submatrix $M_{\alpha\alpha}$ consisting of the elements a_{ij} $(i \in \alpha,\ j \in \alpha)$ of M is nonzero.

Example 5.3. (See Yen and Hung (2001)) Consider problem (5.13) with

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad q = \begin{pmatrix} -1 \\ \mu \end{pmatrix} \in R^2, \quad \mu \neq 0, \quad n = 2.$$

A direct computation shows that the intersection of the cones $\ell(M)$, $\{v : Mv \geq 0\}$ and $\{v : v \geq 0\}$, is the set $\{v = (0, v_2) \in R^2 : v_2 \geq 0\}$. For verifying condition (ii) in Proposition 5.6, there is no loss of generality in assuming that $v = (0, 1)$. It is easy to show that there is no $u^0 \geq 0$ such that $(Mu^0 + q)^T v = 0$ and $\langle Mv, u^0 \rangle = 0$. By Proposition 5.6, $\mathrm{Sol}(M, q)$ is a compact set.

Example 5.4. (See Yen and Hung (2001)) Consider problem (5.13) with

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad q = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \in R^2, \quad n = 2.$$

The intersection of the cones $\ell(M)$, $\{v : Mv \geq 0\}$ and $\{v : v \geq 0\}$, is the set $\{v = (0, v_2) \in R^2 : v_2 \geq 0\}$. It is easy to show that conditions (i)-(iii) in Proposition 5.6 are satisfied if we choose $v = (0, 1)$ and $u^0 = (1, 0)$. By Proposition 5.6, $\mathrm{Sol}(M, q)$ is an unbounded set.

5.5 Commentaries

Problem (5.4) is finite-dimensional. Infinite-dimensional VI problems are not studied in this book. Systematic studies on infinite-dimensional VI problems with applications to mathematical physics (obstacle problems, etc.) can be found, for example, in Kinderlehrer and Stampacchia (1980), Rodrigues (1987).

The important role of finite-dimensional VI problems and complementarity problems in mathematics and in mathematical applications is well known (see, for instance, Harker and Pang (1990), Nagurney (1993), and Patriksson (1999)).

A comprehensive theory on LCP problems was given by Cottle, Pang and Stone (1992). Several key results on LCP problems have been extended to the case of AVI problems.

The first volume of the book by Facchinei and Pang (2003) describes the basic theory on finite-dimensional VI problems and complementarity problems, while its second volume concentrates on iterative algorithms for solving these problems. The book aims at being an enduring reference on the subject and at providing the foundation for its continued growth.

Robinson (see Robinson (1979), Theorem 2, and Robinson (1981), Proposition 1) obtained two fundamental theorems on Lipschitz continuity of the solution map in general AVI problems, which he called the *linear generalized equations.* In Chapter 7 we will study these theorems.

Chapter 6

Solution Existence for Affine Variational Inequalities

In this chapter, some basic theorems on the solution existence of affine variational inequalities will be proved. Different conditions on monotonicity of the linear operator represented by matrix M and the relative position of vector q with respect to the constraint set Δ and the recession cone $0^+\Delta$ will be used in these theorems. As in the preceding chapter, we denote the problem

$$\text{Find } \bar{x} \in \Delta \text{ such that } \langle M\bar{x} + q, y - \bar{x} \rangle \geq 0 \ \ \forall y \in \Delta \tag{6.1}$$

by AVI(M, q, Δ). Here $M \in R^{n \times n}$, $q \in R^n$, and Δ is a nonempty polyhedral convex set in R^n.

6.1 Solution Existence under Monotonicity

Consider problem (6.1). Since Δ is a polyhedral convex set, there exist $m \in N$, $A \in R^{m \times n}$ and $b \in R^m$ such that

$$\Delta = \{x \in R^n \,:\, Ax \geq b\}. \tag{6.2}$$

Theorem 6.1. (See Gowda and Pang (1994a), p. 432) *If the following two conditions are satisfied*

(i) *there exists* $\bar{x} \in \Delta$ *such that* $(M\bar{x}+q)^T v \geq 0$ *for every* $v \in 0^+\Delta$;

(ii) $(y-x)^T M(y-x) \geq 0$ *for all* $x \in \Delta$ *and* $y \in \Delta$;

then the solution set $\mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$ *is nonempty.*

Let

$$\begin{aligned}\tilde{\Delta} &= \left\{ z = \begin{pmatrix} x \\ \lambda \end{pmatrix} \in R^{n+m} : (M \quad -A^T)z = -q, \right. \\ &\qquad\qquad \left. \begin{pmatrix} A & 0 \\ 0 & E \end{pmatrix} z \geq \begin{pmatrix} b \\ 0 \end{pmatrix} \right\} \\ &= \left\{ z = \begin{pmatrix} x \\ \lambda \end{pmatrix} \in R^{n+m} : Mx - A^T\lambda + q = 0, \ \ Ax \geq b, \ \ \lambda \geq 0 \right\},\end{aligned}$$

where I denotes the unit matrix in $R^{m\times m}$. Let

$$f(z) = \frac{1}{2} z^T(\widetilde{M} + \widetilde{M}^T)z + \begin{pmatrix} q \\ -b \end{pmatrix}^T z,$$

where

$$\widetilde{M} = \begin{pmatrix} M & -A^T \\ A & 0 \end{pmatrix} \in R^{(n+m)\times(n+m)}, \ \ z = \begin{pmatrix} x \\ \lambda \end{pmatrix} \in R^{n+m}.$$

Consider the following *auxiliary* quadratic program

$$\min\{f(z) \, : \, z \in \tilde{\Delta}\}. \tag{6.3}$$

Lemma 6.1. *The set* $\tilde{\Delta}$ *is nonempty if and only if there exists* $\bar{x} \in \Delta$ *such that* $(M\bar{x}+q)^T v \geq 0$ *for every* $v \in 0^+\Delta$.

Proof. *Necessity:* If $\tilde{\Delta} \neq \emptyset$ then there exists $\bar{z} = \begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} \in R^{n+m}$ such that

$$M\bar{x} - A^T\bar{\lambda} + q = 0, \ \ A\bar{x} \geq b, \ \ \bar{\lambda} \geq 0. \tag{6.4}$$

Let $v \in 0^+\Delta$. By (6.2), we have $Av \geq 0$. From (6.4) we deduce that

$$0 = (M\bar{x} - A^T\bar{\lambda} + q)^T v = (M\bar{x}+q)^T v - \bar{\lambda}^T Av.$$

Hence $(M\bar{x}+q)^T v = \bar{\lambda}^T Av \geq 0$.

Sufficiency: Suppose that there exists $\bar{x} \in \Delta$ such that $(M\bar{x} + q)^T v \geq 0$ for every $v \in 0^+\Delta = \delta(A)$. Consider the following linear program

$$\min\{c^T y \, : \, y \in \Delta\}, \tag{6.5}$$

where $c := M\bar{x}+q$. From our assumption it follows that $\Delta \neq \emptyset$ and $(M\bar{x}+q)^T v \geq 0$ whenever $v \in R^n$, $Av \geq 0$. By Theorem 2.2, (6.5) has a solution. According to Theorem 3.3, there exists $\bar{\lambda} \in R^m$ such that

$$-A^T\bar{\lambda} + c = 0, \ \ \bar{\lambda} \geq 0. \tag{6.6}$$

Since $\bar{x} \in \Delta$, we have $A\bar{x} \geq b$. Combining this with (6.6) we deduce that $\bar{z} := \begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix}$ belongs to $\tilde{\Delta}$. So $\tilde{\Delta} \neq \emptyset$. □

Lemma 6.2. *If there exists $\bar{x} \in \Delta$ such that $(M\bar{x} + q)^T v \geq 0$ for every $v \in 0^+\Delta$ then the auxiliary quadratic program* (6.3) *has a solution.*

Proof. By Lemma 6.1, from the assumption it follows that $\tilde{\Delta}$ is nonempty. Let $z = \begin{pmatrix} x \\ \lambda \end{pmatrix} \in \tilde{\Delta}$. We have

$$\begin{aligned}
f(z) &= \frac{1}{2}z^T(\widetilde{M} + \widetilde{M}^T)z + \begin{pmatrix} q \\ -b \end{pmatrix}^T z \\
&= \frac{1}{2}\begin{pmatrix} x \\ \lambda \end{pmatrix}^T \left[\begin{pmatrix} M & -A^T \\ A & 0 \end{pmatrix} + \begin{pmatrix} M & -A^T \\ A & 0 \end{pmatrix}^T\right]\begin{pmatrix} x \\ \lambda \end{pmatrix} \\
&\qquad + \begin{pmatrix} q \\ -b \end{pmatrix}^T \begin{pmatrix} x \\ \lambda \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} x \\ \lambda \end{pmatrix}^T \begin{pmatrix} M + M^T & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ \lambda \end{pmatrix} + q^T x - b^T\lambda \\
&= \frac{1}{2}x^T(M + M^T)x + q^T x - b^T\lambda \\
&= x^T M x + q^T x - b^T\lambda \\
&= x^T(Mx + q) - b^T\lambda \\
&= x^T(A^T\lambda) - b^T\lambda \\
&= \lambda^T(Ax - b) \geq 0.
\end{aligned}$$

So $f(z)$ is bounded from below on $\tilde{\Delta}$. By the Frank-Wolfe Theorem (see Theorem 2.1), (6.3) has a solution. □

Proof of Theorem 6.1.

By assumption (i) and by Lemma 6.2, the auxiliary quadratic problem (6.3) has a solution $\bar{z} = \begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix}$. Hence, by Corollary 3.2 there exist Lagrange multipliers $\theta = \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix} \in R^{2m}$ and $\mu \in R^n$ such that

$$\begin{cases} (\widetilde{M} + \widetilde{M}^T)\begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} - \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}^T \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix} \\ \qquad -(M \;\; - A^T)^T \mu + \begin{pmatrix} q \\ -b \end{pmatrix} = 0, \\ \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} \ge \begin{pmatrix} b \\ 0 \end{pmatrix}, \quad \theta \ge 0, \\ \theta^T \left[\begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} - \begin{pmatrix} b \\ 0 \end{pmatrix} \right] = 0. \end{cases}$$

This system can be written as the following one

$$\begin{cases} \begin{pmatrix} M & -A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} + \begin{pmatrix} M^T & A^T \\ -A & 0 \end{pmatrix}\begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} - \begin{pmatrix} A^T & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix} \\ \qquad - \begin{pmatrix} M^T \\ -A \end{pmatrix} \mu + \begin{pmatrix} q \\ -b \end{pmatrix} = 0 \\ A\bar{x} \ge b, \;\; \bar{\lambda} \ge 0, \;\; \theta^1 \ge 0, \;\; \theta^2 \ge 0, \\ (\theta^1)^T(A\bar{x} - b) = 0, \;\; (\theta^2)^T\bar{\lambda} = 0. \end{cases}$$

In its turn, the latter is equivalent to the system (6.7)–(6.10) below:

$$(M\bar{x} - A^T\bar{\lambda} + q) + M^T\bar{x} + A^T\bar{\lambda} - A^T\theta^1 - M^T\mu = 0, \tag{6.7}$$

$$A\bar{x} - A\bar{x} - \theta^2 + A\mu - b = 0, \tag{6.8}$$

$$A\bar{x} \ge b, \;\; \bar{\lambda} \ge 0, \;\; \theta^1 \ge 0, \;\; \theta^2 \ge 0, \tag{6.9}$$

$$(\theta^1)^T(A\bar{x} - b) = 0, \;\; (\theta^2)^T\bar{\lambda} = 0. \tag{6.10}$$

From (6.7) and the inclusion $\bar{z} = \begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix} \in \tilde{\Delta}$ it follows that

$$M^T(\bar{x} - \mu) = A^T(\theta^1 - \bar{\lambda}). \tag{6.11}$$

From (6.8) it follows that

$$A(\bar{x} - \mu) = A\bar{x} - \theta^2 - b. \tag{6.12}$$

By (6.10)–(6.12),

$$\begin{aligned} (\bar{x} - \mu)^T M^T (\bar{x} - \mu) \;\; &= (\bar{x} - \mu)^T A^T (\theta^1 - \bar{\lambda}) \\ &= (\theta^1 - \bar{\lambda})^T A(\bar{x} - \mu) \\ &= (\theta^1 - \bar{\lambda})^T (A\bar{x} - \theta^2 - b) \\ &= \underbrace{(\theta^1)^T(A\bar{x} - b)}_{=0} - (\theta^1)^T\theta^2 \\ &\qquad + \underbrace{(\theta^2)^T\bar{\lambda}}_{=0} - \bar{\lambda}^T(A\bar{x} - b) \\ &= -(\theta^1)^T\theta^2 - \bar{\lambda}^T(A\bar{x} - b). \end{aligned}$$

Hence, by virtue of (6.9), we have

$$(\bar{x}-\mu)^T M^T(\bar{x}-\mu) = -(\theta^1)^T\theta^2 - \bar{\lambda}^T(A\bar{x}-b) \le 0. \tag{6.13}$$

From (6.8) it follows that

$$A\mu - b = \theta^2 \ge 0.$$

So $\mu \in \Delta$. Since $\bar{x} \in \Delta$, we can deduce from assumption (ii) that $(\bar{x}-\mu)^T M^T(\bar{x}-\mu) \ge 0$. Combining this with (6.9) and (6.13) gives $\bar{\lambda}^T(A\bar{x}-b) = 0$. Since $\binom{\bar{x}}{\bar{\lambda}} \in \tilde{\Delta}$, $M\bar{x} - A^T\bar{\lambda} + q = 0$. Thus we have shown that

$$\begin{cases} M\bar{x} - A^T\bar{\lambda} + q = 0, \\ A\bar{x} \ge b, \ \ \bar{\lambda} \ge 0, \\ \bar{\lambda}^T(A\bar{x}-b) = 0. \end{cases}$$

Then, according to Theorem 5.3, $\bar{x} \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$. □

Assumption (ii) is crucial for the validity of the conclusion of the above theorem. It is easily seen that (ii) is equivalent to the requirement that the operator $\phi : \Delta \to R^n$ defined by setting $\phi(x) = Mx + q$ is monotone on Δ (see Definition 5.2).

Definition 6.1 (cf. Cottle et al. (1992), p. 176). By abuse of terminology, we say that matrix $M \in R^{n\times n}$ is *monotone* on a closed convex set $\Delta \subset R^n$ if the linear operator corresponding to M is monotone on Δ, that is

$$(y-x)^T M(y-x) \ge 0 \ \ \forall x \in \Delta, \ \forall y \in \Delta. \tag{6.14}$$

Matrix M is said to be *copositive* on Δ if

$$v^T M v \ge 0 \ \ \forall v \in 0^+\Delta. \tag{6.15}$$

If M is copositive on R^n_+ then one simply says that M is a *copositive matrix.* Matrix M is said to be *strictly copositive* on Δ if

$$v^T M v > 0 \ \ \forall v \in 0^+\Delta \setminus \{0\}. \tag{6.16}$$

Remark 6.1. Monotonicity implies copositivity. But the reverse implication, in general, is false. Indeed, if (6.14) holds and if Δ is nonempty then, for any $\bar{x} \in \Delta$ and $v \in 0^+\Delta$, we have

$$v^T M v = ((\bar{x}+v)-\bar{x})^T M((\bar{x}+v)-\bar{x}) \ge 0.$$

Hence M is copositive on Δ. To show that in general copositivity does not imply monotonicity, we consider the following example. Let

$$M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in R^{2\times 2}, \quad \Delta = R^2_+.$$

For every $v \in 0^+\Delta = R^2_+$ we have $v^T M v \geq 0$. So M is copositive on Δ. But M is not monotone on Δ. Indeed, choosing $x = (0, 1)$ and $y = (1, 0)$, we see that

$$(y - x)^T M (y - x) = (1 \ \ -1) \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -2 < 0.$$

Remark 6.2. If $\mathrm{int}\Delta \neq \emptyset$ then matrix M is monotone on Δ if and only if M is positive semidefinite. Indeed, it is clear that if $M \in R^{n\times n}$ is a positive semidefinite matrix then, for any nonempty closed convex set $\Delta \subset R^n$, M is copositive on Δ. On the other hand, if $\mathrm{int}\Delta \neq \emptyset$ then there exists $\bar{x} \in \Delta$ and $\varepsilon > 0$ such that $B(\bar{x}, \varepsilon) \subset \Delta$. For every $z \in R^n$ there exists $t > 0$ such that $y := \bar{x} + tz \in B(\bar{x}, \varepsilon) \subset \Delta$. Then we have

$$0 \leq (y - \bar{x})^T M (y - \bar{x}) = t^2 z^T M z.$$

Hence $z^T M z \geq 0$ for every $z \in R^n$.

Remark 6.3. It is clear that if M is strictly copositive on Δ then it is copositive on Δ. The converse is not true in general. For example, if $\Delta = R^2_+$ and $M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, then M is copositive but not strictly copositive on Δ. Indeed, choosing $\bar{v} = (0, 1)$ we see that $\bar{v} \in 0^+\Delta \setminus \{0\} = R^2_+ \setminus \{0\}$ but $\bar{v}^T M \bar{v} = 0$.

We now consider a simple example to see how Theorem 6.1 can be used.

Example 6.1. Let

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in R^{2\times 2}, \quad q = (q_1, q_2) \in R^2,$$
$$\Delta = \{x = (x_1, x_2) \in R^2_+ \, : \, x_1 \geq 0, \ x_2 = 0\}.$$

Theorem 6.1 can be applied to this problem. Indeed, since M is monotone on Δ, it suffices to show that there exists $\bar{x} \in \Delta$ such that $(M\bar{x} + q)^T v \geq 0$ for every $v \in 0^+\Delta$. If $q_1 < 0$ then $\bar{x} = (-q_1, 0)$ satisfies the last condition. If $q_1 \geq 0$ then $\bar{x} = (0, 0)$ satisfies that condition. Further investigation on the problem shows that $\mathrm{Sol}(\mathrm{AVI}(M,q,\Delta)) = \{(-q_1, 0)\}$ if $q_1 < 0$ and $\mathrm{Sol}(\mathrm{AVI}(M,q,\Delta)) = \{(0,0)\}$ if $q_1 \geq 0$.

From Theorem 6.1 it is easy to deduce the following result.

Theorem 6.2. (See Gowda and Pang (1994a), Theorem 1) *If M is a positive semidefinite matrix and there exists $\bar{x} \in \Delta$ such that $(M\bar{x} + q)^T v \geq 0$ for every $v \in 0^+\Delta$, then problem* (6.1) *has a solution.*

Corollary 6.1. (See Cottle et al. (1982), Theorem 3.1.2) *Suppose that M is a positive semidefinite matrix. Then the linear complementarity problem* LCP(M,q) *has a solution if and only if there exists $\bar{x}$ such that*

$$\bar{x} \geq 0, \quad M\bar{x} + q \geq 0. \tag{6.17}$$

Proof. Put $\Delta = R^n_+$. Note that $(M\bar{x} + q)^T v \geq 0$ for every $v \in 0^+\Delta = R^n_+$ for some $\bar{x} \in \Delta$ if and only if there exists $\bar{x}$ satisfying (6.17). Applying Theorem 6.2 we obtain the desired conclusion. □

In the terminology of Cottle et al. (1992), if there exists $\bar{x} \in R^n$ satisfying (6.17) then problem LCP(M,q) is said to be *feasible*. The set of all $\bar{x} \in R^n$ satisfying (6.17) is called *the feasible region* of that problem. Corollary 6.1 asserts that *a linear complementarity problem with a positive semidefinite matrix M is solvable if and only it is feasible.*

6.2 Solution Existence under Copositivity

In this section we obtain some existence theorems for the AVI problem (6.1) where M is not assumed to be monotone on Δ. It is assumed only that M is copositive on Δ.

We first establish an existence theorem under strict copositivity.

Theorem 6.3. *If matrix M is strictly copositive on a nonempty polyhedral convex set Δ then, for any $q \in R^n$, problem* AVI(M,q,Δ) *has a solution.*

The following auxiliary fact shows that the strict copositivity assumption in the above theorem is, in fact, equivalent to a coercivity condition of the form (5.7).

Lemma 6.3. *Matrix $M \in R^{n\times n}$ is strictly copositive on a nonempty polyhedral convex set $\Delta \subset R^n$ if and only if there exists $x^0 \in \Delta$ such that*

$$\frac{\langle My - Mx^0, y - x^0\rangle}{\|y - x^0\|} \to +\infty \text{ as } \|y\| \to +\infty,\ y \in \Delta. \tag{6.18}$$

Proof. *Necessity:* Suppose that Δ is nonempty and M is strictly copositive on Δ. If $0^+\Delta = \{0\}$ then, according to Theorem 8.4 in Rockafellar (1970), Δ is compact. So, for an arbitrarily chosen $x^0 \in \Delta$, condition (6.18) is satisfied. Now consider the case where $0^+\Delta \neq \{0\}$. select any $x^0 \in \Delta$. We claim that (6.18) is valid. On the contrary, suppose that (6.18) is false. Then there must exist $\gamma > 0$ and a sequence $\{y^k\} \subset \Delta$ such that $\|y^k\| \to +\infty$ and

$$\frac{\langle My^k - Mx^0, y^k - x^0\rangle}{\|y^k - x^0\|} \le \gamma \quad \forall k \in N. \tag{6.19}$$

Since Δ is a nonempty polyhedral convex set, by Theorems 19.1 and 19.5 from Rockafellar (1970) one can find a compact set $K \subset \Delta$ such that

$$\Delta = K + 0^+\Delta.$$

Hence, for each $k \in N$ there exist $u^k \in K$ and $v^k \in 0^+\Delta$ such that $y^k = u^k + v^k$. It is easily seen that $\|v^k\| \to +\infty$. Therefore, without loss of generality we can assume that

$$u^k \to \bar{u}, \quad v^k \neq 0 \text{ for every } k \in N, \quad \frac{v^k}{\|v^k\|} \to \bar{v},$$

for some $\bar{u} \in \Delta$ and $\bar{v} \in 0^+\Delta$ with $\|\bar{v}\| = 1$. From (6.19) it follows that

$$\frac{\|v^k\|^{-2}\langle Mv^k + M(u^k - x^0), v^k + (u^k - x^0)\rangle}{\|v^k\|^{-1}\|v^k + u^k - x^0\|} \le \frac{1}{\|v^k\|}\gamma$$

for all $k \in N$. Letting $k \to \infty$, from the above inequality we obtain

$$\frac{\langle M\bar{v}, \bar{v}\rangle}{\|\bar{v}\|} \le 0,$$

which contradicts the assumed strict copositivity of M on Δ. We have thus proved that (6.18) is valid.

Sufficiency: Suppose that there exists $x^0 \in \Delta$ such that (6.18) is fulfilled. Let $v \in 0^+\Delta \setminus \{0\}$ be given arbitrarily. Since $y(t) :=$

$x^0 + tv \in \Delta$ for every $t > 0$ and $\|y(t)\| \to +\infty$ as $t \to +\infty$, substituting $y = y(t)$ into (6.18) gives

$$t\frac{\langle Mv, v\rangle}{\|v\|} \to +\infty \quad \text{as } t \to +\infty.$$

This implies that $v^TMv = \langle Mv, v\rangle > 0$. We have thus shown that M is strictly copositive on Δ. □.

Proof of Theorem 6.3.

Suppose that M is strictly copositive on Δ and $q \in R^n$ is an arbitrarily given vector. Consider the affine operator $\phi(x) = Mx + q$. Applying Lemma 6.3 we can assert that there exists $x^0 \in \Delta$ such that the coercivity condition (5.17) is satisfied. According to Theorem 5.2, problem VI(ϕ, Δ) has solutions. Since the latter is exactly the problem AVI(M, q, Δ), the desired conclusion follows. □

One can derive Theorem 6.3 directly from Theorem 5.1 without appealing to Theorem 5.2 and Lemma 6.3.

Another proof of Theorem 6.3.

Suppose that $\Delta \neq \emptyset$, M is strictly copositive on Δ, and $q \in R^n$ is given arbitrarily. Let $m \in N$, $A \in R^{m\times n}$ and $b \in R^m$ be such that Δ has the representation (6.2). Then $0^+\Delta = \{v \in R^n : Av \geq 0\}$ (see Rockafellar (1970), p. 62). Select a point $x^0 \in \Delta$. For each $k \in N$, we set

$$\begin{aligned}\Delta_k \;&= \Delta \cap \{x = (x_1, \ldots, x_n) \in R^n \,:\, x_i^0 - k \leq x_i \leq x_i^0 + k \\ &\qquad \text{for every } i = 1, 2, \ldots, n\}.\end{aligned} \tag{6.20}$$

It is clear that, for every $k \in N$, Δ_k is a nonempty, compact, polyhedral convex set. Given any $k \in N$, we consider the problem AVI(M, q, Δ_k). According to the Hartman-Stampacchia Theorem (see Theorem 5.1), Sol(AVI$(M, q, \Delta_k)) \neq \emptyset$. For each $k \in N$, select a point $x^k \in$ Sol(AVI(M, q, Δ_k)). We claim that the sequence $\{x^k\}$ is bounded. To obtain a contradiction, suppose that $\{x^k\}$ is unbounded. Without restriction of generality we can assume that $x^k \neq 0$ for all k, $\|x^k\| \to +\infty$ as $k \to \infty$, and there exists $\bar{v} \in R^n$ such that

$$\frac{x^k}{\|x^k\|} \to \bar{v}, \quad \|\bar{v}\| = 1.$$

Since $x^0 \in \Delta_k$ for every $k \in N$, we have

$$\langle Mx^k + q, x^0 - x^k \rangle \geq 0 \quad \forall k \in N$$

or, equivalently,

$$\langle Mx^k + q, x^0 \rangle \geq \langle Mx^k + q, x^k \rangle \quad \forall k \in N. \tag{6.21}$$

Dividing the inequality in (6.21) by $\|x^k\|^2$ and letting $k \to \infty$ we get

$$0 \geq \langle M\bar{v}, \bar{v} \rangle. \tag{6.22}$$

Since $x^k \in \Delta$, we have $Ax^k \geq b$. Dividing the last inequality by $\|x^k\|$ and taking the limit as $k \to \infty$ we obtain $A\bar{v} \geq 0$. This shows that $\bar{v} \in 0^+\Delta$. Since $\|\bar{v}\| = 1$, (6.22) contradicts the assumed strict copositivity of M on Δ. We have thus proved that the sequence $\{x^k\}$ is bounded. There is no loss of generality in assuming that $x^k \to \bar{x}$ for some $\bar{x} \in \Delta$. For each $x \in \Delta$ one can find and index $k_x \in N$ such that $x \in \Delta_k$ for all $k \geq k_x$. Consequently, for every $k \geq k_x$, it holds

$$\langle Mx^k + q, x - x^k \rangle \geq 0.$$

Letting $k \to \infty$ we obtain

$$\langle M\bar{x} + q, x - \bar{x} \rangle \geq 0.$$

Since the last inequality is valid for any $x \in \Delta$, we conclude that $\bar{x} \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$. The proof is complete. □

Example 6.2. Let

$$M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in R^{2\times 2}, \quad \Delta = R^2_+.$$

In Remark 6.1 we have observed that M is not monotone on Δ. However, M is strictly copositive on Δ. Indeed, since Δ is a cone, we have $0^+\Delta = \Delta$. For any nonzero vector $v = (v_1, v_2) \in 0^+\Delta = R^2_+$ it holds

$$v^T M v = v_1^2 + 4v_1v_2 + v_2^2 > 0.$$

This shows that M is strictly copositive on Δ. According to Theorem 6.3, for any $q \in R^2$, problem $\mathrm{AVI}(M, q, \Delta)$ is solvable. Note that Theorem 6.1 cannot be applied to this problem because M is not monotone on Δ.

In the following existence theorem we need not to assume that the matrix M strictly copositive on Δ. Instead of the strict copositivity, a weaker assumption is employed.

Theorem 6.4. *If matrix M is copositive on a nonempty polyhedral convex set Δ and there exists no $\bar{v} \in R^n \setminus \{0\}$ such that*

$$\bar{v} \in 0^+\Delta, \quad M\bar{v} \in (0^+\Delta)^+, \quad \bar{v}^T M\bar{v} = 0, \tag{6.23}$$

where

$$(0^+\Delta)^+ = \{\xi \in R^n \,:\, \xi^T v \geq 0 \ \ \forall v \in 0^+\Delta\}$$

then, for any $q \in R^n$, problem AVI(M, q, Δ) *has a solution.*

Proof. Suppose that $\Delta \neq \emptyset$, M is copositive on Δ and there exists no $\bar{v} \in R^n \setminus \{0\}$ satisfying (6.23). Suppose that $q \in R^n$ is given arbitrarily. Let $m \in N$, $A \in R^{m\times n}$ and $b \in R^m$ be such that Δ has the representation (6.2). Then $0^+\Delta = \{v \in R^n \,:\, Av \geq 0\}$. Let $x^0 \in \Delta$. For each $k \in N$, we define

$$\Delta_k = \Delta \cap \bar{B}(x^0, k). \tag{6.24}$$

Note that, for every $k \in N$, Δ_k is a nonempty, compact, convex set. Given any $k \in N$, we consider the VI problem

$$\text{Find } \ x \in \Delta_k \ \text{ such that } \ \langle Mx + q, y - x\rangle \geq 0 \quad \forall y \in \Delta_k$$

and denote its solution set by Sol(VI((M, q, Δ_k))). By Theorem 5.1, Sol(VI((M, q, Δ_k))) $\neq \emptyset$. For each $k \in N$, select a point $x^k \in$ Sol(VI((M, q, Δ_k))). We claim that the sequence $\{x^k\}$ is bounded. Suppose, contrary to our claim, that $\{x^k\}$ is unbounded. There is no loss of generality in assuming that $x^k \neq 0$ for all k, $\|x^k\| \to +\infty$ as $k \to \infty$, and there exists $\bar{v} \in R^n$ such that

$$\frac{x^k}{\|x^k\|} \to \bar{v}, \quad \|\bar{v}\| = 1.$$

Since $x^0 \in \Delta_k$ for every $k \in N$, we have

$$\langle Mx^k + q, x^0 - x^k\rangle \geq 0 \quad \forall k \in N.$$

As in the second proof of Theorem 6.3, from the last property we deduce that

$$0 \geq \langle M\bar{v}, \bar{v}\rangle.$$

Since $x^k \in \Delta$, we have $Ax^k \geq b$ for every $k \in N$. This implies $A\bar{v} \geq 0$. So $\bar{v} \in 0^+\Delta$. Since M is copositive on Δ, from what has already been said it follows that

$$\bar{v}^T M \bar{v} = 0. \tag{6.25}$$

For any $w \in 0^+\Delta \setminus \{0\}$, from (6.24) and the fact that $x^0 + tw \in \Delta$ for every $t \geq 0$ we deduce that

$$x^0 + \|x^k - x^0\| \frac{w}{\|w\|} \in \Delta_k.$$

Since $x^k \in \mathrm{Sol}(\mathrm{VI}((M, q, \Delta_k))$, we have

$$\left\langle Mx^k + q, x^0 + \|x^k - x^0\| \frac{w}{\|w\|} - x^k \right\rangle \geq 0.$$

Dividing this inequality by $\|x^k\|^2$, letting $k \to \infty$ and noting that $\lim_{k\to\infty} \dfrac{\|x^k - x^0\|}{\|x^k\|} \to 1$, by virtue of (6.25) we obtain $\left\langle M\bar{v}, \dfrac{w}{\|w\|} \right\rangle \geq 0$. Hence $\langle M\bar{v}, w\rangle \geq 0$ for every $w \in 0^+\Delta$. This means that $M\bar{v} \in (0^+\Delta)^+$. We see that vector $\bar{v} \in R^n \setminus \{0\}$ satisfies all the three conditions described in (6.23). This contradicts our assumption. We have thus proved that the sequence $\{x^k\}$ is bounded. Without loss of generality we can assume that $x^k \to \bar{x}$ for some $\bar{x} \in \Delta$. For each $x \in \Delta$ there exists $k_x \in N$ such that $x \in \Delta_k$ for all $k \geq k_x$. Consequently, for every $k \geq k_x$, we have $\langle Mx^k + q, x - x^k\rangle \geq 0$. Letting $k \to \infty$ we obtain $\langle M\bar{x}+q, x-\bar{x}\rangle \geq 0$. Since this inequality holds for any $x \in \Delta$, we can assert that $\bar{x} \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$. The proof is complete. □

Example 6.3. Let M and Δ be the same as in Example 6.1. It is a simple matter to verify that there exists no $\bar{v} \in R^n \setminus \{0\}$ satisfying the three conditions in (6.23). Since M is copositive on Δ, Theorem 6.4 asserts that, for any $q = (q_1, q_2) \in R^2$, problem $\mathrm{AVI}(M, q, \Delta)$ has a solution.

In the sequel, sometimes we shall use the following simple fact.

Lemma 6.4. *Let $K \subset R^n$ be a nonempty closed cone. Let $q \in R^n$. Then $q \in \mathrm{int}K^+$, where $\mathrm{int}K^+$ denotes the interior of the positive dual cone K^+ of K, if and only if*

$$v^T q > 0 \quad \forall v \in K \setminus \{0\}. \tag{6.26}$$

Proof. Suppose that $q \in \text{int}K^+$. If there exists $\bar{v} \in K \setminus \{0\}$ such that $\bar{v}^T q \le 0$ then $\bar{v}^T q = 0$ because the condition $q \in K^+$ implies that $v^T q \ge 0$ for every $v \in K$. From this we see that the linear functional $\xi \to \bar{v}^T \xi$ achieves its global minimum on K^+ at q. As $q \in \text{int}K^+$, there exists $\varepsilon > 0$ such that $\bar{B}(q, \varepsilon) \subset K^+$. Then

$$\bar{v}^T \xi \ge 0 \quad \forall \xi \in \bar{B}(q, \varepsilon).$$

This implies that $\bar{v} = 0$, a contradiction. We have thus proved that if $q \in \text{int}K^+$ then (6.26) is valid.

Conversely, assume that (6.26) holds. To obtain a contradiction, suppose that $q \notin \text{int}K^+$. Then there exists a sequence $\{q^k\}$ in $R^n \setminus K^+$ such that $q^k \to q$. Consequently, for each $k \in N$ there exists $v^k \in K$ such that $(v^k)^T q^k < 0$. Without loss of generality we can assume that $\dfrac{v^k}{\|v^k\|} \to \bar{v}$ with $\|\bar{v}\| = 1$. We have

$$\frac{(v^k)^T q^k}{\|v^k\|} < 0, \quad \frac{v^k}{\|v^k\|} \in K \quad \forall k \in N.$$

Taking the limits as $k \to \infty$ we obtain $\bar{v}^T q \le 0$ and $\bar{v} \in K$, contrary to (6.26). □

In the case where Δ is a cone, we have the following existence theorem.

Theorem 6.5. *Assume that Δ is a polyhedral convex cone. If matrix M is copositive on Δ and*

$$q \in \text{int}([\text{Sol}(\text{AVI}(M, 0, \Delta))]^+), \tag{6.27}$$

then problem $\text{AVI}(M, q, \Delta)$ *has a solution.*

Note that $\text{AVI}(M, q, \Delta)$ is a generalized linear complementarity problem (see Definition 5.8). From the definition it follows that $v \in \text{Sol}(\text{AVI}(M, 0, \Delta))$ if and only if

$$v \in \Delta, \quad Mv \in \Delta^+, \quad v^T M v = 0.$$

Hence, applying Lemma 6.4 to the cone $K := \text{Sol}(\text{AVI}(M, 0, \Delta))$ we see that condition (6.27) is equivalent to the requirement that there exists no $\bar{v} \in R^n \setminus \{0\}$ such that

$$\bar{v} \in \Delta, \quad M\bar{v} \in \Delta^+, \quad \bar{v}^T M \bar{v} = 0, \quad q^T \bar{v} \le 0. \tag{6.28}$$

Proof of Theorem 6.5.

Suppose that Δ is a polyhedral convex cone, M is copositive on Δ, and q is such that (6.27) holds. For each $k \in N$, we set

$$\Delta_k = \Delta \cap \{x \in R^n : -k \leq x_i \leq k \text{ for every } i = 1, 2, \ldots, n\}.$$

It is clear that, for each $k \in N$, $0 \in \Delta_k$ and Δ_k is a compact, polyhedral convex set. Consider the problem AVI(M, q, Δ_k). By Theorem 5.1, we can find a point $x^k \in$ Sol(AVI((M, q, Δ_k))). If the sequence $\{x^k\}$ is unbounded then without loss of generality we can assume that $x^k \neq 0$ for all k, $\|x^k\| \to +\infty$ as $k \to \infty$, and there exists $\bar{v} \in R^n$ such that

$$\frac{x^k}{\|x^k\|} \to \bar{v}, \quad \|\bar{v}\| = 1.$$

Since $0 \in \Delta_k$, we have

$$\langle Mx^k + q, 0 - x^k \rangle \geq 0.$$

Hence

$$-q^T x^k \geq (x^k)^T M x^k \quad (\forall k \in N). \tag{6.29}$$

Dividing the inequality in (6.29) by $\|x^k\|^2$ and taking limits as $k \to \infty$ we get

$$0 \geq \bar{v}^T M \bar{v}. \tag{6.30}$$

It is clear that $\bar{v} \in \Delta$. Since M is copositive on Δ, we have $v^T M v \geq 0$ for every $v \in \Delta$. Combining this fact with (6.30) yields

$$\bar{v}^T M \bar{v} = 0. \tag{6.31}$$

From (6.29) and the copositivity of M on Δ it follows that $-q^T x^k \geq 0$ for every $k \in N$. This implies that

$$-q^T \bar{v} \geq 0. \tag{6.32}$$

Fix any $w \in \Delta \setminus \{0\}$. It is evident that

$$\|x^k\| \frac{w}{\|w\|} \in \Delta.$$

Since $x^k \in$ Sol(AVI(M, q, Δ_k)), we have

$$\left\langle Mx^k + q, \|x^k\| \frac{w}{\|w\|} - x^k \right\rangle \geq 0 \quad (\forall k \in N).$$

From this and (6.31) we deduce that $\langle M\bar{v}, w\rangle \geq 0$. Since the last inequality is valid for every $w \in \Delta \setminus \{0\}$, we see that $M\bar{v} \in \Delta^+$. Combining this with (6.31) and (6.32) we can assert that (6.28) is satisfied. Then (6.27) is false. We have arrived at a contradiction. Thus the sequence $\{x^k\}$ must be bounded. Analysis similar to that in the final part of the proof of Theorem 6.4 shows that problem AVI(M, q, Δ) has a solution. □

Example 6.4. Let

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in R^{2\times 2}, \quad q = (1,1) \in R^2,$$
$$\Delta = \{x = (x_1, x_2) \in R^2 \, : \, x_1 \geq 0, \ x_2 \geq 0, \ x_1 = x_2\}.$$

Theorem 6.5 can be applied to the problem AVI(M, q, Δ). Indeed, we have

$$v^T Mv = v_1^2 - v_2^2 = 0 \quad \forall v \in 0^+\Delta = \Delta.$$

This shows that M is copositive on Δ. Furthermore, we have

$$\begin{aligned}\mathrm{Sol}(\mathrm{AVI}(M, 0, 0^+\Delta)) &= \{v \in R^2 \, : \, v \in 0^+\Delta, \ \ Mv \in (0^+\Delta)^+, \\ &\qquad v^T Mv = 0\} = \Delta.\end{aligned}$$

Therefore

$$q^T v = v_1 + v_2 > 0 \quad \forall v = (v_1, v_2) \in \mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta)) \setminus \{0\}.$$

So (6.27) is satisfied. By Theorem 6.5, problem AVI(M, q, Δ) is solvable. In fact, we have

$$\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) = \{(0, 0)\}.$$

It is worth pointing that, since M is not strictly copositive on Δ, Theorem 6.3 cannot be applied to this problem. Since all the three conditions described in (6.23) are satisfied if one chooses $\bar{v} = (1,1) \in R^2 \setminus \{0\}$, Theorem 6.4 also cannot be applied to this problem.

Remark 6.4. In the case where Δ is a polyhedral convex cone, the conclusion of Theorem 6.4 follows from Theorem 6.5. Indeed, in this case, under the assumption of Theorem 6.4 we have

$$\mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta)) = \{0\}.$$

Hence $[\mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta))]^+ = R^n$. So (6.27) is satisfied for any $q \in R^n$. By Theorem 6.5, problem AVI(M, q, Δ) is solvable.

Applying Theorem 6.5 to LCP problems we obtain the following corollary.

Corollary 6.2. *If M is a copositive matrix and*

$$q \in \mathrm{int}([\mathrm{Sol}(M, 0)]^+), \tag{6.33}$$

then the problem LCP(M, q) *has a solution.*

Note that condition (6.33) is stronger than condition (6.34) in the following existence theorem for LCP problems.

Theorem 6.6. (See Cottle et al. (1992), Theorem 3.8.6) *If M is a copositive matrix and*

$$q \in [\mathrm{Sol}(M, 0)]^+, \tag{6.34}$$

then problem LCP(M, q) *has a solution.*

It is clear that (6.34) can be rewritten in the following form:

$$[v \in R^n,\ \ v \geq 0,\ \ Mv \geq 0,\ \ v^TMv = 0] \implies [q^Tv \geq 0].$$

Meanwhile, by Lemma 6.4, condition (6.33) is equivalent to the following one:

$$[v \in R^n \setminus \{0\},\ \ v \geq 0,\ \ Mv \geq 0,\ \ v^TMv = 0] \implies [q^Tv > 0].$$

In connection with Theorems 6.5, the following open question seems to be interesting.

QUESTION: Whether the conclusion of Theorem 6.5 is still valid if in the place of (6.27) one uses the following weaker condition

$$q \in [\mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta))]^+?$$

Note that the last inclusion can be rewritten in the form:

$$[v \in R^n,\ \ v \in 0^+\Delta,\ \ Mv \in \Delta^+,\ \ v^TMv = 0] \implies [q^Tv \geq 0].$$

6.3 Commentaries

In this chapter, we have considered a variety of solution existence theorems for affine variational inequalities. Here the compactness of the constraint set Δ is not assumed. But we have to employ a monotonicity property of the matrix M with respect to Δ. Namely, we have had deal with the monotonicity, the strict monotonicity, and the copositivity of M w.r.t. Δ.

The interested reader is referred to Gowda and Pang (1994a) for an insightful study on existence theorems for AVI problems.

Chapter 7

Upper-Lipschitz Continuity of the Solution Map in Affine Variational Inequalities

In this chapter we shall discuss two fundamental theorems due to Robinson (1979, 1981) on the upper-Lipschitz continuity of the solution map in affine variational inequality problems. The theorem on the upper-Lipschitz continuity of the solution map in linear complementarity problems due to Cottle et al. (1992) is also studied in this chapter. The Walkup-Wets Theorem (see Walkup and Wets (1969)), which we analyze in Section 7.1, is the basis for obtaining these results.

7.1 The Walkup-Wets Theorem

Let $\Delta \subset R^n$ be a nonempty subset. Let $\tau : R^n \to R^m$ be an affine operator; that is there exist a linear operator $A : R^n \to R^m$ and a vector $b \in R^m$ such that $\tau(x) = Ax + b$ for every $x \in R^n$. Define

$$\begin{aligned}\Delta(y) = \tau^{-1}(y) \cap \Delta \ &= \{x \in \Delta \, : \, \tau(x) = y\} \\ &= \{x \in \Delta \, : \, Ax + b = y\}.\end{aligned} \tag{7.1}$$

Definition 7.1. (See Walkup and Wets (1969), Definition 1) A subset $\Delta \subset R^n$ is said to have *property* $\mathcal{L}_j$ if for every affine operator $\tau : R^n \to R^m$, $m \in N$, with $\dim(\ker(\tau)) = j$, the inverse mapping

$y \to \Delta(y)$ is Lipschitz on its effective domain. This means that there exists a constant $\ell > 0$ such that

$$\Delta(y') \subset \Delta(y) + \ell\|y' - y\|\bar{B}_{R^n} \quad \text{whenever } \Delta(y) \neq \emptyset, \ \Delta(y') \neq \emptyset. \tag{7.2}$$

In the above definition, $\dim(\ker(\tau))$ denotes the dimension of the affine set

$$\ker(\tau) = \{x \in R^n \, : \, \tau(x) = 0\}.$$

The following theorem is a key tool for proving other results in this chapter.

Theorem 7.1 (The Walkup-Wets Theorem; see Walkup and Wets (1969), Theorem 1). *Let $\Delta \subset R^n$ be a nonempty closed convex set and let $j \in N$, $1 \le j \le n-1$. Then Δ is a polyhedral convex set if and only if it has property $\mathcal{L}_j$.*

In the sequel, we will use only one assertion of this theorem: *If Δ is a polyhedral convex set, then it has property $\mathcal{L}_j$.* A detailed proof of this assertion can be found in Mangasarian and Shiau (1987).

Corollary 7.1. *If $\Delta \subset R^n$ is a polyhedral convex set and if $\tau : R^n \to R^m$ is an affine operator, then there exists a constant $\ell > 0$ such that* (7.2)*, where $\Delta(y)$ is defined by* (7.1) *for all $y \in R^n$, holds.*

Proof. If $j := \dim(\ker(\tau))$ satisfies the condition $1 \le j \le n-1$, then the conclusion is immediate from Theorem 7.1. If $\dim(\ker(\tau)) = n$ then $\ker(\tau) = R^n$, and we have

$$\Delta(y) = \tau^{-1}(y) \cap \Delta = \begin{cases} \Delta & \text{if } y = 0, \\ \emptyset & \text{if } y \neq 0. \end{cases}$$

This shows that (7.2) is fulfilled with any $\ell > 0$. We now suppose that $\dim(\ker(\tau)) = 0$. Let $\tau(x) = Ax + b$, where $A : R^n \to R^m$ is a linear operator and $b \in R^m$. Since τ is an injective mapping, $Y := \tau(R^n)$ is an affine set in R^m with $\dim Y = n$, and that $n \le m$. Likewise, the set $Y_0 := A(R^n)$ is a linear subspace of R^m with $\dim Y_0 = n$. Let $\widetilde{A} : R^n \to Y_0$ be the linear operator defined by setting $\widetilde{A}x = Ax$ for every $x \in R^n$. It is easily shown that

$$\|\tau^{-1}(y') - \tau^{-1}(y)\| \le \|\widetilde{A}^{-1}\|\|y' - y\|$$

for every $y \in Y$ and $y' \in Y$. From this we deduce that (7.2) is satisfied with $\ell := \|\widetilde{A}^{-1}\|$. □

Remark 7.1. Under the assumptions of Corollary 7.1, for every $y \in R^m$, $\Delta(y)$ is a polyhedral convex set (possibly empty).

Remark 7.2. The conclusion of Theorem 7.1 is not true if one chooses $j = 0$. Namely, the arguments described in the final part of the proof of Corollary 7.1 show that any nonempty set $\Delta \subset R^n$ has property $\mathcal{L}_0$. Similarly, the conclusion of Theorem 7.1 is not valid if $j = n$.

Corollary 7.2. *For any nonempty polyhedral convex set $\Delta \subset R^n$ and any matrix $C \in R^{s\times n}$ there exists a constant $\ell > 0$ such that*

$$\Delta(C, d'') \subset \Delta(C, d') + \ell\|d'' - d'\|\bar{B}_{R^n} \tag{7.3}$$

whenever $\Delta(C, d')$ and $\Delta(C, d'')$ are nonempty; where

$$\Delta(C, d) := \{x \in \Delta \,:\, Cx = d\}$$

for every $d \in R^s$.

Proof. Set $\tau(x) = Cx$. Since

$$\Delta(C, y) = \tau^{-1}(y) \cap \Delta = \Delta(y)$$

where $\Delta(y)$ is defined by (7.1), applying Corollary 7.1 we can find $\ell > 0$ such that the Lipschitz continuity property stated in (7.3) is satisfied. $\square$

Corollary 7.3. *For any nonempty polyhedral convex set $\Delta \subset R^n$, any matrix $A \in R^{m\times n}$ and matrix $C \in R^{s\times n}$ there exists a constant $\ell > 0$ such that*

$$\Delta(A, C, b'', d'') \subset \Delta(A, C, b', d') + \ell(\|b'' - b'\| + \|d'' - d'\|)\bar{B}_{R^n} \tag{7.4}$$

whenever $\Delta(A, C, b', d')$ and $\Delta(A, C, b'', d'')$ are nonempty; where

$$\Delta(A, C, b, d) := \{x \in \Delta \,:\, Ax \geq b,\ Cx = d\}$$

for every $b \in R^m$ and $d \in R^s$.

Proof. Define

$$\widetilde{C} = \begin{pmatrix} A & -E \\ C & 0 \end{pmatrix} \in R^{(m+s)\times(n+m)},$$

where E denotes the unit matrix in $R^{m\times m}$ and 0 denotes the null in $R^{s\times m}$. Let

$$\widetilde{\Delta} = \{(x, w) \in R^n \times R^m \,:\, x \in \Delta,\ w \geq 0\}.$$

By Corollary 7.2, there exists $\ell > 0$ such that

$$\widetilde{\Delta}(\widetilde{C}, b'', d'') \subset \widetilde{\Delta}(\widetilde{C}, b', d') + \ell(\|b'' - b'\| + \|d'' - d'\|)\bar{B}_{R^{n+m}} \quad (7.5)$$

whenever $\widetilde{\Delta}(\widetilde{C}, b', d') \neq \emptyset$ and $\widetilde{\Delta}(\widetilde{C}, b'', d'') \neq \emptyset$, where

$$\widetilde{\Delta}(\widetilde{C}, b, d) := \left\{(x, w) \in \widetilde{\Delta} \, : \, \widetilde{C}\begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}\right\}.$$

Since

$$\begin{aligned} \Delta(A, C, b, d) &= \{x \in \Delta \, : \, \exists w \in R^m, \; w \geq 0, \; Ax - w = b, \; Cx = d\} \\ &= \mathrm{Pr}_{R^n}(\widetilde{\Delta}(\widetilde{C}, b, d)) \end{aligned}$$

where $\mathrm{Pr}_{R^n}(x, w) = x$ for every $(x, w) \in R^n \times R^m$, we see at once that (7.5) implies (7.4). □

7.2 Upper-Lipschitz Continuity with respect to Linear Variables

The notion of polyhedral multifunction was proposed by Robinson (see Robinson (1979, 1981). We now study several basic facts concerning polyhedral multifunctions.

Definition 7.2. If $\Phi : R^n \to 2^{R^m}$ is a multifunction then its *graph* and *effective domain* are defined, respectively, by setting

$$\begin{aligned} \mathrm{graph}\Phi &= \{(x, y) \in R^n \times R^m \, : \, y \in \Phi(x)\}, \\ \mathrm{dom}\Phi &= \{x \in R^n \, : \, \Phi(x) \neq \emptyset\}. \end{aligned}$$

Definition 7.3. A set-valued mapping $\Phi : R^n \to 2^{R^m}$ is called a *polyhedral multifunction* if its graph can be represented as the union of finitely many polyhedral convex sets in $R^n \times R^m$.

The following statement shows that the *normal-cone operator* corresponding to a polyhedral convex set is a polyhedral multifunction.

Proposition 7.1. (See Robinson (1981)) *Suppose that $\Delta \subset R^n$ is a nonempty polyhedral convex set. Then the formula*

$$\Phi(x) = N_\Delta(x) \quad (x \in R^n)$$

defines a polyhedral multifunction $\Phi : R^n \to 2^{R^n}$.

Proof. Let $m \in N$, $A \in R^{n\times n}$ and $b \in R^m$ be such that $\Delta = \{x \in R^n : Ax \geq b\}$. Set $I = \{1, \ldots, m\}$. Let

$$F_\alpha = \{x \in R^n : A_\alpha x = b_\alpha, \;\; A_{I\setminus\alpha} x > b_{I\setminus\alpha}\}$$

be the pseudo-face of Δ corresponding to an index set $\alpha \subset I$. For every $x \in F_\alpha$ we have

$$T_\Delta(x) = \{v \in R^n : A_\alpha v \geq 0\}.$$

(See the proof of Theorem 4.2.) Since

$$N_\Delta(x) = \{\xi \in R^n : \langle \xi, v\rangle \leq 0 \;\; \forall v \in T_\Delta(x)\},$$

we have $\xi \in N_\Delta(x)$ if and only if the inequality $\langle \xi, v\rangle \leq 0$ is a consequence of the inequality system $A_\alpha v \geq 0$. Consequently, applying Farkas' Lemma (see Theorem 3.2) we deduce that $\xi \in N_\Delta(x)$ if and only if there exist $\lambda_1 \geq 0, \ldots, \; \lambda_m \geq 0$ such that

$$\xi = \sum_{i\in\alpha} \lambda_i(-A_i^T),$$

where A_i denotes the $i-$th row of matrix A. (Note that if $\alpha = \emptyset$ and $x \in F_\alpha$, then $x \in \text{int}\Delta$; hence $\xi = 0$ for every $\xi \in N_\Delta(x)$.) Define

$$\Omega_\alpha = \{(x, \xi) \in R^n \times R^n : x \in F_\alpha, \; \xi \in N_\Delta(x)\}.$$

Obviously, $\Omega_\alpha \subset \text{graph}\Phi$. Note that

$$\begin{aligned}\Omega_\alpha \; &= \{(x, \xi) \in R^n \times R^n : A_\alpha x = b_\alpha, \; A_{I\setminus\alpha} x > b_{I\setminus\alpha}, \\ &\qquad \xi = \textstyle\sum_{i\in\alpha} \lambda_i(-A_i^T) \;\; \text{for some} \;\; \lambda_\alpha \in R_+^{|\alpha|}\}.\end{aligned}$$

is a convex set. Here $|\alpha|$ denotes the number of elements in α. It is easily seen that the topological closure $\overline{\Omega}_\alpha$ of Ω_α is given by the formula

$$\begin{aligned}\overline{\Omega}_\alpha \; &= \{(x, \xi) \in R^n \times R^n : A_\alpha x = b_\alpha, \; A_{I\setminus\alpha} x \geq b_{I\setminus\alpha}, \\ &\qquad \xi = \textstyle\sum_{i\in\alpha} \lambda_i(-A_i) \;\; \text{for some} \;\; \lambda_\alpha \in R_+^{|\alpha|}\} \\ &= \text{Pr}_{R^n\times R^n}\{(x, \xi, \lambda_\alpha) \in R^n \times R^n \times R_+^{|\alpha|} : A_\alpha x = b_\alpha, \\ &\qquad A_{I\setminus\alpha} x \geq b_{I\setminus\alpha}, \; \textstyle\sum_{i\in\alpha} \lambda_i A_i^T + \xi = 0\},\end{aligned}$$

where $\text{Pr}_{R^n\times R^n}(x, \xi, \lambda_\alpha) = (x, \xi)$. It is clear that the set in the last curly brackets is a polyhedral convex set. From this fact, the above formula for $\overline{\Omega}_\alpha$ and Theorem 19.1 in Rockafellar (1970) we deduce that $\overline{\Omega}_\alpha$ is a polyhedral convex set (see the proof of Theorem 4.3). Since $\Delta = \bigcup_{\alpha\subset I} F_\alpha$, we have

$$\text{graph}\Phi = \bigcup_{\alpha\subset I} \Omega_\alpha. \tag{7.6}$$

Observe that graphΦ is a closed set. Indeed, suppose that $\{(x^k, \xi^k)\}$ is a sequence satisfying $(x^k, \xi^k) \to (\bar{x}, \bar{\xi}) \in R^n \times R^n$, and $(x^k, \xi^k) \in$ graphΦ for every $k \in N$. On account of formula (1.12), we have

$$\langle \xi^k, y - x^k \rangle \leq 0 \quad \forall y \in \Delta,\ \forall k \in N.$$

Fixing any $y \in \Delta$ and taking limit as $k \to \infty$, from the last inequality we obtain $\langle \bar{\xi}, y - \bar{x} \rangle \leq 0$. Since this inequality holds for each $y \in \Delta$, we see that $\bar{\xi} \in N_\Delta(\bar{x})$. Hence $(\bar{x}, \bar{\xi}) \in$ graphΦ. We have thus proved that the set graphΦ is closed. On account of this fact, from (7.6) we deduce that

$$\text{graph}\Phi = \bigcup_{\alpha\subset I} \overline{\Omega}_\alpha.$$

This shows that graphΦ can be represented as the union of finitely many polyhedral convex sets. The proof is complete. □

The following statement shows that the solution map of a parametric affine variational inequality problem is a polyhedral multifunction (on the linear variables of the problem).

Proposition 7.2. *Suppose that $M \in R^{n\times n}$, $A \in R^{m\times n}$ and $C \in R^{s\times n}$ are given matrices. Then the formula*

$$\Phi(q, b, d) = \text{Sol}(\text{AVI}(M, q, \Delta(b, d))),$$

where $(q, b, d) \in R^n \times R^m \times R^s$, $\Delta(b, d) := \{x \in R^n : Ax \geq b,\ Cx = d\}$ and $\text{Sol}(\text{AVI}(M, q, \Delta(b, d)))$ denotes the solution set of problem (6.1) with $\Delta = \Delta(b, d)$, defines a polyhedral multifunction

$$\Phi : R^n \times R^m \times R^s \to 2^{R^n}.$$

Proof. According to Corollary 5.2, $x \in \text{Sol}(\text{AVI}(M, q, \Delta(b, d)))$ if and only if there exist $\lambda = (\lambda_1, \ldots, \lambda_m) \in R^m$ and $\mu = (\mu_1, \ldots, \mu_s) \in R^s$ such that

$$\begin{cases} Mx - A^T\lambda - C^T\mu + q = 0, \\ Ax \geq b,\ Cx = d,\ \lambda \geq 0, \\ \lambda^T(Ax - b) = 0. \end{cases} \tag{7.7}$$

Let $I = \{1, \dots, m\}$. For each index set $\alpha \subset I$, we define

$$Q_\alpha = \mathrm{Pr}_1\Big(\{(x,q,b,d,\lambda,\mu) \,:\, Mx - A^T\lambda - C^T\mu + q = 0, \\ A_\alpha x = b_\alpha,\ A_{I\setminus\alpha}x \geq b_{I\setminus\alpha}, \\ Cx = d,\ \lambda_\alpha \geq 0,\ \lambda_{I\setminus\alpha} = 0\}\Big), \tag{7.8}$$

where

$$\mathrm{Pr}_1(x,q,b,d,\lambda,\mu) = (x,q,b,d)$$

for all $(x,q,b,d,\lambda,\mu) \in R^n \times R^n \times R^m \times R^s \times R^m \times R^s$. Hence Q_α is a polyhedral convex set. Note that

$$\mathrm{graph}\Phi = \bigcup_{\alpha\subset I} Q_\alpha. \tag{7.9}$$

Indeed, for each $(x,q,b,d) \in \mathrm{graph}\Phi$ we have

$$x \in \mathrm{Sol}(\mathrm{AVI}(M,q,\Delta(b,d))).$$

So there exist $\lambda = (\lambda_1, \dots, \lambda_m) \in R^m$ and $\mu = (\mu_1, \dots, \mu_s) \in R^s$ satisfying (7.7). Let $\alpha = \{i \in I \,:\, A_i x = b_i\}$. For every $i \in I \setminus \alpha$, we have $A_i x > b_i$. Then from the equality $\lambda_i(A_i x - b_i) = 0$ we deduce that $\lambda_i = 0$ for every $i \in I \setminus \alpha$. On account of this remark, we see that (x,q,b,d,λ,μ) satisfies all the conditions described in the curly braces in formula (7.8). This implies that $(x,q,b,d) \in Q_\alpha$. We thus get

$$\mathrm{graph}\Phi \subset \bigcup_{\alpha\subset I} Q_\alpha.$$

Since the reverse inclusion is obvious, we obtain formula (7.9), which shows that graphΦ can be represented as the union of finitely many polyhedral convex sets. □

Theorem 7.2. (See Robinson (1981), Proposition 1) *If $\Phi : R^n \to 2^{R^m}$ is a polyhedral multifunction, then there exists a constant $\ell > 0$ such that for every $\bar{x} \in R^n$ there is a neighborhood $U_{\bar{x}}$ of $\bar{x}$ satisfying*

$$\Phi(x) \subset \Phi(\bar{x}) + \ell\|x - \bar{x}\|\bar{B}_{R^m} \quad \forall x \in U_{\bar{x}}. \tag{7.10}$$

Definition 7.4. (See Robinson (1981)) Suppose that $\Phi : R^n \to 2^{R^m}$ is a multifunction and $\bar{x} \in R^n$ is a given point. If there exist $\ell > 0$ and a neighborhood $U_{\bar{x}}$ of $\bar{x}$ such that property (7.10) is valid, then Φ is said to be *locally upper-Lipschitz* at $\bar{x}$ with the Lipschitz constant ℓ.

The locally upper-Lipschitz property is weaker than the locally Lipschitz property which is described as follows.

Definition 7.5. A multifunction $\Phi : R^n \to 2^{R^m}$ is said to be *locally Lipschitz* at $\bar{x} \in R^n$ if there exist a constant $\ell > 0$ and a neighborhood $U_{\bar{x}}$ of $\bar{x}$ such that

$$\Phi(x) \subset \Phi(u) + \ell\|x - u\|\bar{B}_{R^m} \quad \forall x \in U_{\bar{x}}, \forall u \in U_{\bar{x}}.$$

If there exists a constant $\ell > 0$ such that

$$\Phi(x) \subset \Phi(u) + \ell\|x - u\|\bar{B}_{R^m}$$

for all x and u from a subset $\Omega \subset R^n$, then Φ is said to be *Lipschitz* on Ω.

From Theorem 7.2 it follows that if Φ is a polyhedral multifunction then it is locally upper-Lipschitz at any point in R^n with the same Lipschitz constant. Note that the *diameter* $\mathrm{diam}U_{\bar{x}} := \sup\{\|y - x\| : x \in U_{\bar{x}},\ y \in U_{\bar{x}}\}$ of neighborhood $U_{\bar{x}}$ depends on $\bar{x}$ and it can change greatly from one point to another.

Proof of Theorem 7.2.

Since Φ is a polyhedral multifunction, there exist nonempty polyhedral convex sets $Q_j \subset R^n \times R^m$ $(j = 1, \ldots, k)$ such that

$$\mathrm{graph}\Phi = \bigcup_{j \in J} Q_j, \tag{7.11}$$

where $J = \{1, \ldots, k\}$. For each $j \in J$ we consider the multifunction $\Phi_j : R^n \to 2^{R^m}$ defined by setting

$$\Phi_j(x) = \{y \in R^m : (x, y) \in Q_j\}. \tag{7.12}$$

Obviously, $\mathrm{graph}\Phi_j = Q_j$. From (7.11) and (7.12) we deduce that

$$\mathrm{graph}\Phi = \bigcup_{j \in J} \mathrm{graph}\Phi_j, \quad \Phi(x) = \bigcup_{j \in J} \Phi_j(x).$$

Claim 1. *For each $j \in J$ there exists a constant $\ell_j > 0$ such that*

$$\Phi_j(x) \subset \Phi_j(u) + \ell_j\|x - u\|\bar{B}_{R^m} \tag{7.13}$$

whenever $\Phi_j(x) \neq \emptyset$ and $\Phi_j(u) \neq \emptyset$. (This means that Φ_j is Lipschitz on its effective domain.)

For proving the claim, consider the linear operator $\tau : R^n \times R^m \to R^n$ defined by setting $\tau(x, y) = x$ for every $(x, y) \in R^n \times R^m$. Let

$$Q_j(x) = \{z \in Q_j \ : \ \tau(z) = x\}. \tag{7.14}$$

By Corollary 7.1, there exists $\ell_j > 0$ such that

$$Q_j(x) \subset Q_j(u) + \ell_j \|x - u\| \bar{B}_{R^{n+m}} \tag{7.15}$$

whenever $Q_j(x) \neq \emptyset$ and $Q_j(u) \neq \emptyset$. From (7.12) and (7.14) it follows that

$$Q_j(x) = \{x\} \times \Phi_j(x) \quad \forall x \in R^n. \tag{7.16}$$

In particular, $Q_j(x) \neq \emptyset$ if and only if $\Phi_j(x) \neq \emptyset$. Given any $x \in R^n$, $u \in R^n$ and $y \in \Phi_j(x)$, from (7.15) and (7.16) we see that there exist $v \in \Phi(u)$ such that

$$\|(x, y) - (u, v)\| \leq \ell_j \|x - u\|.$$

Since $\|(x, y) - (u, v)\| = (\|x - u\|^2 + \|y - v\|^2)^{1/2}$, the last inequality implies that $\|y - v\| \leq \ell_j \|x - u\|$. From what has already been proved, it may be concluded that (7.13) holds whenever $\Phi_j(x) \neq \emptyset$ and $\Phi_j(u) \neq \emptyset$.

We set $\ell = \max\{l_j \ : \ j \in J\}$. The proof will be completed if we can establish the following fact.

CLAIM 2. *For each $\bar{x} \in R^n$ there exists a neighborhood $U_{\bar{x}}$ of $\bar{x}$ such that* (7.10) *holds.*

Let $\bar{x} \in R^n$ be given arbitrarily. Define

$$J_0 = \{j \in J \ : \ \bar{x} \in \mathrm{dom}\Phi_j\}, \quad J_1 = J \setminus J_0.$$

Since $\mathrm{dom}\Phi_j = \tau(Q_j)$, where τ is the linear operator defined above, we see that $\mathrm{dom}\Phi_j$ is a polyhedral convex set. This implies that the set $\bigcup_{j \in J_1} \mathrm{dom}\Phi_j$ is closed. (Note that if $J_1 = \emptyset$ then this set is empty.) As $\bar{x} \notin \bigcup_{j \in J_1} \mathrm{dom}\Phi_j$, there must exist $\varepsilon > 0$ such that the neighborhood $U_{\bar{x}} := B(\bar{x}, \varepsilon)$ of $\bar{x}$ does not intersect the set

$$\bigcup_{j \in J_1} \mathrm{dom}\Phi_j.$$

Let $x \in U_{\bar{x}}$. If $x \notin \bigcup_{j\in J_0} \mathrm{dom}\Phi_j$, then

$$\Phi(x) = \left(\bigcup_{j\in J_0} \Phi_j(x)\right) \bigcup \left(\bigcup_{j\in J_1} \Phi_j(x)\right) = \emptyset.$$

So the inclusion (7.10) is valid. If $x \in \bigcup_{j\in J_0} \mathrm{dom}\Phi_j$, then we have

$$\Phi(x) = \bigcup_{j\in J_0} \Phi_j(x) = \bigcup_{j\in J_0'} \Phi_j(x),$$

where $J_0' = \{j \in J_0 : x \in \mathrm{dom}\Phi_j\}$. For each $j \in J_0'$, according to Claim 1, we have

$$\Phi_j(x) \subset \Phi_j(\bar{x}) + \ell_j\|x - \bar{x}\|\bar{B}_{R^m} \subset \Phi(\bar{x}) + \ell\|x - \bar{x}\|\bar{B}_{R^m}.$$

Therefore

$$\Phi(x) = \bigcup_{j\in J_0'} \Phi_j(x) \subset \Phi(\bar{x}) + \ell\|x - \bar{x}\|\bar{B}_{R^m}.$$

Claim 2 has been proved. □

Remark 7.3. From the proof of Theorem 7.2 it is easily seen that Φ is Lipschitz on the set $\bigcap_{j\in J} \mathrm{dom}\Phi_j$ with the Lipschitz constant ℓ.

Combining Theorem 7.2 with Proposition 7.2 we obtain the next result on upper-Lipschitz continuity of the solution map in a general AVI problem where the linear variables are subject to perturbation.

Theorem 7.3. *Suppose that $M \in R^{n\times n}$, $A \in R^{m\times n}$ and $C \in R^{s\times n}$ are given matrices. Then there exists a constant $\ell > 0$ such that the multifunction $\Phi : R^n \times R^m \times R^s \to 2^{R^n}$ defined by the formula*

$$\Phi(q, b, d) = \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta(b, d))),$$

where $(q, b, d) \in R^n \times R^m \times R^s$ and $\Delta(b, d) := \{x \in R^n : Ax \geq b,\ Cx = d\}$, is locally upper-Lipschitz at any point $(\bar{q}, \bar{b}, \bar{d}) \in R^n \times R^m \times R^s$ with the Lipschitz constant ℓ.

Applying Theorem 7.3 to the case where the constraint set $\Delta(b, d)$ of the problem $\mathrm{AVI}(M, q, \Delta(b, d))$ is fixed (i.e., the pair (b, d) is not subject to perturbations), we have the following result.

Corollary 7.4. *Suppose that $M \in R^{n\times n}$ is a given matrix and $\Delta \subset R^n$ is a nonempty polyhedral convex set. Then there exists a constant $\ell > 0$ such that the multifunction $\Phi : R^n \to 2^{R^n}$ defined by the formula*

$$\Phi(q) = \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)),$$

where $q \in R^n$, is locally upper-Lipschitz at any point $\bar{q} \in R^n$ with the Lipschitz constant ℓ.

7.3 Upper-Lipschitz Continuity with respect to all Variables

Our aim in this section is to study some results on locally upper-Lipschitz continuity of the multifunction $\Phi : R^{n\times n} \times R^n \to 2^{R^n}$ defined by the formula

$$\Phi(M, q) = \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)),$$

where $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ denotes the solution set of the problem (6.1). First we consider the case where Δ is a polyhedral convex cone. Then we consider the case where Δ is an arbitrary nonempty polyhedral convex set.

The following theorem specializes to Theorem 7.5.1 in Cottle et al. (1992) about the solution map in parametric linear complementarity problems if $\Delta = R^n_+$.

Theorem 7.4. *Suppose that $\Delta \subset R^n$ is a polyhedral convex cone. Suppose that $M \in R^{n\times n}$ is a given matrix and $q \in R^n$ is a given vector. If M is copositive on Δ and*

$$q \in \mathrm{int}\left([\mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta))]^+\right), \tag{7.17}$$

then there exist constants $\varepsilon > 0$, $\delta > 0$ and $\ell > 0$ such that if $(\widetilde{M}, \widetilde{q}) \in R^{n\times n} \times R^n$, $\widetilde{M}$ is copositive on Δ, and if

$$\max\{\|\widetilde{M} - M\|,\ \|\widetilde{q} - q\|\} < \varepsilon, \tag{7.18}$$

then the set $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$ is nonempty,

$$\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \subset \delta \bar{B}_{R^n}, \tag{7.19}$$

and

$$\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \subset \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) + \ell(\|\widetilde{M} - M\| + \|\widetilde{q} - q\|)\bar{B}_{R^n}. \tag{7.20}$$

Proof. Suppose that M is copositive on Δ and (7.17) is satisfied. Since Δ is a polyhedral convex cone, we see that for every $(\widetilde{M}, \widetilde{q}) \in R^{n\times n} \times R^n$ the problem $\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)$ is an GLCP. In particular, $\mathrm{AVI}(M, 0, \Delta)$ is a GLCP problem and we have

$$\mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta)) = \{v \in \Delta : Mv \in \Delta^+,\ \langle Mv, v\rangle = 0\}.$$

Since Sol(AVI($M, 0, \Delta$)) is a closed cone, Lemma 6.4 shows that (7.17) is equivalent to the following condition

$$q^T v > 0 \quad \forall v \in \text{Sol}(\text{AVI}(M, 0, \Delta)) \setminus \{0\}. \tag{7.21}$$

CLAIM 1. *There exists* $\varepsilon > 0$ *such that if* $(\widetilde{M}, \widetilde{q}) \in R^{n \times n} \times R^n$, $\widetilde{M}$ *is copositive on* Δ, *and if* (7.18) *holds, then the set* Sol(AVI($\widetilde{M}, \widetilde{q}, \Delta$)) *is nonempty.*

Suppose Claim 1 were false. Then we could find a sequence $\{(M^k, q^k)\}$ in $R^{n \times n} \times R^n$ such that M^k is copositive on Δ for every $k \in N$, $(M^k, q^k) \to (M, q)$ as $k \to \infty$, and Sol(AVI(M^k, q^k, Δ)) $= \emptyset$ for every $k \in N$. According to Theorem 6.5, we must have

$$q^k \notin \text{int}\left([\text{Sol}(\text{AVI}(M^k, 0, \Delta))]^+\right) \quad \forall k \in N.$$

Applying Lemma 6.4 we can assert that for each $k \in N$ there exists $v^k \in$ Sol(AVI($M^k, 0, \Delta$)) $\setminus \{0\}$ such that $(q^k)^T v^k \leq 0$. Then we have

$$v^k \in \Delta, \quad M^k v^k \in \Delta^+, \quad \langle M^k v^k, v^k \rangle = 0, \tag{7.22}$$

for every $k \in N$. Without loss of generality we can assume that

$$\frac{v^k}{\|v^k\|} \to \bar{v} \in R^n, \quad \|\bar{v}\| = 1.$$

From (7.22) it follows that

$$\frac{v^k}{\|v^k\|} \in \Delta, \quad M^k \frac{v^k}{\|v^k\|} \in \Delta^+, \quad \left\langle M^k \frac{v^k}{\|v^k\|}, \frac{v^k}{\|v^k\|} \right\rangle = 0.$$

Taking limits as $k \to \infty$ we obtain

$$\bar{v} \in \Delta, \quad M\bar{v} \in \Delta^+, \quad \langle M\bar{v}, \bar{v} \rangle = 0.$$

This shows that $\bar{v} \in$ Sol(AVI($M, 0, \Delta$)). Since $(q^k)^T v^k \leq 0$, we see that $(q^k)^T \dfrac{v^k}{\|v^k\|} \leq 0$ for every $k \in N$. Letting $k \to \infty$ yields $q^T \bar{v} \leq 0$. Since $\bar{v} \in$ Sol(AVI($M, 0, \Delta$)) $\setminus \{0\}$, the last inequality contradicts (7.21). We have thus justified Claim 1.

CLAIM 2. *There exist* $\varepsilon > 0$ *and* $\delta > 0$ *such that if* $(\widetilde{M}, \widetilde{q}) \in R^{n \times n} \times R^n$, $\widetilde{M}$ *is copositive on* Δ, *and if* (7.18) *holds, then inclusion* (7.19) *is satisfied.*

To obtain a contradiction, suppose that there exist a sequence $\{(M^k, q^k)\}$ in $R^{n\times n} \times R^n$ and a sequence $\{x^k\}$ in R^n such that M^k is copositive on Δ for every $k \in N$, $x^k \in \mathrm{Sol}(\mathrm{AVI}(M^k, q^k, \Delta))$ for every $k \in N$, $(M^k, q^k) \to (M, q)$ as $k \to \infty$, and $\|x^k\| \to +\infty$ as $k \to \infty$. Since $x^k \in \mathrm{Sol}(\mathrm{AVI}(M^k, q^k, \Delta))$, we see that

$$x^k \in \Delta, \quad M^k x^k + q^k \in \Delta^+, \quad \langle M^k x^k + q^k, x^k\rangle = 0, \tag{7.23}$$

for every $k \in N$. There is no loss of generality in assuming that

$$\frac{x^k}{\|x^k\|} \to \bar{v} \in R^n, \quad \|\bar{v}\| = 1.$$

From (7.23) it follows that

$$\bar{v} \in \Delta, \quad M\bar{v} \in \Delta^+, \quad \langle M\bar{v}, \bar{v}\rangle = 0.$$

From this we conclude that $\bar{v} \in \mathrm{Sol}(\mathrm{AVI}(M, 0, \Delta))$. Since

$$\langle M^k x^k + q^k, x^k\rangle = 0$$

and since $0^+\Delta = \Delta$ and M^k is copositive on Δ, we have

$$-(q^k)^T x^k = -\langle q^k, x^k\rangle = \langle M^k x^k, x^k\rangle \geq 0.$$

Then

$$q^T\bar{v} = \lim_{k\to\infty}\left((q^k)^T \frac{x^k}{\|x^k\|}\right) \leq 0.$$

This contradicts (7.21). Claim 2 has been proved.

Now we are in a position to show that there exist $\varepsilon > 0$, $\delta > 0$ and $\ell > 0$ such that if $(\widetilde{M}, \widetilde{q}) \in R^{n\times n} \times R^n$, $\widetilde{M}$ is copositive on Δ, and if (7.18) holds, then $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \neq \emptyset$ and (7.19), (7.20) are satisfied.

Combining Claim 1 with Claim 2 we see that there exist $\varepsilon > 0$ and $\delta > 0$ such that if $(\widetilde{M}, \widetilde{q}) \in R^{n\times n} \times R^n$, $\widetilde{M}$ is copositive on Δ, and if (7.18) holds, then $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \neq \emptyset$ and (7.19) is satisfied. According to Corollary 7.4, for the given matrix M and vector q, there exist a constant $\ell_M > 0$ and a neighborhood U_q of q such that

$$\mathrm{Sol}(\mathrm{AVI}(M, q', \Delta)) \subset \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) + \ell_M\|q' - q\|\bar{B}_{R^n} \tag{7.24}$$

for every $q' \in U_q$. Let $(\widetilde{M}, \widetilde{q}) \in R^{n \times n} \times R^n$ be such that $\widetilde{M}$ is copositive on Δ and (7.18) holds. Select any $\widetilde{x} \in \mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$. Setting

$$\bar{q} = \widetilde{q} + (\widetilde{M} - M)\widetilde{x} \tag{7.25}$$

we will show that

$$\widetilde{x} \in \mathrm{Sol}(\mathrm{AVI}(M, \bar{q}, \Delta)). \tag{7.26}$$

Since

$$\langle \widetilde{M}\widetilde{x} + \widetilde{q}, x - \widetilde{x} \rangle \geq 0 \quad \forall x \in \Delta,$$

using (7.25) we deduce that

$$\begin{aligned} 0 \leq \langle \widetilde{M}\widetilde{x} + \widetilde{q}, x - \widetilde{x} \rangle &= \langle \widetilde{M}\widetilde{x} + \bar{q} - \widetilde{M}\widetilde{x} + M\widetilde{x}, x - \widetilde{x} \rangle \\ &= \langle M\widetilde{x} + \bar{q}, x - \widetilde{x} \rangle \end{aligned}$$

for every $x \in \Delta$. This shows that (7.26) is valid. From (7.18), (7.19) and (7.25) it follows that

$$\|\bar{q} - q\| \leq \|\widetilde{q} - q\| + \|\widetilde{M} - M\| \|\widetilde{x}\| \leq \varepsilon(1 + \delta).$$

Consequently, choosing a smaller $\varepsilon > 0$ if necessary, we can assert that $\bar{q} \in U_q$ *whenever* $(\widetilde{M}, \widetilde{q}) \in R^{n \times n} \times R^n$, $\widetilde{M}$ *is copositive on* Δ, (7.18) *holds*. Hence from (7.24) and (7.26) we deduce that there exists $x \in \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ such that

$$\begin{aligned} \|\widetilde{x} - x\| &\leq \ell_M \|\bar{q} - q\| \\ &\leq \ell_M (\|\widetilde{q} - q\| + \|\widetilde{M} - M\| \|\widetilde{x}\|) \\ &\leq \ell_M (\|\widetilde{q} - q\| + \delta \|\widetilde{M} - M\|) \\ &\leq \ell (\|\widetilde{q} - q\| + \|\widetilde{M} - M\|), \end{aligned}$$

where $\ell = \max\{\ell_M, \delta\ell_M\}$. We have thus obtained (7.20). The proof is complete. □

Our next goal is to establish the following interesting result on AVI problems with positive semidefinite matrices.

Theorem 7.5. (See Robinson (1979), Theorem 2) *Let* $M \in R^{n \times n}$ *be a positive semidefinite matrix,* Δ *a nonempty polyhedral convex set in* R^n, *and* $q \in R^n$. *Then the following two properties are equivalent:*

(i) *The solution set* $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ *is nonempty and bounded;*

(ii) *There exists* $\varepsilon > 0$ *such that for each* $\widetilde{M} \in R^{n\times n}$ *and each* $\widetilde{q} \in R^n$ *with*

$$\max\{\|\widetilde{M} - M\|, \ \|\widetilde{q} - q\|\} < \varepsilon, \tag{7.27}$$

the set $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$ *is nonempty.*

For proving the above theorem we shall need the following three auxiliary lemmas in which *it is assumed that* $M \in R^{n\times n}$ *is a positive semidefinite matrix,* $\Delta \subset R^n$ *is a nonempty polyhedral convex set, and* $q \in R^n$. We set $M\Delta = \{Mx : x \in \Delta\}$.

Lemma 7.1. (See, for instance, Best and Chakravarti (1992)) *For any* $\bar{v} \in R^n$, *if* $\bar{v}^T M \bar{v} = 0$ *then* $(M + M^T)\bar{v} = 0$.

Proof. Consider the unconstrained quadratic program

$$\min\left\{ f(x) := \frac{1}{2} x^T (M + M^T) x \, : \, x \in R^n \right\}.$$

From our assumptions it follows that

$$\begin{aligned} \frac{1}{2} x^T (M + M^T) x = x^T M x \ &\geq 0 \\ &= \bar{v}^T M \bar{v} \\ &= \frac{1}{2} \bar{v}^T (M + M^T) \bar{v} \end{aligned}$$

for every $x \in R^n$. Hence $\bar{v}$ is a global solution of the above problem. By Theorem 3.1 we have

$$0 = \nabla f(\bar{v}) = (M + M^T)\bar{v},$$

which completes the proof. □

Lemma 7.2. *The inclusion*

$$q \in \mathrm{int}((0^+\Delta)^+ - M\Delta) \tag{7.28}$$

holds if and only if

$$\forall v \in 0^+\Delta \setminus \{0\} \ \exists x \in \Delta \ \text{ such that } \ \langle Mx + q, v\rangle > 0. \tag{7.29}$$

Proof. *Necessity:* Suppose that (7.28) holds. Then there exists $\varepsilon > 0$ such that

$$B(q, \varepsilon) \subset (0^+\Delta)^+ - M\Delta. \tag{7.30}$$

To obtain a contradiction, suppose that there exists $\bar{v} \in 0^+\Delta \setminus \{0\}$ such that

$$\langle Mx + q, \bar{v} \rangle \leq 0 \quad \forall x \in \Delta.$$

By (7.30), for every $q' \in B(q, \varepsilon)$ there exist $w \in (0^+\Delta)^+$ and $x \in \Delta$ such that $q' = w - Mx$. So we have

$$\langle q' - q, \bar{v} \rangle \geq \langle w, \bar{v} \rangle \geq 0 \quad \forall q' \in B(q, \varepsilon).$$

This clearly forces $\bar{v} = 0$, which is impossible.

Sufficiency: On the contrary, suppose that (7.29) is valid, but (7.28) is false. Then there exists a sequence $\{q^k\} \subset R^n$ such that $q^k \notin (0^+\Delta)^+ - M\Delta$ for all $k \in N$, and $q^k \to q$. From this we deduce that

$$(M\Delta + q^k) \cap (0^+\Delta)^+ = \emptyset \quad \forall k \in N.$$

Since $M\Delta + q^k$ and $(0^+\Delta)^+$ are two disjoint polyhedral convex sets, by Theorem 11.3 from Rockafellar (1970) there exists a hyperplane separating these sets properly. Since $(0^+\Delta)^+$ is a cone, by Theorem 11.7 from Rockafellar (1970) there exists a hyperplane which separates the above two sets properly and passes through the origin. So there exists $v^k \in R^n$ with $\|v^k\| = 1$ such that

$$\langle v^k, Mx + q^k \rangle \leq 0 \leq \langle v^k, w \rangle \quad \forall x \in \Delta,\ \forall w \in (0^+\Delta)^+. \tag{7.31}$$

(Actually, the above-mentioned hyperplane is defined by the formula $H = \{z \in R^n : \langle v^k, z \rangle = 0\}$). Without loss of generality we can assume that $v^k \to \bar{v} \in R^n$, $\|\bar{v}\| = 1$. From (7.31) it follows that

$$\langle \bar{v}, Mx + q \rangle \leq 0 \quad \forall x \in \Delta \tag{7.32}$$

and

$$\langle \bar{v}, w \rangle \geq 0 \quad \forall w \in (0^+\Delta)^+. \tag{7.33}$$

By Theorem 14.1 from Rockafellar (1970), from (7.33) it follows that $\bar{v} \in 0^+\Delta$. Combining this with (7.32) we see that (7.29) is false, which is impossible. □

Lemma 7.3. (See Gowda-Pang (1994a), Theorem 7) *The solution set* Sol(AVI(M, q, Δ)) *is nonempty and bounded if and only if* (7.28) *holds.*

Proof. *Necessity:* To obtain a contradiction, suppose that the set Sol(AVI(M, q, Δ)) is nonempty and bounded, but (7.28) does not hold. Then, by Lemma 7.2 there exists $\bar{v} \in 0^+\Delta \setminus \{0\}$ such that

(7.32) holds. Select a point $x^0 \in \mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$. For each $t > 0$, we set $x_t = x^0 + t\bar{v}$. Since $\bar{v} \in 0^+\Delta$, we have $x_t \in \Delta$ for every $t > 0$. Substituting x_t for x in (7.32) we get

$$\langle \bar{v}, Mx^0 + q\rangle + t\langle \bar{v}, M\bar{v}\rangle \le 0 \quad \forall t > 0.$$

This implies that $\langle \bar{v}, M\bar{v}\rangle \le 0$. Besides, since M is positive semidefinite, we have $\langle \bar{v}, M\bar{v}\rangle \ge 0$. So

$$\langle \bar{v}, M\bar{v}\rangle = 0. \tag{7.34}$$

By Lemma 7.1, from (7.34) we obtain

$$(M + M^T)\bar{v} = 0. \tag{7.35}$$

Fix any $x \in \Delta$. On account of (7.32), (7.34), (7.35) and the fact that $x^0 \in \mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$, we have

$$\begin{aligned}
\langle Mx_t + q, x - x_t\rangle &= \langle Mx^0 + q + tM\bar{v}, x - x^0 - t\bar{v}\rangle \\
&= \langle Mx^0 + q, x - x^0\rangle + t\langle M\bar{v}, x - x^0\rangle \\
&\quad -t\langle Mx^0 + q, \bar{v}\rangle - t^2 \underbrace{\langle M\bar{v}, \bar{v}\rangle}_{=0} \\
&= \langle Mx^0 + q, x - x^0\rangle - t\underbrace{\langle \bar{v}, Mx + q\rangle}_{\le 0} \\
&\quad -t\langle \underbrace{(M + M^T)\bar{v}}_{=0}, x^0\rangle \\
&\ge \langle Mx^0 + q, x - x^0\rangle \\
&\ge 0.
\end{aligned}$$

Since this holds for every $x \in \Delta$, $x_t \in \mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$. As the last inclusion is valid for each $t > 0$, we conclude that $\mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$ is unbounded, a contradiction.

Sufficiency: Suppose that (7.28) holds. We have to show that the set $\mathrm{Sol}(\mathrm{AVI}(M,q,\Delta))$ is nonempty and bounded. By (7.28),

$$q \in (0^+\Delta)^+ - M\Delta.$$

Hence there exist $w \in (0^+\Delta)^+$ and $\bar{x} \in \Delta$ such that $q = w - M\bar{x}$. Since $M\bar{x} + q = w \in (0^+\Delta)^+$, for every $v \in 0^+\Delta$ it holds

$$\langle M\bar{x} + q, v\rangle = \langle w, v\rangle \ge 0.$$

Since M is a positive semidefinite matrix, we see that both conditions (i) and (ii) in Theorem 6.1 are satisfied. Hence the set

Sol(AVI(M, q, Δ)) is nonempty. To show that Sol(AVI(M, q, Δ)) is bounded we suppose, contrary to our claim, that there exists a sequence $\{x^k\}$ in Sol(AVI(M, q, Δ)) such that $\|x^k\| \to +\infty$. There is no loss of generality in assuming that $x^k \neq 0$ for each $k \in N$, and

$$\frac{x^k}{\|x^k\|} \to \bar{v} \in R^n, \quad \|\bar{v}\| = 1.$$

Let $m \in N$, $A \in R^{m \times n}$ and $b \in R^m$ be such that $\Delta = \{x \in R^n : Ax \geq b\}$. Since $Ax^k \geq b$ for every $k \in N$, dividing the inequality by $\|x^k\|$ and letting $k \to \infty$ we obtain $A\bar{v} \geq 0$. This shows that $\bar{v} \in 0^+\Delta$. We have

$$\langle Mx^k + q, x - x^k \rangle \geq 0 \quad \forall x \in \Delta \; \forall k \in N.$$

Hence

$$\langle Mx^k + q, x \rangle \geq \langle Mx^k, x^k \rangle + \langle q, x^k \rangle \quad \forall x \in \Delta \; \forall k \in N. \qquad (7.36)$$

Dividing the last inequality by $\|x^k\|^2$ and letting $k \to \infty$ we get $0 \geq \langle M\bar{v}, \bar{v} \rangle$. Since M is positive semidefinite, from this we see that $\langle M\bar{v}, \bar{v} \rangle = 0$. Thus, by Lemma 7.1 we have

$$M\bar{v} = -M^T\bar{v}. \qquad (7.37)$$

Fix a point $x \in \Delta$. Since $\langle Mx^k, x^k \rangle \geq 0$ for every $k \in N$, (7.36) implies that

$$\langle Mx^k + q, x \rangle \geq \langle q, x^k \rangle \quad \forall k \in N.$$

Dividing the last inequality by $\|x^k\|$ and letting $k \to \infty$ we obtain

$$\langle M\bar{v}, x \rangle \geq \langle q, \bar{v} \rangle.$$

Combining this with (7.37) we can assert that

$$\langle Mx + q, \bar{v} \rangle \leq 0 \quad \forall x \in \Delta.$$

Since $\bar{v} \in (0^+\Delta) \setminus \{0\}$, from the last fact and Lemma 7.2 it follows that (7.28) does not hold. We have thus arrived at a contradiction. The proof is complete. □

Proof of Theorem 7.5.

We first prove the implication (i) ⇒ (ii). To obtain a contradiction, suppose that Sol(AVI(M, q, Δ)) is nonempty and bounded, while there exists a sequence $(M^k, q^k) \in R^{n\times n} \times R^n$ such that $(M^k, q^k) \to (M, q)$ and

$$\mathrm{Sol}(\mathrm{AVI}(M^k, q^k, \Delta)) = \emptyset \quad \forall k \in N. \tag{7.38}$$

Since Δ is nonempty, for $j \in N$ large enough, the set

$$\Delta_j := \Delta \cap \{x \in R^n : \|x\| \le j\}$$

is nonempty. Without restriction of generality we can assume that $\Delta_j \neq \emptyset$ for every $j \in N$. By the Hartman-Stampacchia Theorem (Theorem 5.1) we can find a point, denoted by $x^{k,j}$, in the solution set $\mathrm{Sol}(\mathrm{AVI}(M^k, q^k, \Delta_j))$. We have

$$\langle M^k x^{k,j} + q^k, x - x^{k,j}\rangle \ge 0 \quad \forall x \in \Delta_j. \tag{7.39}$$

Note that

$$\|x^{k,j}\| = j \quad \forall j \in N. \tag{7.40}$$

Indeed, if $\|x^{k,j}\| < j$ then there exists $\mu > 0$ such that $\bar{B}(x^{k,j}, \mu) \subset \bar{B}(0, j)$. Hence from (7.39) it follows that

$$\langle M^k x^{k,j} + q^k, x - x^{k,j}\rangle \ge 0 \quad \forall x \in \Delta \cap \bar{B}(x^{k,j}, \mu).$$

By Proposition 5.3, this implies that $x^{k,j} \in \mathrm{Sol}(\mathrm{AVI}(M^k, q^k, \Delta))$, which is impossible because (7.38) holds. Fixing an index $j \in N$ we consider the sequence $\{x^{k,j}\}_{k\in N}$. From (7.40) we deduce that this sequence has a convergent subsequence. There is no loss of generality in assuming that

$$\lim_{k\to\infty} x^{k,j} = x^j, \quad x^j \in R^n, \quad \|x^j\| = j. \tag{7.41}$$

Letting $k \to \infty$ we deduce from (7.39) that

$$\langle Mx^j + q, x - x^j\rangle \ge 0 \quad \forall x \in \Delta_j. \tag{7.42}$$

On account of (7.41), without loss of generality we can assume that

$$\frac{x^j}{\|x^j\|} \to \bar{v} \in R^n, \quad \|\bar{v}\| = 1.$$

Let us fix a point $x \in \Delta$. It is clear that there exists an index $j_x \in N$ such that $x \in \Delta_j$ for every $j \ge j_x$. From (7.42) we deduce that

$$\langle Mx^j + q, x - x^j\rangle \ge 0 \quad \forall j \ge j_x.$$

Hence

$$\langle Mx^j + q, x \rangle \geq \langle Mx^j, x^j \rangle + \langle q, x^j \rangle \quad \forall j \geq j_x. \tag{7.43}$$

As in the last part of the proof of Lemma 7.3, we can show that $\bar{v} \in (0^+\Delta) \setminus \{0\}$ and deduce from (7.43) the following inequality

$$\langle Mx + q, \bar{v} \rangle \leq 0.$$

Since the latter holds for every $x \in \Delta$, applying Lemma 7.2 we see that the inclusion (7.28) cannot hold. According to Lemma 7.3, the last fact implies that the set Sol(AVI(M, q, Δ)) cannot be nonempty and bounded. This contradicts our assumption.

We now prove the implication (ii) $\Rightarrow$ (i). Suppose that there exists $\varepsilon > 0$ such that if matrix $\widetilde{M} \in R^{n\times n}$ and vector $\widetilde{q} \in R^n$ satisfy condition (7.27) then the set Sol(AVI($\widetilde{M}, \widetilde{q}, \Delta$)) is nonempty. Consequently, for any $\widetilde{q} \in R^n$ satisfying $\|\widetilde{q} - q\| < \varepsilon$, the set Sol(AVI($M, \widetilde{q}, \Delta$)) is nonempty. Let $\widetilde{x} \in$ Sol(AVI($M, \widetilde{q}, \Delta$)). For any $v \in 0^+\Delta$ we have

$$(M\widetilde{x} + \widetilde{q})^T v = \langle M\widetilde{x} + \widetilde{q}, (\widetilde{x} + v) - \widetilde{x} \rangle \geq 0.$$

Hence $M\bar{x} + \widetilde{q} \in (0^+\Delta)^+$. So we have $\widetilde{q} \in (0^+\Delta)^+ - M\Delta$. Since this inclusion is valid for each $\widetilde{q}$ satisfying $\|\widetilde{q} - q\| < \varepsilon$, we conclude that

$$q \in \operatorname{int}((0^+\Delta)^+ - M\Delta).$$

By Lemma 7.3, the set Sol(AVI(M, q, Δ)) is nonempty and bounded. The proof is complete. □

Let us consider three illustrative examples.

Example 7.1. Setting $\Delta = [0, +\infty) \subset R^1$, $M = (-1)$, and $q = 0$, we have Sol(AVI(M, q, Δ)) $= \{0\}$. Note that matrix M is not positive semidefinite. Taking $\widetilde{M} = M$ and $\widetilde{q} = -\theta$, where $\theta > 0$, we check at once that Sol(AVI($\widetilde{M}, \widetilde{q}, \Delta$)) $= \emptyset$. So, for this AVI problem, property (i) in Theorem 7.5 holds, but property (ii) does not hold. This example shows that, in Theorem 7.5, one cannot omit the assumption that M is a positive semidefinite matrix.

Example 7.2. Setting $\Delta = (-\infty, +\infty) = R^1$, $M = (0)$, and $q = 0$, we have Sol(AVI(M, q, Δ)) $= \Delta$. So property (i) in Theorem 7.5 does not hold for this example. Taking $\widetilde{M} = (0)$ and $\widetilde{q} = \theta$, where $\theta > 0$, we have Sol(AVI($\widetilde{M}, \widetilde{q}, \Delta$)) $= \emptyset$. This shows that, for the

AVI problem under consideration, property (ii) in Theorem 7.5 fails to hold.

Example 7.3. Setting $\Delta = [1, +\infty) \subset R^1$, $M = (0)$, and $q = 0$, we have $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) = \Delta$. Taking $\widetilde{M} = M$ and $\widetilde{q} = \theta$, where $\theta > 0$, we see that $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) = \{1\}$. But taking $\widetilde{M} = (-\theta)$ and $\widetilde{q} = 0$, where $\theta > 0$, we see that $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) = \emptyset$. So, for this problem, both the properties (i) and (ii) in Theorem 7.5 do not hold.

In connection with Theorem 7.5, it is natural to raise the following open question.

QUESTION: Is it true that property (i) in Theorem 7.5 implies that there exists $\varepsilon > 0$ such that if matrix $\widetilde{M} \in R^{n \times n}$ and vector $\widetilde{q} \in R^n$ satisfy condition (7.27) then the set $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$ is bounded (may be empty)?

The next example shows that property (i) in Theorem 7.5 does not imply that the solution sets $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$, where $(\widetilde{M}, \widetilde{q})$ is taken from a neighborhood of (M, q), are uniformly bounded.

Example 7.4. (See Robinson (1979), pp. 139–140) Let $\Delta = [0, +\infty) \subset R^1$, $M = (0)$, and $q = 1$. It is clear that

$$\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) = \{0\}.$$

Taking $\widetilde{M} = (-\mu)$ and $\widetilde{q} = 1$, where $\mu > 0$, we have

$$\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) = \left\{0, \frac{1}{\mu}\right\}.$$

From this we conclude that there exist no $\varepsilon > 0$ and $\delta > 0$ such that if matrix $\widetilde{M} \in R^{1 \times 1}$ and vector $\widetilde{q} \in R^1$ satisfy condition (7.27) then $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \subset \delta \bar{B}_{R^1}$.

The following theorem is one of the main results on solution stability of AVI problems. One can observe that this theorem and Theorem 7.4 are independent results.

Theorem 7.6. (See Robinson (1979), Theorem 2) *Suppose that $\Delta \subset R^n$ is a nonempty polyhedral convex set. Suppose that $M \in R^{n \times n}$ is a given matrix and $q \in R^n$ is a given vector. If M is a positive semidefinite matrix and if the solution set* $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ *is nonempty and bounded, then there exist constants $\varepsilon > 0$, $\delta > 0$ and $\ell > 0$ such that if $(\widetilde{M}, \widetilde{q}) \in R^{n \times n} \times R^n$, $\widetilde{M}$ is positive semidefinite, and if*

$$\max\{\|\widetilde{M} - M\|,\ \|\widetilde{q} - q\|\} < \varepsilon, \tag{7.44}$$

then the set $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$ *is nonempty,*

$$\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \subset \delta \bar{B}_{R^n}, \tag{7.45}$$

and

$$\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta)) \subset \mathrm{Sol}(\mathrm{AVI}(M, q, \Delta)) + \ell(\|\widetilde{M} - M\| + \|\widetilde{q} - q\|)\bar{B}_{R^n}. \tag{7.46}$$

Proof. Since M is positive semidefinite and $\mathrm{Sol}(\mathrm{AVI}(M, q, \Delta))$ is nonempty and bounded, by Lemmas 7.2 and 7.3 we have

$$\forall v \in 0^+\Delta \setminus \{0\} \ \exists x \in \Delta \ \text{ such that } \ \langle Mx + q, v \rangle > 0. \tag{7.47}$$

Moreover, according to Theorem 7.5, there exists $\varepsilon_0 > 0$ such that for each matrix $\widetilde{M} \in R^{n \times n}$ and each $\widetilde{q} \in R^n$ satisfying

$$\max\{\|\widetilde{M} - M\|, \ \|\widetilde{q} - q\|\} < \varepsilon_0,$$

the set $\mathrm{Sol}(\mathrm{AVI}(\widetilde{M}, \widetilde{q}, \Delta))$ is nonempty. We claim that there exist constants $\varepsilon > 0$ and $\delta > 0$ such that (7.45) holds for every $(\widetilde{M}, \widetilde{q}) \in R^{n \times n} \times R^n$ satisfying condition (7.44) and the requirement that $\widetilde{M}$ is a positive semidefinite matrix. Indeed, if the claim were false we would find a sequence $\{(M^k, q^k)\}$ in $R^{n \times n} \times R^n$ and a sequence $\{x^k\}$ in R^n such that M^k is positive semidefinite for every $k \in$, $(M^k, q^k) \to (M, q)$, $x^k \in \mathrm{Sol}(\mathrm{AVI}(M^k, q^k, \Delta))$ for every $k \in N$, and $\|x^k\| \to +\infty$ as $k \to \infty$. For each $x \in \Delta$, we have

$$\langle M^k x^k + q^k, x - x^k \rangle \geq 0 \quad \forall k \in N. \tag{7.48}$$

Without loss of generality we can assume that $x^k \neq 0$ for every $k \in N$, and

$$\frac{x^k}{\|x^k\|} \to \bar{v} \in R^n, \quad \|\bar{v}\| = 1.$$

It is easily seen that $\bar{v} \in (0^+\Delta)^+$. From (7.48) it follows that

$$\langle M^k x^k + q^k, x \rangle \geq \langle M^k x^k, x^k \rangle + \langle q^k, x^k \rangle \quad \forall k \in N. \tag{7.49}$$

Dividing the last inequality by $\|x^k\|^2$ and letting $k \to \infty$ we get $0 \geq \langle M\bar{v}, \bar{v} \rangle$. Since M is positive semidefinite, from this we see that $\langle M\bar{v}, \bar{v} \rangle = 0$. By Lemma 7.1 we have

$$M\bar{v} = -M^T\bar{v}. \tag{7.50}$$

Fix a point $x \in \Delta$. Since M^k is positive semidefinite, we have $\langle M^k x^k, x^k\rangle \geq 0$ for every $k \in N$. Hence (7.49) implies that

$$\langle M^k x^k + q^k, x\rangle \geq \langle q^k, x^k\rangle \quad \forall k \in N.$$

Dividing the last inequality by $\|x^k\|$ and letting $k \to \infty$ we obtain

$$\langle M\bar{v}, x\rangle \geq \langle q, \bar{v}\rangle.$$

Combining this with (7.50) we get

$$\langle Mx + q, \bar{v}\rangle \leq 0 \quad \forall x \in \Delta.$$

Since $\bar{v} \in (0^+\Delta)^+ \setminus \{0\}$, the last fact contradicts (7.47). Our claim has been proved. We can now proceed analogously to the proof of Claim 3 in the proof of Theorem 7.4 to find the required constants $\varepsilon > 0$, $\delta > 0$ and $\ell > 0$. □

7.4 Commentaries

As it has been noted in Robinson (1981), p. 206, the class of polyhedral multifunctions is closed under finite addition, scalar multiplication, and finite composition. This means that if $\Phi : R^n \to 2^{R^m}$, $\Psi : R^m \to 2^{R^s}$, $\Phi_j : R^n \to 2^{R^m}$ $(j = 1, \dots, m)$ are some given polyhedral multifuntions and $\lambda \in R$ is a given scalar, then the formulae

$$(\lambda\Phi)(x) = \lambda\Phi(x) \quad (\forall x \in R^n),$$

$$(\Phi_1 + \dots + \Phi_k)(x) = \Phi_1(x) + \dots + \Phi_k(x) \quad (\forall x \in R^n),$$

and

$$(\Psi \circ \Phi)(x) = \Psi(\Phi(x)) \quad (\forall x \in R^n),$$

create new polyhedral multifunctions which are denoted by $\lambda\Phi$, $\Phi_1 + \dots + \Phi_k$ and $\Psi \circ \Phi$, respectively.

The proof of Theorem 7.4 is similar in spirit to the proof of Theorem 7.5.1 in Cottle et al. (1992).

The 'elementary' proof of the results of Robinson (see Theorems 7.5 and 7.6) on the solution stability of AVI problems with positive semidefinite matrices given in this chapter is new. We hope that it can expose furthermore the beauty of these results. The original proof of Robinson is based on a general solution stability theorem for variational inequalities in Banach spaces (see Robinson (1979), Theorem 1).

Results presented in this chapter deal only with upper-Lipschitz continuity properties of the solution map of parametric AVI problems. For multifunctions, the lower semicontinuity, the upper semicontinuity, the openness, the Aubin property, the metric regularity, and the single-valuedness are other interesting properties which have many applications (see Aubin and Frankowska (1990), Mordukhovich (1993), Rockafellar and Wets (1998), and references therein). It is of interest to characterize these properties of the solution map in parametric AVI problems (in particular, of the solution map in parametric LCP problems). Some results in this direction have been obtained (see, for instance, Jansen and Tijs (1987), Gowda (1992), Donchev and Rockafellar (1996), Oettli and Yen (1995), Gowda and Sznajder (1996)). We will study the lower semicontinuity and the upper semicontinuity the solution map of parametric AVI problems in Chapter 18.