QUADRATIC PROGRAMMING
AND AFFINE VARIATIONAL
INEQUALITIES

A Qualitative Study
Nonconvex Optimization and Its Applications

VOLUME 78

Managing Editor:

Panos Pardalos
University of Florida, U.S.A.

Advisory Board:

J. R. Birge
University of Michigan, U.S.A.

Ding-Zhu Du
University of Minnesota, U.S.A.

C. A. Floudas
Princeton University, U.S.A.

J. Mockus
Lithuanian Academy of Sciences, Lithuania

H. D. Sherali
Virginia Polytechnic Institute and State University, U.S.A.

G. Stavroulakis
Technical University Braunschweig, Germany

H. Tuy
National Centre for Natural Science and Technology, Vietnam
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
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>ix</td>
</tr>
<tr>
<td>Notations and Abbreviations</td>
<td>xi</td>
</tr>
<tr>
<td>1 Quadratic Programming Problems</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Mathematical Programming Problems</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Convex Programs and Nonconvex Programs</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Smooth Programs and Nonsmooth Programs</td>
<td>14</td>
</tr>
<tr>
<td>1.4 Linear Programs and Nonlinear Programs</td>
<td>19</td>
</tr>
<tr>
<td>1.5 Quadratic Programs</td>
<td>21</td>
</tr>
<tr>
<td>1.6 Commentaries</td>
<td>27</td>
</tr>
<tr>
<td>2 Existence Theorems for Quadratic Programs</td>
<td>29</td>
</tr>
<tr>
<td>2.1 The Frank-Wolfe Theorem</td>
<td>29</td>
</tr>
<tr>
<td>2.2 The Eaves Theorem</td>
<td>36</td>
</tr>
<tr>
<td>2.3 Commentaries</td>
<td>43</td>
</tr>
<tr>
<td>3 Necessary and Sufficient Optimality Conditions for Quadratic Programs</td>
<td>45</td>
</tr>
<tr>
<td>3.1 First-Order Optimality Conditions</td>
<td>45</td>
</tr>
<tr>
<td>3.2 Second-Order Optimality Conditions</td>
<td>50</td>
</tr>
<tr>
<td>3.3 Commentaries</td>
<td>62</td>
</tr>
<tr>
<td>4 Properties of the Solution Sets of Quadratic Programs</td>
<td>65</td>
</tr>
<tr>
<td>4.1 Characterizations of the Unboundedness of the Solution Sets</td>
<td>65</td>
</tr>
<tr>
<td>4.2 Closedness of the Solution Sets</td>
<td>76</td>
</tr>
<tr>
<td>4.3 A Property of the Bounded Infinite Solution Sets</td>
<td>77</td>
</tr>
</tbody>
</table>
10.3 A Special Case ........................................ 168
10.4 Sufficient Conditions for the usc Property of $S(\cdot)$ ...... 174
10.5 Corollaries and Examples ................................ 179
10.6 USC Property of $S(\cdot)$: The General Case ............. 182
10.7 Commentaries ............................................ 193

11 Lower Semicontinuity of the KKT Point Set Mapping ........ 195
11.1 The Case of Canonical QP Problems ..................... 195
11.2 The Case of Standard QP Problems ....................... 199
11.3 Commentaries ............................................ 211

12 Continuity of the Solution Map in Quadratic Programming .... 213
12.1 USC Property of the Solution Map ....................... 213
12.2 LSC Property the Solution Map ......................... 216
12.3 Commentaries ............................................. 222

13 Continuity of the Optimal Value Function in Quadratic Programming ............................................. 223
13.1 Continuity of the Optimal Value Function ................. 223
13.2 Semicontinuity of the Optimal Value Function ............ 233
13.3 Commentaries ............................................. 237

14 Directional Differentiability of the Optimal Value Function .... 239
14.1 Lemmas .................................................. 239
14.2 Condition (G) ........................................... 246
14.3 Directional Differentiability of $\varphi(\cdot)$ ............... 250
14.4 Commentaries ........................................... 257

15 Quadratic Programming under Linear Perturbations:
I. Continuity of the Solution Maps .......................... 259
15.1 Lower Semicontinuity of the Local Solution Map ........ 260
15.2 Lower Semicontinuity of the Solution Map ............... 261
15.3 Commentaries ............................................. 268

16 Quadratic Programming under Linear Perturbations:
II. Properties of the Optimal Value Function ................ 269
16.1 Auxiliary Results ....................................... 269
16.2 Directional Differentiability .............................. 274
16.3 Piecewise Linear-Quadratic Property ..................... 278
<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.4 Proof of Proposition 16.2</td>
<td>287</td>
</tr>
<tr>
<td>16.5 Commentaries</td>
<td>289</td>
</tr>
<tr>
<td><strong>17 Quadratic Programming under Linear Perturbations:</strong></td>
<td></td>
</tr>
<tr>
<td>III. The Convex Case</td>
<td>291</td>
</tr>
<tr>
<td>17.1 Preliminaries</td>
<td></td>
</tr>
<tr>
<td>17.2 Projection onto a Moving Polyhedral Convex Set</td>
<td>293</td>
</tr>
<tr>
<td>17.3 Application to Variational Inequalities</td>
<td>297</td>
</tr>
<tr>
<td>17.4 Application to a Network Equilibrium Problem</td>
<td>300</td>
</tr>
<tr>
<td>17.5 Commentaries</td>
<td>305</td>
</tr>
<tr>
<td><strong>18 Continuity of the Solution Map in Affine Variational Inequalities</strong></td>
<td></td>
</tr>
<tr>
<td>18.1 USC Property of the Solution Map</td>
<td>307</td>
</tr>
<tr>
<td>18.2 LSC Property of the Solution Map</td>
<td>321</td>
</tr>
<tr>
<td>18.3 Commentaries</td>
<td>327</td>
</tr>
</tbody>
</table>

References: 329

Index: 343
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
Gue Myung Lee, Nguyen Nang Tam, and Nguyen Dong Yen
Notations and Abbreviations

$N$ the set of the positive integers
$\mathbb{R}$ the real line
$\overline{\mathbb{R}}$ the extended real line
$\mathbb{R}^n$ the $n$—dimensional Euclidean space
$\mathbb{R}_{+}^{n}$ the nonnegative orthant in $\mathbb{R}^n$
$\emptyset$ 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 $A$
$\text{rank} A$ the rank of matrix $A$
$\|A\|$ the norm of matrix $A$
$\mathbb{R}^{m \times n}$ the set of the $m \times n$-matrices
$\mathbb{R}_{S}^{n \times n}$ the set of the symmetric $n \times n$-matrices
$\text{det} A$ the determinant of a square matrix $A$
$E$ the unit matrix in $\mathbb{R}^{n \times n}$
$B(x, \delta)$ the open ball centered at $x$ with radius $\delta$
$\overline{B}(x, \delta)$ the closed ball centered at $x$

with radius $\delta$
$\mathbb{B}_{\mathbb{R}^n}$ the closed unit ball in $\mathbb{R}^n$
$\text{int} \Omega$ the interior of $\Omega$
$\overline{\Omega}$ the closure of $\Omega$
$\text{bd} \Omega$ the boundary of $\Omega$
$\text{co} \Omega$ the convex hull of $\Omega$
$\text{dist}(x, \Omega)$ the distance from $x$ to $\Omega$
$\text{cone} M$ the cone generated by $M$
$\text{ri} \Delta$ the relative interior of a convex set $\Delta$
$\text{aff} \Delta$ the affine hull of $\Delta$
$\text{extr} \Delta$ the set of the extreme points of $\Delta$
$0^+ \Delta$ the recession cone of $\Delta$
$\mathcal{T}_\Delta(\bar{x})$ the tangent cone to $\Delta$ at $\bar{x}$
$\mathcal{N}_\Delta(\bar{x})$ the normal cone to $\Delta$ at $\bar{x}$
$M^\perp$ the linear subspace of $\mathbb{R}^n$ orthogonal to $M \subset \mathbb{R}^n$
$\text{Pr}_K(\cdot)$ or $\text{P}_K(\cdot)$ the metric projection from $\mathbb{R}^n$
on to a closed convex subset $K \subset \mathbb{R}^n$
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>dom$f$</td>
<td>the effective domain of function $f$</td>
</tr>
<tr>
<td>$f'(\bar{x}; v)$</td>
<td>the directional derivative of $f$ at $\bar{x}$ in direction $v$</td>
</tr>
<tr>
<td>$f^0(\bar{x}; v)$</td>
<td>the Clarke generalized directional derivative of $f$ at $\bar{x}$ in direction $v$</td>
</tr>
<tr>
<td>$\partial f(\bar{x})$</td>
<td>the subdifferential of a convex function $f$ at $\bar{x}$, or the Clarke generalized gradient of a locally Lipschitz function $f$ at $\bar{x}$</td>
</tr>
<tr>
<td>$\nabla f(\bar{x})$</td>
<td>the gradient of $f$ at $\bar{x}$</td>
</tr>
<tr>
<td>$\nabla^2 f(\bar{x})$</td>
<td>the Hessian matrix of $f$ at $\bar{x}$</td>
</tr>
<tr>
<td>Sol$(P)$</td>
<td>the solution set of problem $(P)$</td>
</tr>
<tr>
<td>loc$(P)$</td>
<td>the local solution set of problem $(P)$</td>
</tr>
<tr>
<td>$S(P)$</td>
<td>the KKT point set of problem $(P)$</td>
</tr>
<tr>
<td>$v(P)$</td>
<td>the optimal value of problem $(P)$</td>
</tr>
<tr>
<td>QP</td>
<td>quadratic programming</td>
</tr>
<tr>
<td>QP$(D, A, c, b)$</td>
<td>quadratic program defined by matrices $D$, $A$ and vectors $c$, $b$</td>
</tr>
<tr>
<td>$S(D, A, c, b)$</td>
<td>the KKT point set of a quadratic program</td>
</tr>
<tr>
<td>Sol$(D, A, c, b)$</td>
<td>the solution set of a quadratic program</td>
</tr>
<tr>
<td>Sol$(c, b)$</td>
<td>the solution set of a quadratic program</td>
</tr>
<tr>
<td>$\varphi(D, A, c, b)$ or $\varphi(c, b)$</td>
<td>the optimal value function of a quadratic program</td>
</tr>
<tr>
<td>VI</td>
<td>variational inequality</td>
</tr>
<tr>
<td>loc$(D, A, c, b)$</td>
<td>the local-solution set of a quadratic program</td>
</tr>
<tr>
<td>VI$(\phi, \Delta)$</td>
<td>the VI defined by operator $\phi$ and set $\Delta$</td>
</tr>
<tr>
<td>AVI</td>
<td>affine variational inequality</td>
</tr>
<tr>
<td>Sol$(\text{VI}(\phi, \Delta))$</td>
<td>the solution set of $\text{VI}(\phi, \Delta)$</td>
</tr>
<tr>
<td>AVI$(M, q, \Delta)$</td>
<td>the AVI defined by matrix $M$, vector $q$, and set $\Delta$</td>
</tr>
<tr>
<td>Sol$(\text{AVI}(M, q, \Delta))$</td>
<td>the solution set of $\text{AVI}(M, q, \Delta)$</td>
</tr>
<tr>
<td>Sol$(M, A, q, b)$</td>
<td>the solution set of $\text{AVI}(M, q, \Delta)$ where $\Delta = {x : Ax \geq b}$</td>
</tr>
<tr>
<td>LCP</td>
<td>linear complementarity</td>
</tr>
<tr>
<td>LCP$(M, q)$</td>
<td>the LCP problem defined by matrix $M$ and vector $q$</td>
</tr>
<tr>
<td>Sol$(M, q)$</td>
<td>the solution set of LCP$(M, q)$</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>NCP</td>
<td>nonlinear complementarity</td>
</tr>
<tr>
<td>NCP(φ, Δ)</td>
<td>the NCP problem defined by φ and Δ</td>
</tr>
<tr>
<td>LFVO</td>
<td>linear fractional vector optimization</td>
</tr>
<tr>
<td>VVI</td>
<td>vector variational inequality</td>
</tr>
<tr>
<td>Sol(VP)</td>
<td>the efficient solution set of the LFVO problem (VP)</td>
</tr>
<tr>
<td>Sol(VP)(^w)</td>
<td>the weakly efficient solution set of the LFVO problem (VP)</td>
</tr>
<tr>
<td>lsc</td>
<td>lower semicontinuous</td>
</tr>
<tr>
<td>lsc property</td>
<td>lower semicontinuity property</td>
</tr>
<tr>
<td>usc</td>
<td>upper semicontinuous</td>
</tr>
<tr>
<td>usc property</td>
<td>upper semicontinuity property</td>
</tr>
<tr>
<td>OD—pair</td>
<td>origin-destination pair</td>
</tr>
<tr>
<td>plq</td>
<td>piecewise linear-quadratic</td>
</tr>
<tr>
<td>PVI</td>
<td>parametric variational inequality</td>
</tr>
</tbody>
</table>
Chapter 1

Quadratic Programming Problems

Quadratic programming problems constitute a special class of non-linear 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

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

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

\[
\|x\| = \left( \sum_{i=1}^{n} x_i^2 \right)^{1/2}
\]
for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and the scalar product

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

for all $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. 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 $\mathbb{R}^n$ 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

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

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

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

We say that $U \subseteq \mathbb{R}^n$ is a neighborhood of $x \in \mathbb{R}^n$ if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. 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 $\Delta$ the constraint set (also the feasible region) of $(P)$. Elements of $\Delta$ are said to be the feasible vectors of $(P)$. If $\Delta = \mathbb{R}^n$ 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

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

The set of all the solutions (resp., the local solutions) of $(P)$ is denoted by $\text{Sol}(P)$ (resp., $\text{loc}(P)$). We say that two mathematical programming problems are equivalent if the solution set of the first problem coincides with that of the second one.
1.1 Mathematical Programming Problems

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

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

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

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

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

**Remark 1.2.** It may happen that $\text{loc}(P) \setminus \text{Sol}(P) \neq \emptyset$. 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

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

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

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

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 $\text{Sol}(P) = \emptyset$. For example, if $\Delta = [1, +\infty) \subset \mathbb{R}$ and

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

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

There are different ways to classify mathematical programming problems:

- Convex vs. Nonconvex
- Smooth vs. Nonsmooth
- Linear vs. Nonlinear.
1. Quadratic Programming Problems

1.2 Convex Programs and Nonconvex Programs

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

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

The smallest convex set containing a set \( \Omega \subset R^n \) is called the convex hull of \( \Omega \) and it is denoted by \( \text{co}\Omega \).

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

\[
\text{epif} := \{(x, \alpha) : x \in R^n, \alpha \in R, \alpha \geq f(x)\} \quad (1.4)
\]

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 \( f : R^n \rightarrow \overline{R} \) 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),

\[
\alpha + (+\infty) = (+\infty) + \alpha = +\infty \quad \text{for} \quad -\infty < \alpha \leq +\infty,
\]
\[
\alpha + (-\infty) = (-\infty) + \alpha = -\infty \quad \text{for} \quad -\infty \leq \alpha < +\infty,
\]
\[
\alpha(+\infty) = (+\infty)\alpha = +\infty, \quad \alpha(-\infty) = (-\infty)\alpha = -\infty,
\]
\[
\text{for } 0 < \alpha \leq +\infty,
\]
\[
\alpha(+\infty) = (+\infty)\alpha = -\infty, \quad \alpha(-\infty) = (-\infty)\alpha = +\infty,
\]
\[
\text{for } -\infty \leq \alpha < 0,
\]
\[
0(+\infty) = (+\infty)0 = 0 = 0(-\infty) = (-\infty)0,
\]
\[
(-\infty) = +\infty, \quad \inf \emptyset = +\infty, \quad \sup \emptyset = -\infty.
\]

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

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

\[
f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \quad \forall x, y \in R^n, \quad \forall t \in (0,1). \quad (1.5)
\]

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

\[
t(x, \alpha) + (1 - t)(y, \beta) \in \text{epif}
\]
for all $t \in (0, 1)$ and for all $x, y \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$ 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 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ 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 $x_1, \ldots, x_k \in \mathbb{R}^n$ 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 $\Delta$ is a convex set and $f$ is a convex function.

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

$$\text{Sol}(P) = \text{loc}(P). \quad (1.6)$$

**Proof.** It suffices to show that $\text{loc}(P) \subseteq \text{Sol}(P)$ whenever $(P)$ is a convex program. Let $\bar{x} \in \text{loc}(P)$ and let $U$ be a neighborhood of $\bar{x}$ such that (1.1) holds. If $\bar{x} \notin \text{Sol}(P)$ 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 $f(\hat{x}) \in \mathbb{R} \cup \{-\infty\}$.

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

$$f(t\hat{x} + (1-t)\bar{x}) \leq tf(\hat{x}) + (1-t)f(\bar{x})$$

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 $\alpha \in \mathbb{R}$, since $(\hat{x}, \alpha) \in \text{epi} f$ and $(\bar{x}, f(\bar{x})) \in \text{epi} f$, we have

$$t(\hat{x}, \alpha) + (1-t)(\bar{x}, f(\bar{x})) \in \text{epi} f.$$
Example 1.1. Consider the problem

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

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 $\Delta$ is nonconvex. It is clear that (1.8) is equivalent to the following problem

$$\min \{ \| x - c \| : x \in \Delta \}. \quad (1.9)$$

One 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 $f_1(x) = -x + 2$, $f_2(x) = x + \frac{3}{2}$, $x \in \mathbb{R}$. Define $f(x) = \min \{ f_1(x), f_2(x) \}$ and choose $\Delta = [0, 2] \subseteq \mathbb{R}$. For these $f$ and $\Delta$, we have

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

Note that in this example $f$ is a nonconvex function, while $\Delta$ 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 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, the set

$$\text{dom}f := \{ x \in \mathbb{R}^n : -\infty < f(x) < +\infty \} \quad (1.10)$$

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

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

(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 $v \in \mathbb{R}^n$ then $f$ is said to be directionally differentiable at $\bar{x}$.

In the next two theorems, $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper convex function.
Theorem 1.1. (See Rockafellar (1970), Theorem 10.1) If $\bar{x} \in \mathbb{R}^n$ and $\delta > 0$ are such that the open ball $B(\bar{x}, \delta)$ is contained in $\text{dom} f$, 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 $\bar{x} \in \text{dom} f$ then for any $v \in \mathbb{R}^n$ 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 $\Delta \subset \mathbb{R}^n$ at a point $\bar{x} \in \mathbb{R}^n$ is defined by the formula

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

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

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

Definition 1.11. A subset $M \subset \mathbb{R}^n$ is called an affine set if $tx + (1 - t)y \in M$ for every $x \in M$, $y \in M$ and $t \in \mathbb{R}$. For a convex set $\Delta \subset \mathbb{R}^n$, the affine hull $\text{aff} \Delta$ of $\Delta$ is the smallest affine set containing $\Delta$. The relative interior of $\Delta$ 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 $\mathbb{R}^n$. If $x \notin \text{dom} f$ then $\partial f(x)$ is empty. If $x \in \text{ri} (\text{dom} f)$ then $\partial f(x)$ is nonempty and

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

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

$$x \in \text{int} (\text{dom} f),$$
1. Quadratic Programming Problems

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

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 $\mathbb{R}^n$. If

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

then

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

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 $\mathbb{R}^n$ and $\Delta \subset \mathbb{R}^n$ is a nonempty convex set. If the inclusion

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

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

$$\text{ri}(\text{dom} f) \cap \text{ri} \Delta \neq \emptyset \quad (1.15)$$

then (1.14) is a necessary and sufficient condition for $\bar{x} \in \mathbb{R}^n$ to be a solution of $(P)$. In particular, if $\Delta = \mathbb{R}^n$ 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 $\mathbb{R}^2$ 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 $\mathbb{R}^2$
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:

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

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 \( \text{dom}f_i = \mathbb{R}^2 \ (i = 1, 2, 3) \), 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,

\[
\partial f_1(\bar{x}) = \overline{B}_{\mathbb{R}^2}, \quad \partial f_2(\bar{x}) = \left\{ \frac{a - b}{\|a - b\|} \right\}, \quad \partial f_3(\bar{x}) = \left\{ \frac{a - c}{\|a - c\|} \right\}.
\]

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

\[
1 \geq \|u^*\|^2 = \langle u^*, u^* \rangle = \langle v^* + w^*, v^* + w^* \rangle = \|v^*\|^2 + \|w^*\|^2 + 2\langle v^*, w^* \rangle.
\]

As \( \|v^*\| = 1 \) and \( \|w^*\| = 1 \), this yields \( \langle v^*, w^* \rangle \leq -\frac{1}{2} \). Denoting by \( \alpha \) 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

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

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
\( \tilde{u}^* := v^* + w^* \) belongs to \( \tilde{B}_{R^2} \). 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 \( \bar{x} \not\equiv a, \bar{x} \not\equiv b \) and \( \bar{x} \not\equiv c \), i.e. \( M \) does not coincide with anyone from the three vertexes \( A, B, C \) of the triangle \( ABC \). In this case, as

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

(1.17) is equivalent to the equality

\[
0 = u^* + v^* + w^*,
\]

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),

\[
1 = \|u^*\|^2 = \langle u^*, u^* \rangle = \langle -v^* - w^*, -v^* - w^* \rangle = \|v^*\|^2 + \|w^*\|^2 + 2\langle v^*, w^* \rangle.
\]

Since \( \|v^*\| = 1 \) and \( \|w^*\| = 1 \), this implies that \( \langle v^*, w^* \rangle = -\frac{1}{2} \). Hence the geometric angle \( \alpha \) between \( v^* \) and \( w^* \) is \( 2\pi/3 \). Similarly, we deduce from (1.20) that the geometric angle \( \beta \) (resp., \( \gamma \)) 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 \( 120^\circ \).) 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 \( 120^\circ \), then \( M \equiv A \) is the unique solution of our problem.

(ii) If all the three angles of the triangle \( ABC \) are smaller than \( 120^\circ \), 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 $120^\circ$. (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 $\mathbb{R}^2$ 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 $\lim_{\|x\| \to +\infty} f(x) = +\infty$. Fix any $z \in \mathbb{R}^2$ and put $\gamma = f(z)$. Let $q \in [\|z\|, +\infty)$ be such that

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

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

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

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

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

\[
(x_1 - y_1)a_2 - (x_2 - y_2)a_1 = x_1y_2 - x_2y_1, \\
(x_1 - y_1)b_2 - (x_2 - y_2)b_1 = x_1y_2 - x_2y_1, \\
(x_1 - y_1)c_2 - (x_2 - y_2)c_1 = x_1y_2 - x_2y_1.
\]

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

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

is a straight line in \( \mathbb{R}^2 \). 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

\[
f(tx + (1-t)y) = \|tx + (1-t)y - a\| + \|tx + (1-t)y - b\| \\
+ \|tx + (1-t)y - c\| \\
= \|t(x-a) + (1-t)(y-a)\| \\
+ \|t(x-b) + (1-t)(y-b)\| \\
+ \|t(x-c) + (1-t)(y-c)\| \\
< t\|x-a\| + (1-t)\|y-a\| \\
+ t\|x-b\| + (1-t)\|y-b\| \\
+ t\|x-c\| + (1-t)\|y-c\| \\
= tf(x) + (1-t)f(y).
\]

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 $\Delta$ 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 $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and

$$\Delta = \{x \in \mathbb{R}^n : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, h_1(x) = 0, \ldots, h_s(x) = 0\},$$

(1.23)

where $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ is a convex function, $h_j : \mathbb{R}^n \to \mathbb{R}$ for $j = 1, \ldots, s$ is an affine function, i.e. there exist $a_j \in \mathbb{R}^n$ and $\alpha_j \in \mathbb{R}$ such that $h_j(x) = \langle a_j, x \rangle + \alpha_j$ for every $x \in \mathbb{R}^n$. 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 $\Delta$ is given by (1.23). Let the above assumptions on $f$, $g_i$ ($i = 1, \ldots, m$) and $h_j$ ($j = 1, \ldots, s$) be satisfied. Assume that there exists a vector $z \in \mathbb{R}^n$ such that

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

(1.24)

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

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

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

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

If \( m = 0 \) then (1.24) is equivalent to the requirement that \( \Delta \) 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 \( f : R^n \rightarrow R \) 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

\[ \nabla f(\bar{x}) = \left( \begin{array}{c} \frac{\partial f(\bar{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\bar{x})}{\partial x_n} \end{array} \right), \]

where \( \frac{\partial f(\bar{x})}{\partial x_i} \) 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

\[ \nabla^2 f(\bar{x}) = \begin{pmatrix} \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\bar{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\bar{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_n \partial x_n} \end{pmatrix}, \]

where \( \frac{\partial^2 f(\bar{x})}{\partial x_j \partial x_i} \) 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 $\mathbb{R}^n$ then $f$ is directionally differentiable on $\mathbb{R}^n$ (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 $\bar{x} \in \mathbb{R}^n$ and $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$.

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

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

**Definition 1.13.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be a locally Lipschitz near $\bar{x} \in \mathbb{R}^n$ 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 $\mathbb{R}^n$ then $f$ is said to be a locally Lipschitz function on $\mathbb{R}^n$. If $f$ is locally Lipschitz near $\bar{x}$ then the generalized directional derivative of $f$ at $\bar{x}$ in direction $v \in \mathbb{R}^n$ is defined by

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

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

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

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

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

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

for every $v \in \mathbb{R}^n$.

(b) If $f$ is a $C^1$-function then $f$ is a locally Lipschitz function and

$$\partial f(\bar{x}) = \{ \nabla f(\bar{x}) \}, \quad f^0(\bar{x}; v) = \langle \nabla f(\bar{x}), v \rangle$$

for all $\bar{x} \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$.

(c) If $f$ is convex then $f$ is a locally Lipschitz function and, for every $\bar{x} \in \mathbb{R}^n$, 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 $v \in \mathbb{R}^n$.

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 $C \subset \mathbb{R}^n$ be a nonempty subset. The Clarke tangent cone $T_C(x)$ to $C$ at $x \in C$ is the set of all $v \in \mathbb{R}^n$ satisfying $d^0_C(x; v) = 0$, where $d^0_C(x; v)$ 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.

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

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

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

(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 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 $\mathbb{R}^n$ converging to $v$ such that $x_k + t_k v_k \in C$ for all $k$. 

1.3 Smooth Programs and Nonsmooth Programs

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

\[\Delta = \{x \in C : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, h_1(x) = 0, \ldots, h_s(x) = 0\},\]

where \(C \subset \mathbb{R}^n\) is a nonempty subset, \(g_i : \mathbb{R}^n \to \mathbb{R} (i = 1, \ldots, m)\) and \(h_j : \mathbb{R}^n \to \mathbb{R} (j = 1, \ldots, s)\) 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

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

and

\[\lambda_i g_i(\bar{x}) = 0 \quad \text{for all } i = 1, 2, \ldots, m. \quad (1.27)\]

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 \(\lambda_0\) 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 \(\lambda_0\) and setting \(\tilde{\lambda}_i = \lambda_i / \lambda_0\) for \(i = 1, \ldots, m, \tilde{\mu}_j = \mu_j / \lambda_0\) for \(j = 1, \ldots, s,\) we obtain

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

and

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

Similarly as in the case of convex programs (see Theorem 1.6), if (1.28) and (1.29) are fulfilled then the numbers \(\tilde{\lambda}_1 \geq 0, \ldots, \tilde{\lambda}_m \geq 0, \tilde{\mu}_1 \in \mathbb{R}, \ldots, \tilde{\mu}_s \in \mathbb{R}\) 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).