Introduction to Optimization and Semidifferential Calculus

MOS-SIAM Series on Optimization

This series is published jointly by the Mathematical Optimization Society and the Society for Industrial and Applied Mathematics. It includes research monographs, books on applications, textbooks at all levels, and tutorials. Besides being of high scientific quality, books in the series must advance the understanding and practice of optimization. They must also be written clearly and at an appropriate level for the intended audience.

Series Volumes

Delfour, M. C., *Introduction to Optimization and Semidifferential Calculus*

Ulbrich, Michael, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*

Biegler, Lorenz T., *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*

Shapiro, Alexander, Dentcheva, Darinka, and Ruszczyński, Andrzej, *Lectures on Stochastic Programming: Modeling and Theory*

Conn, Andrew R., Scheinberg, Katya, and Vicente, Luis N., *Introduction to Derivative-Free Optimization*

Ferris, Michael C., Mangasarian, Olvi L., and Wright, Stephen J., *Linear Programming with MATLAB*

Attouch, Hedy, Buttazzo, Giuseppe, and Michaille, Gérard, *Variational Analysis in Sobolev and* BV *Spaces: Applications to PDEs and Optimization*

Wallace, Stein W. and Ziemba, William T., editors, *Applications of Stochastic Programming*

Grötschel, Martin, editor, *The Sharpest Cut: The Impact of Manfred Padberg and His Work*

Renegar, James, *A Mathematical View of Interior-Point Methods in Convex Optimization*

Ben-Tal, Aharon and Nemirovski, Arkadi, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*

Conn, Andrew R., Gould, Nicholas I. M., and Toint, Phillippe L., *Trust-Region Methods*

Introduction to Optimization and Semidifferential Calculus

M. C. Delfour
Centre de Recherches Mathématiques
and
Département de Mathématiques et de Statistique
Université de Montréal
Montréal, Canada

siam®

Mathematical
Optimization Society

Society for Industrial and Applied Mathematics
Philadelphia

Mathematical Optimization Society
Philadelphia

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Delfour, Michel C., 1943-
Introduction to optimization and semidifferential calculus / M. C. Delfour.
p. cm. – (MOS-SIAM series on optimization)
Includes bibliographical references and index.
ISBN 978-1-611972-14-6
1. Mathematical optimization. 2. Differential calculus. I. Title.
QA402.5.D348 2012
515'.642–dc23
2011040535

To Francis and Guillaume

Contents

List of Figures

Preface

A Great and Beautiful Subject

Optimization refers to finding, characterizing, and computing the *minima* and/or *maxima* of a function with respect to a *set of admissible points*.

Its early steps were intertwined with the ones of the differential calculus and the mathematical analysis. The first idea of the differential calculus and the rule for the computation of the minima and maxima could be attributed to Fermat in 1638. The concept of derivative was introduced in that context by Leibniz and Newton almost fifty years later. So, the condition obtained by Fermat for the extremum of an algebraic function was de facto generalized in the form $f'(x) = 0$. With the introduction of the notion of differentiable function of several variables and of differentiable functions defined on Hilbert and topological vector spaces, the rule of Fermat remains valid. One of the important areas of optimization is the *calculus of variations*, which deals with the minimization/maximization of *functionals*, that is, functions of functions. It was also intertwined with the development of *classical analysis* and *functional analysis*.

But, optimization is not just *mathematical analysis*. Many decision-making problems in operations research, engineering, management, economics, computer sciences, and statistics are formulated as *mathematical programs* requiring the maximization or minimization of an *objective function* subject to constraints. Such programs[1] often have special structures: linear, quadratic, convex, nonlinear, semidefinite, dynamic, integer, stochastic programming, etc. This was the source of more theory and efficient algorithms to compute solutions. With the easier access to increasingly more powerful computers, larger and more complex problems were tackled thus creating a demand for efficient computer software to solve *large-scale systems*.

To give a few landmarks, the modern form of the multipliers rule goes back to Lagrange[2] in his path-breaking *Mécanique analytique* in 1788 and the *steepest descent method* to Gauss.[3] The *simplex algorithm* to solve *linear programming*[4] problems was created by

[1] The term *programming* in this context does not refer to computer programming. Rather, the term comes from the use of program by the United States military to refer to proposed training and logistics schedules, which were the problems that Dantzig was studying at the time.

[2] Joseph Louis, comte de Lagrange (in Italian Giuseppe Lodovico Lagrangia) (1736–1813).

[3] Johann Carl Friedrich Gauss (1777–1855).

[4] Much of the theory had been introduced by Leonid Vitaliyevich Kantorovich (1912–1986) in 1939 (L .V. KANTOROVICH [**1**, **2**]).

George Dantzig[5] and the *theory of the duality* was developed by John von Neumann[6] both in 1947. The necessary conditions for inequality-constrained problems were first published in the Masters thesis of William Karush in 1939, although they became renowned after a seminal conference paper by Harold W. Kuhn and Albert W. Tucker in 1951.

Intended Audience and Objectives of the Book

This book is intended as a textbook for a one-term course at the undergraduate level for students in Mathematics, Physics, Engineering, Economics, and other disciplines with a basic knowledge of mathematical analysis and linear algebra. It is intentionally limited to the optimization with respect to variables belonging to finite-dimensional spaces. This is what we call *finite-dimensional optimization*. It provides a lighter exposition deferring at the graduate level technical questions of Functional Analysis associated with the Calculus of Variations. The useful background material has been added at the end of the first chapter to make the book self-sufficient. The book can also be used for a first year graduate course or as a companion to other textbooks.

Being limited to one term, choices had to be made. The classical themes of optimization are covered emphasizing the semidifferential calculus while staying at a level accessible to an undergraduate student. In the making of the book, some material has been added to the original lecture notes. For a one-term basic program the sections and subsections beginning with the black triangle ► can be skipped. The book is structured in such a way that the basic program only requires very basic notions of analysis and the Hadamard semidifferential that is easily accessible to nonmathematicians as an extension of their elementary one-dimensional differential calculus. The added material makes the book more interesting and provides connections with *convex analysis* and, to a lesser degree, *subdifferentials*. Yet, the book does not pretend or aim at covering everything. The added material is not mathematically more difficult since it only involves more *liminf* and *limsup* in the definitions of lower and upper semidifferentials, but it might be too much for a basic undergraduate course.

For a first initiation to nondifferentiable optimization, *semidifferentials* have been preferred over *subdifferentials*[7] that necessitate a good command of set-valued analysis. The emphasis will be on Hadamard semidifferentiable[8] functions for which the resulting semidifferential calculus retains all the nice features of the classical differential calculus, including the good old chain rule. Convex continuous and semiconvex functions are Hadamard semidifferentiable and an explicit expression of the semidifferential of an extremum with respect to parameters can be obtained. So, it works well for most nondifferentiable optimization problems including semiconvex or semiconcave problems. The Hadamard semidifferential calculus readily extends to functions defined on differential manifolds and on groups that naturally occur in optimization problems with respect to the shape or the geometry.[9]

[5] George Bernard Dantzig (1914–2005) (G. B. DANTZIG [**1**, **3**]).

[6] John von Neumann (1903–1957).

[7] For a treatment of finite-dimensional optimization based on subdifferentials and the *generalized gradient*, the reader is referred to the original work of R. T. ROCKAFELLAR [**1**] and F. H. CLARKE [**2**] and to the more recent book of J. M. BORWEIN and A. S. LEWIS [**1**].

[8] The differential in the sense of Hadamard goes back to the beginning of the 20th century. We shall go back to the original papers of J. HADAMARD [**2**] in 1923 and of M. FRÉCHET [**3**] in 1937.

[9] The reader is referred to the book of M. C. DELFOUR and J.-P. ZOLÉSIO [**1**].

The book is written in the mathematical style of definitions, theorems, and detailed proofs. It is not necessary to go through all the proofs, but it was felt important to have all the proofs in the book. Numerous examples and exercises are incorporated in each chapter to illustrate and better understand the subject material. In addition, the answer to all the exercises is provided in Appendix B. This considerably expands the set of examples and enriches the theoretical content of the book. More exercises along with examples of applications in various fields can be found in other books such as the ones of S. BOYD and L. VANDENBERGHE [**1**] and F. BONNANS [**1**].

The purpose of the historical commentaries and landmarks is mainly to put the subject in perspective and to situate it in time.

Numbering and Referencing Systems

The *numbering* of equations, theorems, lemmas, corollaries, definitions, examples, and remarks is by chapter. When a reference to another chapter is necessary it is always followed by the words *in Chapter* and the *number of the chapter*. For instance, "equation (7.5) from Theorem 7.4(i) of Chapter 2" or "Theorem 5.2 from section 5 in Chapter 3." The text of theorems, lemmas, and corollaries is slanted; the text of definitions, examples, and remarks is normal shape and ended by a square □. This makes it possible to aesthetically emphasize certain words especially in definitions. The bibliography is by author in alphabetical order. For each author or group of coauthors there is a numbering in square brackets starting with [**1**]. A reference to an item by a single author is of the form J. HADAMARD [**2**] and a reference to an item with several coauthors is of the form H. W. KUHN and A. W. TUCKER [**1**]. *Boxed formulae* or *statements* are used in some chapters for two distinct purposes. First, they emphasize certain important definitions, results, or identities; second, in long proofs of some theorems, lemmas, or corollaries they isolate key intermediary results which will be necessary to more easily follow the subsequent steps of the proof.

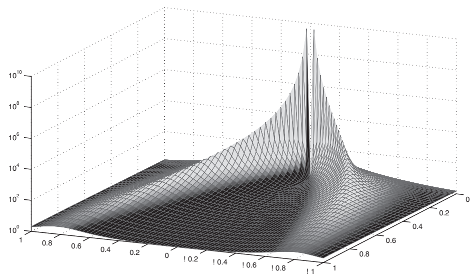
Acknowledgments

This book is based on an undergraduate course created in 1975 at the University of Montreal by Andrzej Manitius (George Mason University) who wrote a first set of lecture notes. The course and its content were reorganized and modified in 1984 by the author and this book is the product of their evolution.

The author is grateful to the referees and to Thomas Liebling, Editor-in-Chief of the MOS-SIAM Series on Optimization, for their very constructive suggestions, without forgetting the students whose contribution has been invaluable.

Michel Delfour

Montreal, May 16, 2011



Chapter 1
Introduction

1 Minima and Maxima

A first objective is to seek the weakest conditions for the existence of an admissible point achieving the extremum (a minimum or a maximum) of an objective function. This will require notions such as continuity, compactness, and convexity and their relaxation to weaker notions of semicontinuity, bounded lower or upper sections, and quasiconvexity. When the set of admissible points is specified by differentiable constraint functions, *dualizing* the necessary optimality condition naturally leads to the introduction of the *Lagrange multipliers* and the *Lagrangian* for which the number of unknowns is increased by the number of multipliers associated with each constraint function.

A second objective is the characterization of the points achieving an extremum. In most problems this requires the differentiability of the objective function and that the set of admissible points be specified by functions that are also differentiable. Otherwise, it is still possible to obtain a characterization of the *extremizers* by going to weaker notions of *semidifferential* and using a local approximation of the set of admissible points by tangent cones. In particular, the convex continuous functions are semidifferentiable. Thus, the minimum of a convex continuous objective function over a convex set of admissible points can be completely characterized by using semidifferentials. When the objective function is not convex, fairly general necessary optimality conditions can also be obtained by using still weaker notions of upper or lower semidifferentials.

As it is seldom possible to explicitly solve the equations that characterize an extremum, it is natural to use numerical methods. It is a broad area of activity strongly stimulated by the availability of and access to more and more powerful computers.

2 Calculus of Variations and Its Offsprings

The sources of this section are B. van Brunt,[1] J. Ferguson,[2] and Wikipedia: Calculus of Variations and Optimization (mathematics).

The first idea of the differential calculus and the rule for the computation of the minima and maxima[3] seems to go back to Fermat[4] in 1638. It is generally accepted that the concept of derivative is due to Leibniz[5] who published in 1684[6] and Newton[7] who published over a longer period of time.[8] So, the condition obtained by Fermat for the extremum of an algebraic function was de facto generalized in the form $f'(x) = 0$. With the introduction of the concept of differentiable function of several variables of Jacobi and of differentiable functions defined on Hilbert and topological vector spaces, the rule of Fermat and Leibniz remains valid. During three centuries, it was applied, justified, adapted, and generalized in the context of the theory of optimization, of the calculus of variations, and of the theory of optimal control.

The *calculus of variations* deals with the minimization/maximization of *functionals*, that is, functions of functions. The simplest example of this type of problems is to find the curve of minimum length between two points. In the absence of constraints, the solution is a straight line between the points. However, when the curve is constrained to stay on a surface in the space, the solution is less obvious and not necessarily unique. Such solutions are called *geodesics*. A problem of that nature is illustrated by *Fermat's principle* in optics: the light follows the shortest path of *optical length* between two points, where the optical length depends on the material of the physical medium. A notion of the same nature in mechanics is the *least action principle*. Several important problems involve functions of several variables. For instance, solutions of boundary value problems for the Laplace equation satisfy the Dirichlet principle.

According to some historians, the calculus of variations begins with the brachistochrone problem of Johann Bernoulli in 1696. It immediately attracted the attention of Jakob Bernoulli and of the Marquis de l'Hôpital, but it is Euler who was the first to develop this subject. His contributions began in 1733 and it is his *Elementa Calculi Variationum* that gave its name to this discipline. In 1744, Euler published his landmark book *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti*.[9] In two papers read to the Académie des Sciences in 1744, and to the Prussian Academy in 1746, Maupertuis proclaimed the *principle of least action*.

[1] Bruce van Brunt, *The Calculus of Variations*, Springer-Verlag, New York, 2004.

[2] James Ferguson, *A. Brief Survey of the History of the Calculus of Variations and its Applications*, University of Victoria, Canada, 2004 (arXiv:math/0402357).

[3] *Methodus ad disquirendam Maximam et Minimam*, 1638 (see P. DE FERMAT [**1**]).

[4] Pierre de Fermat (1601–1665).

[5] Gottfried Wilhelm Leibniz (1646–1716).

[6] *Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractas nec irrationales quantitates moratur, et singulare pro illis calculi genus* (A new method for maxima and minima and their tangents, that are not limited to fractional or irrational expressions, and a remarkable type of calculus for these), in Acta Eruditorum, 1684, a journal created in Leipzig two years earlier.

[7] Sir Isaac Newton (1643–1728).

[8] The *Method of Fluxions* completed in 1671 and published in 1736 and *Philosophiæ Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy), often called the *Principia* (Principles), 1687 and 1726 (third edition).

[9] A method for discovering curved lines that enjoy a maximum or minimum property, or the solution of the isoperimetric problem taken in the widest sense.

Lagrange extensively contributed to the theory and Legendre[10] in 1786 laid the foundations of the characterization of maxima and minima. In his path-breaking *Mécanique analytique*[11] in 1788, Lagrange summarized all the work done in the field of classical mechanics since Newton.[12] It is in that book that he clearly expressed the multipliers rule in its modern form.

In the 18th century several mathematicians have contributed to this enterprise, but perhaps the most important work of the century is that of Weierstrass starting in 1870. He was the first to give a completely correct proof of a sufficient condition for the existence of a minimum. His celebrated course on the theory is epoch-making, and it may be asserted that he was the first to place it on a firm and unquestionable foundation. The 20th and the 23rd Hilbert problems published in 1900 enticed further development. In the 20th century, Noether, Tonelli, Lebesgue, and Hadamard, among others, made significant contributions. Morse applied the calculus of variations in what is now called Morse theory. Pontryagin, Rockafellar, and Clarke developed new mathematical tools for the *optimal control theory*, a generalization of the calculus of variations.

In problems of *geometrical design* and *shape optimization*, the modelling, optimization, or control variable is no longer a vector of scalars or functions but the shape of a geometrical object.[13] In this category, we find for instance *Plateau's problem*. This problem was raised by Lagrange in 1760, but it is named after Joseph Plateau who was interested in soap films around 1873.[14] It consists in finding the surface of minimum area with a given boundary in space. Such surfaces are easy to find experimentally, but their mathematical interpretation is extremely complex. In this family of problems we also encounter several identification problems such as the reconstruction and the processing of images from two dimensional or three dimensional scanned material or biological (biometrics) objects. Another area is the *enhancing of images* such as in the *Very Large Telescope* project. To deal with nonparametrized geometrical objects, nonlinear and nonconvex spaces and a *shape differential calculus* are required.

3 Contents of the Book

Chapter 2 is devoted to the existence of minimizers of a real-valued function. Natural notions of lower and upper semicontinuities and several notions of convexity are introduced to relax the conditions of the classical Weierstrass theorem specialized to the case of the minimum. Some elements of *convex analysis*, such as the notion of *Fenchel–Legendre transform*, the *primal* and *dual problems*, and the *Fenchel duality theorem* are also included along with *Ekeland's variational principle* and some of its consequences.

Chapter 3 first provides a review of the differentiability of functions of one and several variables. Notions of semidifferentials and, specifically, of Hadamard semidifferentials are introduced. The connection is made with the classical Gateaux[15] and Fréchet differentials. A semidifferential calculus that extends the classical differential calculus is developed. In this framework, Hadamard semidifferentials of lower and upper envelopes of finitely many

[10] *Sur la manière de distinguer les maxima des minima dans le calcul des variations* (On the method of distinguishing maxima from minima in the calculus of variations).

[11] J. L. LAGRANGE [**2**].

[12] *Philosophiæ naturalis principia mathematica*.

[13] Cf., for instance, M. C. DELFOUR and J.-P. ZOLÉSIO [**1**].

[14] Joseph Antoine Ferdinand Plateau (1801–1883). Cf. J. A. PLATEAU [**1**].

[15] Without circumflex accent (see footnote 19 on page 80).

Hadamard semidifferentiable functions exist and the *chain rule* for the composition of Hadamard semidifferentiable functions remains valid. Moreover, the convex continuous functions and, more generally, the semiconvex functions are Hadamard semidifferentiable. Upper and lower notions of semidifferentials are also included in connection with semiconvex functions. Finally, we give a fairly general theorem on the semidifferentiability of extrema with respect to a parameter. It is applied to get the explicit expression of the Hadamard subdifferential of the extremum of quadratic functions. The differentials and semidifferentials introduced in this chapter are summarized and compared in the last section.

Chapter 4 focuses on optimality conditions to characterize an unconstrained or a constrained extremum via the semidifferential or differential of the objective function. It first considers twice differentiable functions without constraints along with several examples. A special attention is given to the generic example of the least and greatest eigenvalues of a symmetric matrix. The explicit expression of their Hadamard semidifferentials is provided with the help of the general theorems of Chapter 3. This is followed by the necessary and sufficient optimality condition for convex differentiable objective functions and convex sets of admissible points. It is specialized to linear subspaces, affine subspaces, and convex cones at the origin. Finally, a necessary and sufficient condition for an arbitrary convex objective function and an arbitrary set of constraints is given in terms of the lower semidifferential.

The second part of the chapter gives a general necessary optimality condition for a local minimum using the upper Hadamard semidifferentiability of the objective function and the Bouligand's tangent cone of admissible directions. It is quite remarkable that such simple notions be sufficient to cover most of the so-called *nondifferentiable* optimization. This condition is dualized by introducing the notion of dual cone. The dual necessary optimality condition is then applied to the *linear programming* problem where the constraints are specified by a finite number of equalities and inequalities on affine functions. At this juncture, the *Lagrangian* is introduced along with its connections to two-person zero-sum games and to Fenchel's primal and dual problems of Chapter 2. A general form of Farkas' lemma is given in preparation of the next chapter. The constructions and results are extended to the *quadratic programming* problem and to Fréchet differentiable objective functions. At the end of this chapter, a glimpse is given into optimization via subdifferentials that involves *set-valued functions*.

Chapter 5 is devoted to differentiable optimization where the set of admissible points is specified by a finite number of differentiable constraint functions. By using the dual necessary optimality condition, we recover the Lagrange multiplier theorem for equality constraints, the Karush–Kuhn–Tucker theorem for inequality constraints, and the general theorem for the mixed case of equalities and inequalities.

4 Some Background Material in Classical Analysis

This section puts together a compact summary of some basic elements of classical analysis that will be needed in the other chapters. They come from several sources (for instance, among others, W. H. FLEMING [**1**], W. RUDIN [**1**], or L. SCHWARTZ [**1**]). The differential calculus will be completely covered from scratch in Chapter 3 and does not require any prerequisite. The various notions of convexity that will be needed will be introduced in each chapter, but the reader is also referred to specialized books such as, for instance, F. A. VALENTINE [**1**], R. T. ROCKAFELLAR [**1**], L. D. BERKOVITZ [**1**], S. R. LAY [**1**], H. TUY [**1**], S. BOYD and L. VANDENBERGHE [**1**].

4.1 Greatest Lower Bound and Least Upper Bound

Let $\mathbb{R}$ denote the set of real numbers and let $|x|$ denote the *absolute value* of x. The following notation will be used for positive and strictly positive real numbers

$$\mathbb{R}_+ \stackrel{\text{def}}{=} \{x \in \mathbb{R} : x \geq 0\} \quad \text{and} \quad \mathbb{R}^+ \stackrel{\text{def}}{=} \{x \in \mathbb{R} : x > 0\}$$

and the notation $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ for the *extended real numbers*.

Definition 4.1.
Let $\varnothing \neq A \subset \mathbb{R}$.

(a) $b^0 \in \mathbb{R}$ is a *least upper bound* of A if

(i) b^0 is an upper bound of A and

(ii) for all upper bounds M of A, we have $b^0 \leq M$.

The least upper bound b^0 of A is unique and is denoted $\sup A$. If A is not bounded above, set $\sup A = +\infty$.

(b) $b_0 \in \mathbb{R}$ is a *greatest lower bound* of A if

(i) b_0 is a lower bound of A and

(ii) for all lower bounds m of A, we have $b_0 \geq m$.

The greatest lower bound b_0 of A is unique and is denoted $\inf A$. If A is not bounded below, set $\inf A = -\infty$. □

Remark 4.1. (i) When $A \neq \varnothing$, we always have $-\infty \leq \inf A \leq \sup A \leq +\infty$. By definition, $\sup A \in \mathbb{R}$ if and only if A is bounded above and $\inf A \in \mathbb{R}$ if and only if A is bounded below.

(ii) When $A = \varnothing$, we write by convention $\sup A = -\infty$ and $\inf A = +\infty$. At first sight it might be shocking to have $\sup A < \inf A$, but, from a mathematical point of view, it is the right choice since $\sup A < \inf A$ if and only if $A = \varnothing$ or, equivalently, $\sup A \geq \inf A$ if and only if $A \neq \varnothing$. □

We shall often use the following equivalent conditions.

Theorem 4.1. *Let $\varnothing \neq A \subset \mathbb{R}$.*

(a) b^0 *is the* least upper bound *of A if and only if*

(i) b^0 *is an upper bound of A and*

(ii′) *for all M such that $b^0 > M$, there exists $x_0 \in A$ such that $b^0 \geq x_0 > M$.*

(b) b_0 *is the* greatest lower bound *of A if and only if*

(i) b_0 *is a lower bound of A and*

(ii′) *for all m such that $b_0 < m$, there exists $x_0 \in A$ such that $b_0 \leq x_0 < m$.*

(c) $\sup A = +\infty$ *if and only if, for all $M \in \mathbb{R}$, there exists $x_0 \in A$ such that $x_0 > M$.*

(d) $\inf A = -\infty$ *if and only if, for all $m \in \mathbb{R}$, there exists $x_0 \in A$ such that $x_0 < m$.*

4.2 Euclidean Space

Most results in this book remain true in general vector spaces of functions and in groups of transformations of infinite dimension. In this book the scope is limited to vector spaces of finite dimension that will be identified with the Cartesian product $\mathbb{R}^n$. For instance, such spaces include the space of all polynomials of degree less than or equal to $n-1$, $n \geq 1$, an integer. In this section, we recall some definitions, notions, and theorems from Classical Analysis.

4.2.1 Cartesian Product, Balls, and Continuity

Given an integer $n \geq 1$, let

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}} \tag{4.1}$$

be the *Cartesian product* of dimension n with the following notation:

an element $\quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ or in vectorial form $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

the *norm* $\quad \|x\|_{\mathbb{R}^n} = \left[\sum_{i=1}^n x_i^2\right]^{1/2}$ and the *inner product* $x \cdot y = \sum_{i=1}^n x_i y_i$. $\qquad$ (4.2)

The norm will be simply written $\|x\|$ when no confusion arises and the arrow on top of the vector $\vec{x}$ will often be dropped. When $n = 1$, $\|x\|_{\mathbb{R}^1}$ coincides with the absolute value $|x|$. $\mathbb{R}^n$ with the *scalar multiplication* and the *addition*

$$\forall \alpha \in \mathbb{R}, x \in \mathbb{R}^n, \quad \alpha x \stackrel{\text{def}}{=} (\alpha x_1, \ldots, \alpha x_n)$$
$$\forall x, y \in \mathbb{R}^n, \quad x + y \stackrel{\text{def}}{=} (x_1 + y_1, \ldots, x_n + y_n)$$

is a *vector space* on $\mathbb{R}$ of dimension n.

Definition 4.2.
The *canonical orthonormal basis* of $\mathbb{R}^n$ is the set $\{e_i^n \in \mathbb{R}^n : 1 \leq i \leq n\}$ defined by

$$(e_i^n)_j \stackrel{\text{def}}{=} \delta_{ij}, \quad \delta_{ij} \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j, \end{cases}$$

that is,

$$e_1^n = (1,0,0,\ldots,0,0), \quad e_2^n = (0,1,0,\ldots,0,0), \quad \ldots, \quad e_n^n = (0,0,0,\ldots,0,1).$$

In particular, $e_i^n \cdot e_j^n = \delta_{ij}$. □

When no confusion arises, we simply write $\{e_i\}$ without the supersript n.

A *Euclidean space* is a vector space E that can be identified with $\mathbb{R}^n$ via a linear bijection for some integer $n \geq 1$. For instance, we can identify with $\mathbb{R}^n$ the space $P^{n-1}[0,1]$ of polynomials of order less than or equal to $n-1$ on the interval $[0,1]$:

$$p \mapsto (p(0), p'(0), \ldots, p^{(n-1)}(0)) : P^{n-1}[0,1] \to \mathbb{R}^n$$
$$(p_0, p_1, \ldots, p_{n-1}) \mapsto p(x) \stackrel{\text{def}}{=} \sum_{i=0}^{n-1} p_i \frac{x^i}{i\,!} : \mathbb{R}^n \to P^{n-1}[0,1].$$

4.2.2 Open, Closed, and Compact Sets

The notions *open* and *closed sets* in $\mathbb{R}^n$ can be defined starting from balls.
Ball at x of radius $r > 0$:

$$\begin{array}{ll} \text{open} & B_r(x) = \{y \in \mathbb{R}^n : \|y - x\| < r\} \\ \text{closed} & \overline{B_r(x)} = \{y \in \mathbb{R}^n : \|y - x\| \leq r\}. \end{array}$$

Unit ball at 0:

$$\text{open } B = \{y \in \mathbb{R}^n : \|y\| < 1\}, \quad \text{closed } \overline{B} = \{y \in \mathbb{R}^n : \|y\| \leq 1\}.$$

Punched open ball at x:

$$B'_r(x) = \{y \in \mathbb{R}^n : 0 < \|y - x\| < r\}$$

Definition 4.3.
Let $U \subset \mathbb{R}^n$.

(i) $a \in \mathbb{R}^n$ is an *interior point* of U, if there exists $r > 0$ such that $B_r(a) \subset U$.

(ii) The *interior* of U is the set of all interior points of U. It will be denoted by $\operatorname{int} U$. By definition $\operatorname{int} U \subset U$.

(iii) $V(x)$ is a *neighborhood* of x if there exists $r > 0$ such that $B_r(x) \subset V(x)$.

(iv) A is an *open subset* of $\mathbb{R}^n$ if for all $x \in A$, there exists a neighborhood $V(x)$ of x such that $V(x) \subset A$.

(v) The family $\mathcal{T}$ of all open sets in $\mathbb{R}^n$ is called the *topology* on $\mathbb{R}^n$ generated by the norm. □

The *topology* $\mathcal{T}$ of $\mathbb{R}^n$ is equal to the family of all finite intersections and arbitrary unions of open balls in $\mathbb{R}^n$.

Definition 4.4. (i) A sequence $\{x_n\}$ in $\mathbb{R}^n$ is *convergent* if there exists a point $x \in \mathbb{R}^n$ such that

$$\forall \varepsilon > 0, \exists N, \forall n > N, \quad \|x_n - x\|_{\mathbb{R}^n} < \varepsilon.$$

The point x is unique and called the *limit point* of $\{x_n\}$.

(ii) $\{x_n\}$ in $\mathbb{R}^n$ is a *Cauchy sequence* if

$$\forall \varepsilon > 0, \exists N, \forall n, m > N, \quad \|x_n - x_m\|_{\mathbb{R}^n} < \varepsilon. \qquad \square$$

A convergent sequence is a Cauchy sequence. All Cauchy sequences in $\mathbb{R}^n$ are convergent to points in $\mathbb{R}^n$. We say that the space $\mathbb{R}^n$ is *complete* for the topology $\mathcal{T}$.

The notions of *limit point* and *closed set* can be specified in several ways. We do it via the notions of *accumulation point* and *isolated point*.

Definition 4.5.
Let U be a subset of $\mathbb{R}^n$.

(i) $a \in U$ is an *isolated point* of U if there exists $r > 0$ such that $B'_r(a) \cap U = \varnothing$.

(ii) $a \in \mathbb{R}^n$ is an *accumulation point* of U if $B'_r(a) \cap U \neq \varnothing$ for all $r > 0$. $\square$

Definition 4.6. (i) $a \in \mathbb{R}^n$ is a *limit point* of U if $B_r(a) \cap U \neq \varnothing$ for all $r > 0$.

(ii) The *closure* of U is the set of all limit points of U. It will be denoted $\overline{U}$.

(iii) F is a *closed set* if it contains all its limit points. $\square$

Remark 4.2. (i) Equivalently, x is a *limit point* of U if, for all neighborhoods $V(x)$ of x, $V(x) \cap U \neq \varnothing$.

(ii) The closure of U is equal to the union of all its isolated and accumulation points. Hence $U \subset \overline{U}$.

(iii) The only subsets of $\mathbb{R}^n$ that are both open and closed are $\varnothing$ and $\mathbb{R}^n$. $\square$

Definition 4.7.
Let A and B be two subsets of $\mathbb{R}^n$.

(i) $A \backslash B \stackrel{\text{def}}{=} \{x \in A : x \notin B\}$. When $A = \mathbb{R}^n$, we write $\complement B$ or $\mathbb{R}^n \backslash B$ and say that $\complement B$ is the *complement* of B in $\mathbb{R}^n$.

(ii) The *boundary* of $U \subset \mathbb{R}^n$ is defined as $\overline{U} \cap \overline{\complement U}$. It will be denoted ∂U. $\square$

It is easy to check that $\partial U = \overline{U} \backslash \operatorname{int} U$, $\overline{U} = \operatorname{int} U \cup \partial U$, and $\overline{\complement U} = \operatorname{int} \complement U \cup \partial U$.

Definition 4.8. (i) A family $\{G_\alpha\}$ of open subsets of $\mathbb{R}^n$ is an *open cover* of $X \subset \mathbb{R}^n$ if $X \subset \cup_\alpha G_\alpha$.

(ii) A nonempty subset K of $\mathbb{R}^n$ is said to be *compact* if each open covering $\{G_\alpha\}$ of K has a finite subcover $\{G_{\alpha_i} : 1 \leq i \leq k\}$. $\square$

Theorem 4.2 (Heine–Borel)**.** *Let $\varnothing \neq U \subset \mathbb{R}^n$. U is compact if and only if U is closed and bounded.*[16] [17]

[16]Heinrich Eduard Heine (1821–1881).
[17]Félix Edouard Justin Émile Borel (1871–1956).

In a normed vector space E, a compact subset U of E is closed and bounded, but the converse is generally not true except in finite-dimensional normed vector spaces.

Theorem 4.3 (Bolzano–Weierstrass)**.** *If U is an infinite subset of a compact set K in a metric space, then U has a limit point in K.*[18] [19]

As a consequence of this theorem, we have the following useful result.

Theorem 4.4. *Let K, $\varnothing \neq K \subset \mathbb{R}^n$, be compact. Any sequence $\{x_n\}$ in K has a convergent subsequence $\{x_{n_k}\}$ to a point x of K:*

$$\exists \{x_{n_k}\} \text{ and } \exists x \in K \text{ such that } x_{n_k} \to x \in K.$$

4.3 Functions

4.3.1 Definitions and Convention

Let $n \geq 1$ and $m \geq 1$ be two integers and $f : D_f \to \mathbb{R}^m$ be a function defined on a *domain* D_f in $\mathbb{R}^n$. Since it is always possible to arbitrarily extend the definition of a function from its initial domain of definition D_f to all of $\mathbb{R}^n$, we adopt the following convention.

Convention 4.1.
All functions $f : D_f \subset \mathbb{R}^n \to \mathbb{R}^m$ in this book have domain $D_f = \mathbb{R}^n$. □

Definition 4.9. (i) A *real-valued function of a real variable* is a function $f : \mathbb{R} \to \mathbb{R}$, and a *real-valued function of several real variables* is a function $f : \mathbb{R}^n \to \mathbb{R}$, $n \geq 2$.

(ii) A *vector function* is a function $f : \mathbb{R}^n \to \mathbb{R}^m$ for some $m \geq 2$. □

Definition 4.10.
Let $n \geq 1$ and $m \geq 1$ be two integers. Denote by

$$\{e_i^m \in \mathbb{R}^m : 1 \leq j \leq m\} \quad \text{and} \quad \{e_i^n \in \mathbb{R}^n : 1 \leq i \leq n\}$$

the respective canonical orthonormal bases associated with $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively.

(i) A function $A : \mathbb{R}^n \to \mathbb{R}^m$ is *linear* if

$$\forall x, y \in \mathbb{R}^n, \forall \alpha, \beta \in \mathbb{R}, \quad A(\alpha x + \beta y) = \alpha\, A(x) + \beta\, A(y). \tag{4.3}$$

(ii) By convention, the $m \times n$ matrix $\{A_{ij}\}$ associated with A will also be denoted A:

$$A_{ij} \overset{\text{def}}{=} e_i^m \cdot A e_j^n, \quad Ax \cdot y = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} x_j\, y_i. \tag{4.4}$$

□

[18] Bernard Placidus Johann Nepomuk Bolzano (1781–1848).

[19] Karl Theodor Wilhelm Weierstrass (1815–1897) was the leader of a famous school of mathematicians who undertook the systematic revision of various sectors of mathematical analysis.

4.3.2 Continuity of a Function

Definition 4.11.
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ for two integers $n \geq 1$ and $m \geq 1$. The function f is *continuous* at $x \in \mathbb{R}^n$ if $\forall \varepsilon > 0$, $\exists \delta(x) > 0$ such that

$$\forall y \text{ such that } \|y - x\|_{\mathbb{R}^n} < \delta(x), \quad \|f(y) - f(x)\|_{\mathbb{R}^m} < \varepsilon.$$

The function f is continuous on $U \subset \mathbb{R}^n$ if it is continuous at every point of U. □

The notion of continuity for a function $f : \mathbb{R}^n \to \mathbb{R}^m$ can be defined in terms of open balls. Indeed, the definition involves the open ball $B_\varepsilon(f(x))$ in $\mathbb{R}^m$ and the open ball $B_{\delta(x)}(x)$ in $\mathbb{R}^n$. The condition becomes: for each open ball $B_\varepsilon(f(x))$ of radius $\varepsilon > 0$, there exists an open ball $B_{\delta(x)}(x)$ in $\mathbb{R}^n$ such that $B_{\delta(x)}(x) \subset f^{-1}\{B_\varepsilon(f(x))\}$. This yields the following equivalent criterion in terms of neighborhoods.

Theorem 4.5. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ for two integers $n \geq 1$ and $m \geq 1$. The function f is* continuous *at $x \in \mathbb{R}^n$ if and only if for each neighborhood W of $f(x)$ in $\mathbb{R}^m$, $f^{-1}\{W\}$ is a neighborhood of x in $\mathbb{R}^n$.*

Proof. For any neighborhood W of $f(x)$, there exists $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subset W$. If f is continuous at x, then, by definition, there exists $\delta(x) > 0$ such that $B_{\delta(x))}(x) \subset f^{-1}\{B_\varepsilon(f(x))\}$. Since $f^{-1}\{B_\varepsilon(f(x))\} \subset f^{-1}\{W\}$, $f^{-1}\{B_\varepsilon(f(x))\}$ is indeed a neighborhood of x by definition. Conversely, for all $\varepsilon > 0$, the open ball $B_\varepsilon(f(x))$ is a neighborhood of $f(x)$. Then $f^{-1}\{B_\varepsilon(f(x))\}$ is a neighborhood of x. So there exists an open ball $B_{\delta(x)}(x)$ of radius $\delta(x) > 0$ such that $B_{\delta(x)}(x) \subset f^{-1}\{B_\varepsilon(f(x))\}$. Hence we get the ε-δ definition of the continuity of f at x. □

Theorem 4.6. *A linear function $A : \mathbb{R}^n \to \mathbb{R}^m$ is continuous on $\mathbb{R}^n$.*

Proof. It is sufficient to prove it for a linear function $A : \mathbb{R}^n \to \mathbb{R}$. Any point $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ can be written

$$x = \sum_{i=1}^{n} x_i \, e_i$$

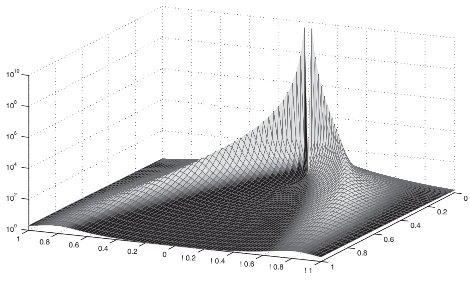
and, by linearity of A,

$$Ax = \sum_{i=1}^{n} x_i \, Ae_i = x \cdot g, \quad g \stackrel{\text{def}}{=} \begin{bmatrix} Ae_1 \\ \vdots \\ Ae_n \end{bmatrix}.$$

The vector g is unique. For any $\varepsilon > 0$, choose $\delta = \varepsilon/(\|g\| + 1)$. Thence,

$$\begin{aligned} \forall y, \|y - x\| < \delta \quad \Rightarrow |Ay - Ax| = |A(y - x)| &= |g \cdot (y - x)| \\ &\leq \|g\| \, \|y - x\| \\ &< \|g\| \, \delta = \frac{\|g\|}{\|g\| + 1} \varepsilon < \varepsilon. \end{aligned}$$

Therefore, A is continuous on $\mathbb{R}^n$. □



Chapter 2
Existence, Convexities, and Convexification

1 Introduction

In this chapter, $\mathbb{R}^n$ will be the Cartesian product endowed with the scalar product and the norm (4.2) of Chapter 1, $f : \mathbb{R}^n \to \mathbb{R}$ or $\mathbb{R} \cup \{+\infty\}$ an *objective function*, and U a nonempty subset of $\mathbb{R}^n$.

The Weierstrass theorem provides conditions on U and f for the existence of points in U achieving both the *infimum* $\inf f(U)$ and the *supremum* $\sup f(U)$: compactness of U and continuity of f on U. In fact, it is sufficient to consider the minimization problem since the one of maximization can be obtained by minimizing the negative of the objective function. By restricting our analysis to the infimum, the class of objective functions can be enlarged to functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and the continuity on U can be relaxed to the weaker notion of *lower semicontinuity* that includes many discontinuous functions. Growth conditions at infinity will complete the results when U is closed but not bounded. In the absence of compactness, we also give *Ekeland's variational principle* and some of its ramifications such as the existence theorem of Takahashi and the fixed point theorem of Caristi. All results are true in finite-dimensional vector spaces and basic ideas and constructions generalize to function spaces.

The last part of the chapter is devoted to *convexity* that plays a special role in the context of a minimization problem. If, in addition to existence, the convexity of f is strict, then the minimizing point is unique. For convex objective functions all local infima are global and hence equal. This suggests *convexifying* the objective function and finding the infimum of the convexified function rather than the global infimum of the original function that can have several local infima. This leads to the work of Legendre, Fenchel, and Rockafellar, the introduction of the *Fenchel–Legendre transform*, the primal and dual problems, and the *Fenchel duality theorem* that will be seen again in Chapter 4 in the context of linear and quadratic programming.

2 Weierstrass Existence Theorem

The fact that $\inf f(U)$ is finite does not guarantee the existence of a point $a \in U$ that achieves the infimium, $f(a) = \inf f(U)$, as illustrated in the following example.

Example 2.1.
Let $U = \mathbb{R}$ and consider the function

$$f(x) = 1 \text{ if } x \leq 0 \text{ and } f(x) = x \text{ if } x > 0$$

for which $\inf f(U) = 0$ and $f(x) \neq 0$ for all $x \in U = \mathbb{R}$. □

The Weierstrass[1] theorem that will be proved later as Theorem 5.1 is fundamental in optimization. It gives sufficient conditions on U and f for the existence of minimizers and maximizers in U.

Theorem 2.1 (Weierstrass)**.** *Given a compact nonempty subset U of $\mathbb{R}^n$ and a real-valued function $f : U \to \mathbb{R}$ that is continuous at U,*[2]

(i) $\exists a \in U$ *such that* $f(a) = \sup_{x \in U} f(x)$,

(ii) $\exists b \in U$ *such that* $f(b) = \inf_{x \in U} f(x)$.

But it is a little too strong since it gives the existence of both minimizing and maximizing points. Indeed, it is not necessary to simultaneously seek the existence of the two types of points since a supremum can always be formulated as an infimum of the negative of the function and vice versa: $\sup f(U) = -\inf -f(U)$. It will be sufficient to find conditions for the existence of a minimizing point of f in U. In so doing, it will be possible to weaken the continuity assumption that is clearly not necessary for the piecewise continuous functions of Figure 2.1 that reach a minimum at a point of $[0,1]$. Notice that at the points of discontinuity, we have chosen to give the function the lower value and not the upper value that would not have resulted in the existence of a minimizing point.

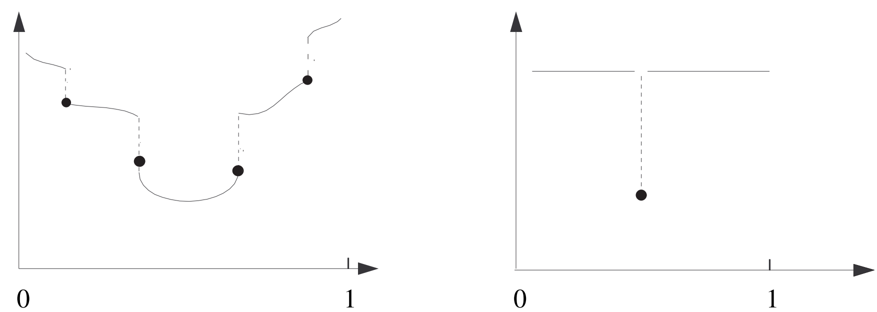


Figure 2.1. *Discontinuous functions having a minimizing point in* $[0,1]$.

3 Extrema of Functions with Extended Values

The $\inf f(U)$ and the $\sup f(U)$ have been defined for real-valued functions, that is, $f(U) \subset \mathbb{R}$. When the set $f(U)$ is unbounded below, $\inf f(U) = -\infty$ and when $f(U)$ is unbounded above, $\sup f(U) = +\infty$ (see Definition 4.1 and Remark 4.1 of Chapter 1).

[1] Karl Theodor Wilhelm Weierstrass (1815–1897).

[2] We can also work with an $f : \mathbb{R}^n \to \mathbb{R}$ continuous on U since, for U closed in $\mathbb{R}^n$, any $f : U \to \mathbb{R}$ continuous on U for the *relative topology* can be extended to a continuous function on $\mathbb{R}^n$.

The idea of extended real-valued objective functions implicitly having effective domains is due to R. T. Rockafellar[3] and J. J. Moreau.[4] In order to consider functions $f : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ that are possibly equal to $+\infty$ or $-\infty$ at some points, the definitions of the $\inf f(U)$ and the $\sup f(U)$ have to be extended.

Definition 3.1.
Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $U \subset \mathbb{R}^n$.

(i) Associate with f its *effective domain*

$$\operatorname{dom} f \overset{\text{def}}{=} \left\{ x \in \mathbb{R}^n : -\infty < f(x) < +\infty \right\}. \tag{3.1}$$

It will also be simply referred to as the *domain* of f.

(ii) The *infimum* of f with respect to U is defined as follows:

$$\inf f(U) \overset{\text{def}}{=} \begin{cases} +\infty, & \text{if } U = \varnothing, \\ \inf f(U), & \text{if } U \neq \varnothing \text{ and } f(U) \subset \mathbb{R}, \\ +\infty, & \text{if } U \neq \varnothing \text{ and } \forall x \in U,\ f(x) = +\infty, \\ -\infty, & \text{if } \exists x \in U \text{ such that } f(x) = -\infty. \end{cases}$$

We shall also use the notation $\inf_{x \in U} f(x)$.

The *supremum* of f with respect to U is defined as follows:

$$\sup f(U) \overset{\text{def}}{=} \begin{cases} -\infty, & \text{if } U = \varnothing, \\ \sup f(U), & \text{if } U \neq \varnothing \text{ and } f(U) \subset \mathbb{R}, \\ -\infty, & \text{if } U \neq \varnothing \text{ and } \forall x \in U,\ f(x) = -\infty, \\ +\infty, & \text{if } \exists x \in U \text{ such that } f(x) = +\infty. \end{cases}$$

We shall also use the notation $\sup_{x \in U} f(x)$.

Infima and suprema constitute the set of *extrema* of f in U.

(iii) When there exists $a \in U$ such that $f(a) = \inf f(U)$, f is said to reach its *minimum* at a point of U and it is written as

$$\min f(U) \text{ or } \min_{x \in U} f(x).$$

---

[3] Ralph Tyrrell Rockafellar (1935–). "Moreau and I independently in those days at first, but soon in close exchanges with each other, made the crucial changes in outlook which, I believe, created convex analysis out of convexity. For instance, he and I passed from the basic objects in Fenchel's work, which were pairs consisting of a convex set and a finite convex function on that set, to extended real-valued functions implicitly having effective domains, for which we moreover introduced set-valued subgradient mappings." R. T. Rockafellar, http://www.convexoptimization.com/wikimization/index.php/Rockafellar.

[4] Jean Jacques Moreau (1923–) "...appears as a rightful heir to the founders of differential calculus and mechanics through the depth of his thinking in the field of nonsmooth mechanics and the size of his contribution to the development of nonsmooth analysis. His interest in mechanics has focused on a wide variety of subjects: singularities in fluid flows, the initiation of cavitation, plasticity, and the statics and dynamics of granular media. Allied to this is his investment in mathematics in the fields of convex analysis, calculus of variations and differential measures" (see P. ALART, O. MAISONNEUVE, and R. T. ROCKAFELLAR [**1**]).

The *set of all minimizing points of f in U* is denoted

$$\operatorname{argmin} f(U) \stackrel{\text{def}}{=} \{a \in U : f(a) = \inf f(U)\}. \tag{3.2}$$

When there exists $b \in U$ such that $f(b) = \sup f(U)$, f is said to reach its *maximum* at a point of U, and it is written as

$$\max f(U) \text{ or } \max_{x \in U} f(x).$$

The *set of all maximizing points of f in U* is denoted

$$\operatorname{argmax} f(U) \stackrel{\text{def}}{=} \{b \in U : f(b) = \sup f(U)\}. \tag{3.3}$$

□

With the above extensions, $\sup f(U)$ can still be replaced by $-\inf(-f(U))$ and vice versa as can be seen from the next theorem.

Theorem 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ and $U \subset \mathbb{R}^n$.*

$$\sup f(U) = -\inf(-f)(U) \quad \textit{and} \quad \operatorname{argmax} f(U) = \operatorname{argmin}(-f)(U). \tag{3.4}$$

Proof. If $U = \varnothing$, then $f(U) = \varnothing$. By convention, $\sup f(U) = -\infty$ and $\inf(-f)(U) = +\infty$. Hence $\sup f(U) = -\infty = -\inf(-f)(U)$.

Assume now that $U \neq \varnothing$. First eliminate the trivial cases. If there exists $x \in U$ such that $f(x) = +\infty$, then $\sup f(U) = +\infty$, $-f(x) = -\infty$, and $\inf -f(U) = -\infty$. The second case is $f(x) = -\infty$ for all $x \in U$ which implies that $\sup f(U) = -\infty$ and, for all $x \in U$, $-f(x) = +\infty$ and $\inf -f(U) = +\infty$.

The last case is $U \neq \varnothing$ and $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ for which there exists $x \in U$ such that $f(x) > -\infty$. Therefore $\sup f(U) > -\infty$.

(i) Assume that $b^0 = \sup f(U) \in \mathbb{R}$. From Definition 4.1 of Chapter 1, b^0 is an upper bound, that is, for all $x \in U$, $f(x) \leq b^0$, and all upper bounds M of $f(U)$ are such that $b^0 \leq M$. Therefore for all $x \in U$, $-f(x) \geq -b^0$ and $-b^0$ is a lower bound of $-f(U) = \{-f(x) : x \in U\}$. Let m be a lower bound of $-f(U)$. Then, $-m$ is an upper bound of $f(U)$ and since b^0 is the least upper bound, $b^0 \leq -m$. Thence, $-b^0 \geq m$ and $-b^0$ is the largest lower bound of $-f(U)$. This yields $\inf -f(U) = -b^0 = -\sup f(U)$ and $-\inf -f(U) = b^0 = \sup f(U)$.

(ii) By convention in Remark 4.1 of Chapter 1, the case $b^0 = \sup f(U) = +\infty$ corresponds to $f(U)$ not bounded above. Therefore, there exists a sequence $\{x_n\} \subset U$ such that $f(x_n) \to +\infty$. This implies that $-f(x_n) \to -\infty$ and hence the set $-f(U)$ is not bounded below. By convention, $\inf -f(U) = -\infty$ and $\sup f(U) = +\infty = -\inf -f(U)$. □

Introducing objective functions with values $\pm\infty$ makes it possible to replace an infimum of f with respect to U by an infimum over all of $\mathbb{R}^n$ by introducing the new function

$$x \mapsto f_U(x) \stackrel{\text{def}}{=} \left\{ \begin{array}{ll} f(x), & \text{if } x \in U \\ +\infty, & \text{if } x \in \mathbb{R}^n \backslash U \end{array} \right\} : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}. \tag{3.5}$$

Theorem 3.2. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Then*

$$\inf f(U) = \inf f_U(U) = \inf f_U(\mathbb{R}^n) \quad \text{and} \quad \operatorname{argmin} f(U) = U \cap \operatorname{argmin} f_U(\mathbb{R}^n).$$

If, in addition, $\inf(U) < +\infty$*, then* $\operatorname{argmin} f(U) = \operatorname{argmin} f_U(\mathbb{R}^n)$.

Proof. By definition, $f(x) \le f_U(x)$ for all $x \in \mathbb{R}^n$ and

$$\inf f(U) = \inf f_U(U) \ge \inf f_U(\mathbb{R}^n).$$

If $\inf f_U(\mathbb{R}^n) = +\infty$, we trivially have equality. If $m \overset{\text{def}}{=} \inf f_U(\mathbb{R}^n) \in \mathbb{R}$, then for all $n \in \mathbb{N}$, there exists $x_n \in \mathbb{R}^n$ such that

$$m + \frac{1}{n} > f_U(x_n) \ge m.$$

Since $f_U(x_n)$ is finite, $x_n \in U$ and $f_U(x_n) = f(x_n)$. Thence

$$\inf f(U) \ge m > f(x_n) - \frac{1}{n} \ge \inf f(U) - \frac{1}{n}.$$

By letting n go to infinity, we get equality.

As $\inf f(U) = \inf f_U(U) = \inf f_U(\mathbb{R}^n)$ and $U \subset \mathbb{R}^n$, we have $\operatorname{argmin} f(U) = \operatorname{argmin} f_U(U) \subset \operatorname{argmin} f_U(\mathbb{R}^n)$. If, in addition, $\inf f(U) < +\infty$, then, by definition of f_U, $\operatorname{argmin} f_U(\mathbb{R}^n) \subset \operatorname{argmin} f_U(U) = \operatorname{argmin} f(U)$. □

Remark 3.1.
For a supremum, extend f by $-\infty$ by considering the function

$$f^U(x) \overset{\text{def}}{=} \left\{ \begin{array}{ll} f(x), & \text{if } x \in U \\ -\infty, & \text{if } x \in \mathbb{R}^n \backslash U \end{array} \right\} = -(-f)_U(x). \tag{3.6}$$

□

For an infimum, two cases are trivial:

(i) there exists $x \in U$ such that $f(x) = -\infty$ that yields $\inf f_U(\mathbb{R}^n) = \inf f(U) = -\infty$ and $x \in \operatorname{argmin} f(U)$;

(ii) for all $x \in U$, $f(x) = +\infty$ that yields $\inf f_U(\mathbb{R}^n) = \inf f(U) = +\infty$ and $\operatorname{argmin} f(U) = U$.

In order to exclude cases (i) and (ii) for the infimum, we introduce the notion of *proper function for an infimum* that is the natural extension of the notion of proper function for convex functions (see R. T. ROCKAFELLAR [**1**]). The dual notion of *proper function for a supremum* of f is obtained by considering the notion of proper function for the infimum of $-f$.

Definition 3.2.
Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$.

(i) f is said to be *proper for the infimum* if

(a) for all $x \in \mathbb{R}^n$, $f(x) > -\infty$ and

(b) there exists $x \in \mathbb{R}^n$ such that $f(x) < +\infty$.

This is equivalent to $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and dom $f \neq \varnothing$.

(ii) f is said to be *proper for the supremum* if

(a) for all $x \in \mathbb{R}^n$, $f(x) < +\infty$ and

(b) there exists $x \in \mathbb{R}^n$ such that $f(x) > -\infty$.

This is equivalent to $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ and dom $f \neq \varnothing$.

Whenever no confusion arises, we shall simply say that the function is *proper*. □

Another trivial case occurs when dom f is a singleton; that is, there is only one point where the function f is finite.

4 Lower and Upper Semicontinuities

In order to consider the infimum of a discontinuous function, we weaken the notion of continuity by breaking it into two weaker notions.

Recall that a real-valued function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous at $a \in \mathbb{R}^n$ if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall x \in B_\delta(a), \quad |f(x) - f(a)| < \varepsilon. \tag{4.1}$$

The open ball $B_\delta(a)$ of radius $\delta > 0$ is a neighborhood of a. Letting $V(a) = B_\delta(a)$, the condition on f yields the following two conditions:

$$\begin{aligned} &\forall x \in V(a), \quad -\varepsilon < f(x) - f(a) && \Rightarrow f(a) - \varepsilon < f(x) \\ &\forall x \in V(a), \quad f(x) - f(a) < \varepsilon && \Rightarrow f(x) < f(a) + \varepsilon. \end{aligned} \tag{4.2}$$

The first condition says that $f(a)$ is *below* all limit points of $f(x)$ as x goes to a, while the second one says that $f(a)$ is *above*, thus yielding the decomposition of the continuity into lower semicontinuity and upper semicontinuity.

Definition 4.1. (i) $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is *lower semicontinuous* at $a \in \mathbb{R}^n$ if

$$\forall h < f(a), \exists \text{ a neighborhood } V(a) \text{ of } a \text{ such that } \forall x \in V(a), \ h < f(x). \tag{4.3}$$

f is *lower semicontinuous* on $U \subset \mathbb{R}^n$ if it is lower semicontinuous at every point of U. By convention, the function identically equal to $-\infty$ is lower semicontinuous.

(ii) $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ is *upper semicontinuous* at $a \in \mathbb{R}^n$ if

$$\forall k > f(a), \exists \text{ a neighborhood } V(a) \text{ of } a \text{ such that } \forall x \in V(a), \ k > f(x). \tag{4.4}$$

f is *upper semicontinuous* on $U \subset \mathbb{R}^n$ if it is upper semicontinuous at every point of U. By convention, the function identically equal to $+\infty$ is upper semicontinuous.

In short, we shall write lsc for lower semicontinuous and usc for upper semicontinuous. □

Functions of Figure 2.1 are lower semicontinuous (lsc) in $]0,1[$. The function identically equal to $+\infty$ is lsc and the one identically equal to $-\infty$ is usc on $\mathbb{R}^n$. As we have seen before, definition (4.1) of the continuity of a function $f:\mathbb{R}^n\to\mathbb{R}$ at a point $a\in\mathbb{R}^n$ is equivalent to the two conditions (4.2): the first one is the lower semicontinuity at a with $h=f(a)-\varepsilon<f(a)$ and the second one is the upper semicontinuity at a with $k=f(a)+\varepsilon>f(a)$.

As for the infimum and the supremum where $\sup f(U)=-\inf -f(U)$, f is usc at x if and only if $-f$ is lsc at x. So it is sufficient to study the properties of lsc functions.

Theorem 4.1. (i) *$f:\mathbb{R}^n\to\mathbb{R}\cup\{-\infty\}$ is* usc *at x if and only if $-f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is* lsc *at x.*

(ii) *$f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is* lsc *at x if and only if $-f:\mathbb{R}^n\to\mathbb{R}\cup\{-\infty\}$ is* usc *at x.*

Proof. As $f:\mathbb{R}^n\to\mathbb{R}\cup\{-\infty\}$, then $-f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$. Given $h<-f(x)$, then $f(x)<-h$. As f is usc at x, there exists a neighborhood $V(x)$ of x such that for all $y\in V(x)$, $f(y)<-h$. As a result, for all $y\in V(x)$, $-f(y)>h$. By definition, $-f$ is lsc at x. The proof of the converse is similar. □

It is easy to check the following properties of lsc functions (see Exercises 10.1 to 10.4) by using the convention $(+\infty)+(+\infty)=+\infty$, $(+\infty)+a=+\infty$ for all $a\in\mathbb{R}$, and $(+\infty)a=(a/\|a\|)\infty$ for all $a\in\mathbb{R}$ not equal to 0.

Theorem 4.2. (i) *For all $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ and $g:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$* lsc *at $a\in\mathbb{R}^n$, the function*

$$(f+g)(x)\overset{\text{def}}{=}f(x)+g(x),\quad \forall x\in\mathbb{R}^n,$$

is lsc *at a.*

(ii) *For all $\lambda\geq 0$ and $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$* lsc *at $a\in\mathbb{R}^n$, the function*

$$(\lambda f)(x)\overset{\text{def}}{=}\begin{Bmatrix}\lambda f(x), & \text{if } \lambda>0\\ 0, & \text{if } \lambda=0\end{Bmatrix},\qquad \forall x\in\mathbb{R}^n,$$

is lsc *at a.*

(iii) *Given a family $\{f_\alpha\}_{\alpha\in A}$ (where A is an index set possibly infinite) of functions $f_\alpha:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$* lsc *at $a\in\mathbb{R}^n$, the* upper envelope

$$\left\{\sup_{\alpha\in A} f_\alpha\right\}(x)\overset{\text{def}}{=}\sup_{\alpha\in A} f_\alpha(x),\quad x\in\mathbb{R}^n,$$

is lsc *at $a\in\mathbb{R}^n$.*

(iv) *Given a finite family $f_i:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$, $1\leq i\leq m$, of functions* lsc *at $a\in\mathbb{R}^n$, the* lower envelope

$$\left\{\min_{1\leq i\leq m} f_i\right\}(x)\overset{\text{def}}{=}\min_{1\leq i\leq m} f_i(x),\quad x\in\mathbb{R}^n,$$

is lsc *at $a\in\mathbb{R}^n$.*

(v) *Given a function* $f : \mathbb{R}^n \to \mathbb{R}$ *and a point* $a \in \mathbb{R}^n$,

$$f \text{ is continuous at } a \iff f \text{ is lsc and usc at } a.$$

(vi) *Given a linear map* $A : \mathbb{R}^m \to \mathbb{R}^n$ *and a function* $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ *lsc at* Ax, *then* $f \circ A : \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ *is lsc at* x.

Property (iv) is not necessarily true for the lower envelope of an infinite number of lsc functions as can be seen from the following example.

Example 4.1.
Define for each integer $k \geq 1$, the continuous function

$$f_k(x) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } x \in [0,1], \\ 1 - k(x-1), & \text{if } x \in [1, 1+1/k], \\ 0, & \text{if } x \in [1+1/k, 2]. \end{cases}$$

It is easy to check that

$$\inf_{k \geq 1} f_k(x) = \begin{cases} 1, & \text{if } x \in [0,1], \\ 0, & \text{if } x \in]1,2], \end{cases}$$

is usc but not lsc at $x = 1$. □

The lower semicontinuity (resp., upper semicontinuity) can also be characterized in terms of the lim inf (resp., lim sup).

Definition 4.2.
Given a function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ (resp., $f : \mathbb{R}^n \to \mathbb{R}\cup\{-\infty\}$),

$$\liminf_{x\to a} f(x) \stackrel{\text{def}}{=} \sup_{\varepsilon>0} \inf_{\substack{x\neq a \\ \|x-a\|<\varepsilon}} f(x) \left(\text{resp., } \limsup_{x\to a} f(x) \stackrel{\text{def}}{=} \inf_{\varepsilon>0} \sup_{\substack{x\neq a \\ \|x-a\|<\varepsilon}} f(x) \right).$$

□

Theorem 4.3. $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ (resp., $f : \mathbb{R}^n \to \mathbb{R}\cup\{-\infty\}$) *is* lsc (resp., usc) *at* a *if and only if*

$$\liminf_{x\to a} f(x) \geq f(a) \quad \left(\text{resp., } \limsup_{x\to a} f(x) \leq f(a) \right).$$

Proof. ($\Rightarrow$) If f is lsc at a, for all $h < f(a)$, there exists a neighborhood $V(a)$ of a such that for all $x \in V(a)$, $f(x) > h$. As $V(a)$ is a neighborhood of a, there exists a ball $B_\varepsilon(a)$, $\varepsilon > 0$, such that $B_\varepsilon(a) \subset V(a)$ and

$$\forall x \in B_\varepsilon(a), \quad f(x) > h \quad \Rightarrow \inf_{\substack{x\in B_\varepsilon(a) \\ x\neq a}} f(x) \geq h \quad \Rightarrow \sup_{\varepsilon>0} \inf_{\substack{x\in B_\varepsilon(a) \\ x\neq a}} f(x) \geq h.$$

Since the inequality is true for all $h < f(a)$, letting h go to $f(a)$, we get

$$\liminf_{x\to a} f(x) \geq h \quad \Rightarrow \liminf_{x\to a} f(x) \geq f(a).$$

($\Leftarrow$) For all h such that $f(a) > h$, by assumption,

$$\liminf_{x\to a} f(x) = \sup_{\varepsilon>0}\left[\inf_{\substack{x\in B_\varepsilon(a)\\ x\neq a}} f(x)\right] \geq f(a) > h.$$

By definition of the sup, for that h, there exists $\varepsilon > 0$ such that

$$\sup_{\varepsilon>0}\left[\inf_{\substack{x\in B_\varepsilon(a)\\ x\neq a}} f(x)\right] \geq \inf_{\substack{x\in B_\varepsilon(a)\\ x\neq a}} f(x) > h \quad \Rightarrow \forall x \in B_\varepsilon(a),\ f(x) > h.$$

As $B_\varepsilon(a)$ is a neighborhood of a, f is lsc at a. $\square$

We have as a corollary the following characterization of the epigraph.

Lemma 4.1. *$f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is* lsc *on $\mathbb{R}^n$ if and only if the* epigraph *of f,*

$$\operatorname{epi} f \overset{\text{def}}{=} \left\{(x,\mu) \in \mathbb{R}^n \times \mathbb{R} : x \in \operatorname{dom} f \text{ and } \mu \geq f(x)\right\}, \tag{4.5}$$

is closed in $\mathbb{R}^n \times \mathbb{R}$. The epigraph epi f *is nonempty if and only if* dom $f \neq \varnothing$*, that is, when f is proper for the infimum.*

Remark 4.1.
The effective domain dom f of an lsc function is not necessarily closed, as can be seen from the example of the function $f(x) = 1/|x|$ if $x \neq 0$ and $+\infty$ if $x = 0$, where dom $f = \mathbb{R}\backslash\{0\}$. $\square$

Proof. If f is lsc on $\mathbb{R}^n$, consider a Cauchy sequence $(x_n,\mu_n) \in \operatorname{epi} f$. By definition, $\mu_n \geq f(x_n)$ and there exists $(x,\mu) \in \mathbb{R}^n \times \mathbb{R}$ such that $x_n \to x$ and $\mu_n \to \mu$. As f is lsc on $\mathbb{R}^n$,

$$\mu = \lim_{n\to\infty} \mu_n = \liminf_{n\to\infty} \mu_n \geq \liminf_{n\to\infty} f(x_n) \geq f(x)$$

and $(x,\mu) \in \operatorname{epi} f$. Hence the epigraph of f is closed in $\mathbb{R}^n \times \mathbb{R}$. Conversely, assume that epi f is closed in $\mathbb{R}^n \times \mathbb{R}$. Let $x \in \mathbb{R}^n$ and $h < f(x)$. Then the point $(x,h) \notin \operatorname{epi} f$. Therefore, there exists a neighborhood $W(x,h)$ such that $W(x,h) \cap \operatorname{epi} f = \varnothing$. In particular, there exists a neighborhood $V(x)$ of x such that $V(x) \times \{h\} \subset W(x,h)$ and hence for all $y \in V(x)$, $f(y) > h$ and f is lsc on $\mathbb{R}^n$. $\square$

We now give a few characterizations of lower semicontinuity in preparation of the proof of Theorem 5.1.

Lemma 4.2. *$f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is* lsc *at $a \in \mathbb{R}^n$ if and only if*

$$\forall h < f(a), \quad G_h \overset{\text{def}}{=} \{x \in \mathbb{R}^n : f(x) > h\} \tag{4.6}$$

is a neighborhood of a (see Figure 2.2*).*

Proof. The proof is by definition. $\square$

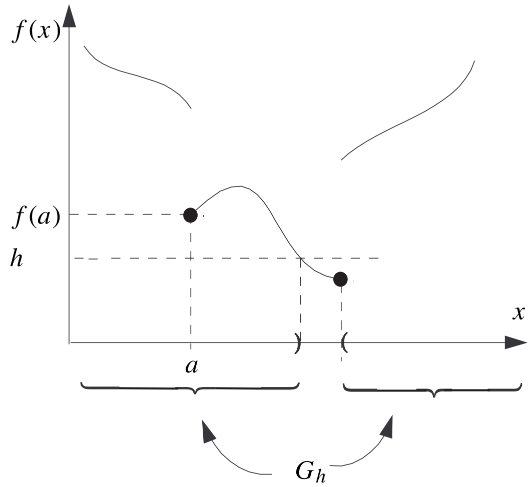


Figure 2.2. *Example of an lsc function.*

Lemma 4.3. *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. *The following conditions are equivalent:*

(i) f *is* lsc *on* $\mathbb{R}^n$;

(ii) $\forall k \in \mathbb{R}$, $G_k = \{x \in \mathbb{R}^n : f(x) > k\}$ *is open in* $\mathbb{R}^n$;

(iii) $\forall k \in \mathbb{R}$, $F_k = \{x \in \mathbb{R}^n : f(x) \leq k\}$ *is closed in* $\mathbb{R}^n$.

Proof. (i) $\Rightarrow$ (ii). If $G_k = \varnothing$, G_k is open. If $G_k \neq \varnothing$, for all $a \in G_k$, $f(a) > k$ and, as f is lsc at a, there exists a neighborhood $V(a)$ of a such that

$$\forall x \in V(a), \quad f(x) > k \quad \Rightarrow \quad V(a) \subset G_k.$$

Therefore $a \in \operatorname{int} G_k$ and G_k is open.

(ii) $\Rightarrow$ (i). By definition, $\mathbb{R}^n = \cup_{k \in \mathbb{R}} G_k$ and for each $a \in \mathbb{R}^n$, there exists $k \in \mathbb{R}$ such that $k < f(a)$. In particular, $a \in G_k \neq \varnothing$. As G_k is open by assumption, G_k is a neighborhood of a. Finally, by definition of G_k, for all $x \in G_k$, $f(x) > k$. Choose $V(a) = G_k$ and, always by definition, f is lsc at a and hence on $\mathbb{R}^n$.

(ii) $\Longleftrightarrow$ (iii) is obvious. $\square$

To use the above lemmas for a function $f : U \to \mathbb{R} \cup \{+\infty\}$ that is lsc only on the subset U of $\mathbb{R}^n$ for the relative topology on U requires the following lemma.

Lemma 4.4. *Let* U, $\varnothing \neq U \subset \mathbb{R}^n$, *be closed and let* $f : U \to \mathbb{R} \cup \{+\infty\}$ *be* lsc *for the relative topology on* U. *The function* f_U *is* lsc *on* $\mathbb{R}^n$.

Proof. Given $a \in \mathbb{R}^n \backslash U$, f_U is lsc on $\mathbb{R}^n \backslash U$. Indeed $V(a) = \mathbb{R}^n \backslash U$ is a nonempty open set containing a since U is closed. So it is a neighborhood of a. For all $h < f(a) = +\infty$,

$$\forall x \in V(a), \quad h < +\infty = f(x)$$

and, by definition, f_U is lsc on $\mathbb{R}^n \backslash U$. As for all $a \in U$, $f_U(a) = f(a)$ and f is lsc on U. So there exists a neighborhood $V(a)$ of a in $\mathbb{R}^n$ such that

$$\forall x \in V(a) \cap U, \quad h < f(x) \quad \Rightarrow \forall x \in V(a), \quad h < f(x) \leq f_U(x),$$

since, by construction, $f_U(x) = +\infty$ in $\mathbb{R}^n \setminus U$. So f_U is also lsc on U. □

Example 4.2.
The *indicator function* of a closed subset U of $\mathbb{R}^n$,

$$I_U(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } x \in U, \\ +\infty, & \text{if } x \notin U, \end{cases}$$

is lsc on $\mathbb{R}^n$. In fact, $I_U = f_U$ for the function $x \mapsto f(x) = 0 : \mathbb{R}^n \to \mathbb{R}$. □

To complete this section, we introduce the "*lower semicontinuous hull*" $\mathrm{cl}\, f$ of a function f, in the terminology of R. T. ROCKAFELLAR [**1**], that corresponds to the *lower semicontinuous regularization* of f in I. EKELAND and R. TEMAM [**1**].

Definition 4.3. (i) The *lsc regularization* of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined as the upper envelope of all lsc functions less than or equal to f:

$$\mathrm{cl}\, f(x) \stackrel{\text{def}}{=} \sup_{\substack{g \text{ lsc and} \\ g \leq f \text{ on } \mathbb{R}^n}} g(x). \tag{4.7}$$

If there exists g lsc on $\mathbb{R}^n$ such that $g \leq f$ on $\mathbb{R}^n$, then $\mathrm{cl}\, f$ is lsc on $\mathbb{R}^n$. Otherwise, set $\mathrm{cl}\, f(x) = -\infty$, by convention.

(ii) The *usc regularization* of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ is defined as the lower envelope of all usc functions greater than or equal to f:

$$\mathrm{cl}_{\,\mathrm{usc}} f(x) \stackrel{\text{def}}{=} \inf_{\substack{g \text{ usc and} \\ f \leq g \text{ on } \mathbb{R}^n}} g(x). \tag{4.8}$$

If there exists g usc on $\mathbb{R}^n$ such that $f \leq g$ on $\mathbb{R}^n$, then $\mathrm{cl}_{\,\mathrm{usc}} f$ is usc on $\mathbb{R}^n$. Otherwise, set $\mathrm{cl}_{\,\mathrm{usc}} f(x) = +\infty$, by convention. □

Note that the definition of the usc regularization amounts to $\mathrm{cl}_{\,\mathrm{usc}} f = -\mathrm{cl}\,(-f)$.

5 Existence of Minimizers in U

5.1 U Compact

We can now weaken the assumptions of the Weierstrass theorem (Theorem 2.1) by separating infimum problems from supremum problems.

Theorem 5.1. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be compact.*

(i) *If $f : U \to \mathbb{R}\cup\{+\infty\}$ is* lsc *on U, then*

$$\exists a \in U \text{ such that } f(a) = \inf f(U). \tag{5.1}$$

If $U \cap \operatorname{dom} f \neq \varnothing$, then $\inf f(U) \in \mathbb{R}$.

(ii) *If $f : U \to \mathbb{R}\cup\{-\infty\}$ is* usc *on U, then*

$$\exists b \in U \text{ such that } f(b) = \sup f(U). \tag{5.2}$$

If $U \cap \operatorname{dom} f \neq \varnothing$, then $\sup f(U) \in \mathbb{R}$.

As U is compact, it is closed and, by Lemma 4.4, f_U is lsc on $\mathbb{R}^n$ without changing the infimum since $\inf f(U) = \inf f_U(\mathbb{R}^n)$ by Theorem 3.2. So we could work with an lsc function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$. Similarly, for $f : \mathbb{R}^n \to \mathbb{R}\cup\{-\infty\}$ usc on closed U, the function f^U defined in (3.6) of Remark 3.1 is usc on $\mathbb{R}^n$.

Proof of Theorem 5.1. Let $m = \inf f(U)$. As U is compact, it is closed. By Lemma 4.4, the function f_U associated with f defined by (3.5) is lsc on $\mathbb{R}^n$ and by Theorem 3.2 we have $m = \inf f(U) = \inf f_U(U) = \inf f_U(\mathbb{R}^n)$.

If $m = +\infty$, then f is identically $+\infty$ on U and all points of U are minimizers. If $m < +\infty$, then for all reals $k > m$, the set $F_k = \{x \in U : f(x) \leq k\} = \{x \in \mathbb{R}^n : f_U(x) \leq k\}$ is closed by the lower semicontinuity of f_U on $\mathbb{R}^n$ (Lemma 4.3). It is also nonempty since, by definition of the inf, for all k such that $m < k$, there exists $f(x) \in f(U)$ such that $m = \inf f(U) \leq f(x) < k$.

Since U is compact, the closed subsets $F_k \subset U$ are also compact. By definition of the F_k's,

$$m < k_1 \leq k_2 \implies F_{k_1} \subset F_{k_2}$$

and hence any finite family of sets F_k has a nonempty intersection.

We claim that $\cap_{k>m} F_k \neq \varnothing$. By contradiction, if the intersection is empty, then for any $K > m$,

$$F_K \cap \Big[\cap_{\substack{m<k\\ k\neq K}} F_k\Big] = \varnothing \quad \Rightarrow F_K \subset \complement\Big[\cap_{\substack{m<k\\ k\neq K}} F_k\Big] = \cup_{\substack{m<k\\ k\neq K}} \complement F_k.$$

Therefore, $\{\complement F_k : m < k \text{ and } k \neq K\}$ is an open cover of the compact F_K. So, there exists a finite subcover of F_K:

$$F_K \subset \cup_{j=1}^m \complement F_{k_j} = \complement\Big[\cap_{j=1}^m F_{k_j}\Big] \quad \Rightarrow F_K \cap F_{k_1} \cap \cdots \cap F_{k_m} = \varnothing.$$

This contradicts the nonempty finite intersection property.

So any point

$$a \in \cap_{k>m} F_k \subset U,$$

belongs to U and

$$\forall k > m,\;\; f(a) \leq k \quad \Rightarrow \quad f(a) \leq m = \inf_{x\in U} f(x) \leq f(a).$$

Hence $a \in U$ is a minimizer and $\operatorname{argmin} f(U) = \cap_{k>m} F_k$. $\square$

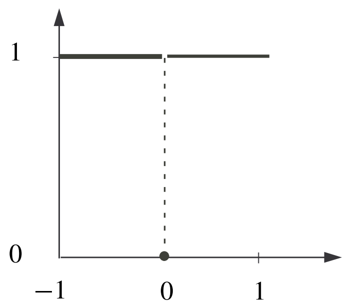


Figure 2.3. *Lsc function that is not* usc *at* 0.

In general, the infimum over U cannot be replaced by the infimum over $\overline{U}$ even if f is lsc. We only have

$$\inf f(U) \geq \inf f(\overline{U})$$

as seen from the following example.

Example 5.1 (see Figure 2.3)**.**
Consider $U = [-1,1] \backslash \{0\}$ and the lsc function

$$f(x) = \begin{cases} 1, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Then $\overline{U} = [-1,1]$ and

$$\inf f(U) = 1 > 0 = \inf f(\overline{U}).$$

The function f is not usc at 0 since for $1/2 > 0 = f(0)$, the set

$$\{x \in [-1,1] : f(x) < 1/2\} = \{0\}$$

is not a neighborhood of 0. □

However, we have the following sufficient condition.

Theorem 5.2. *Let* U, $\varnothing \neq U \subset \mathbb{R}^n$ *and let* $f : \mathbb{R}^n \to \mathbb{R}$ *be* usc *on* $\overline{U}$. *Then*

$$\inf f(U) = \inf f(\overline{U}).$$

Proof. As $U \subset \overline{U}$, we have

$$\inf f(\overline{U}) \leq \inf f(U).$$

As $U \neq \varnothing$, both $\inf f(\overline{U})$ and $\inf f(U)$ are bounded above.

If $\inf f(U) = -\infty$, then $\inf f(\overline{U}) = -\infty$ and the equality is verified. If $\inf f(U)$ is finite, assume that

$$\inf f(\overline{U}) < \inf f(U).$$

By definition of $\inf f(\overline{U})$, there exists $x_0 \in \overline{U}$ such that

$$\inf f(\overline{U}) \le f(x_0) < \inf f(U).$$

As f is usc, there exists a neighborhood $V(x_0)$ of x_0 such that

$$\forall x \in V(x_0), \quad f(x) < \inf f(U).$$

But $x_0 \in \overline{U}$ is a limit point of U for which $V(x_0) \cap U \neq \varnothing$. Therefore, there exists $u \in U$ such that $f(u) < \inf f(U)$. This contradicts the definition of $\inf f(U)$. □

5.2 U Closed but not Necessarily Bounded

By Theorem 5.1, the functions of Figure 2.1 have minimizers at least at one point of the compact subset $U = [0,1]$ of $\mathbb{R}$. However, in its present form, this theorem is a little restrictive since it does not apply to the following simple example:

$$\inf f(\mathbb{R}^n), \quad f(x) \stackrel{\text{def}}{=} \|x\|^2, \quad x \in \mathbb{R}^n.$$

The difficulty arises from the fact that, as $\mathbb{R}^n$ is not bounded, it is not compact.

Remark 5.1.
Going over the proof of the theorem, it is readily seen that the compactness of U is not really necessary. Since the family $\{F_k : k > m\}$ of closed subsets of U is an "increasing sequence"

$$F_{k_1} \subset F_{k_2} \subset U, \quad \forall k_2 \ge k_1 > m,$$

then, for each $\bar{k} > m$, $\cap_{\bar{k} \ge k > m} F_k = \cap_{k>m} F_k$. So it is sufficient to find some $\bar{k} \in \mathbb{R}$ for which the *lower section* $F_{\bar{k}} = \{x \in U : f(x) \le \bar{k}\}$ is nonempty and bounded (hence compact[5]) instead of making the assumption on U. □

Definition 5.1.
Let U, $\varnothing \neq U \subset \mathbb{R}^n$.

(i) $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ has a *bounded lower section* in U if there exists $k \in \mathbb{R}$ such that the lower section

$$F_k = \{x \in U : f(x) \le k\} \tag{5.3}$$

is nonempty and bounded.

(ii) $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ has a *bounded upper section* in U if there exists $k \in \mathbb{R}$ such that the upper section

$$F_k = \{x \in U : f(x) \ge k\} \tag{5.4}$$

is nonempty and bounded. □

[5] In finite dimension a set is compact if and only if it is closed and bounded by Heine–Borel theorem (Theorem 4.2 of Chapter 1).

When U is a nonempty compact subset of $\mathbb{R}^n$ (that is, bounded and closed), any function f proper for the infimum has a bounded lower section in U.

Theorem 5.3. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be closed.*

(i) *If $f : U \to \mathbb{R}\cup\{+\infty\}$ is* lsc *on U with a* bounded lower section in U, *then*

$$\exists a \in U \text{ such that } f(a) = \inf_{x\in U} f(x) \in \mathbb{R}. \tag{5.5}$$

(ii) *If $f : U \to \mathbb{R}\cup\{-\infty\}$ is* usc *on U with a* bounded upper section in U, *then*

$$\exists b \in U \text{ such that } f(b) = \sup_{x\in U} f(x) \in \mathbb{R}. \tag{5.6}$$

Example 5.2 (distance function)**.**
Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be closed. Given $x \in \mathbb{R}^n$, there exists $\hat{x} \in U$ such that

$$d_U(x) \overset{\text{def}}{=} \inf_{y\in U} \|x-y\| = \|x-\hat{x}\| \tag{5.7}$$

and

$$\forall x_1, x_2 \in \mathbb{R}^n, \quad |d_U(x_2) - d_U(x_1)| \le \|x_2 - x_1\|. \tag{5.8}$$

To show this, consider the infimum

$$\inf_{y\in U} f(y), \quad f(y) \overset{\text{def}}{=} \|y-x\|^2.$$

The function f is continuous and hence lsc on $\mathbb{R}^n$. For any $y_0 \in U$ and $k = \|y_0 - x\|^2$, the lower section

$$F_k \overset{\text{def}}{=} \left\{ y \in U : \|y-x\|^2 \le \|y_0 - x\|^2 \right\}$$

is not empty, since $y_0 \in F_k$, and bounded since

$$\forall y \in F_k, \quad \|y\| \le \|x\| + \|y-x\| \le \|x\| + \|y_0 - x\| \le \|x\| + \sqrt{k}.$$

f has a bounded lower section in U. By Theorem 5.3(i), there exists a minimizer $\hat{x} \in U$. For all $y \in U$,

$$\begin{aligned} \|x_2 - y\| \le \|x_1 - y\| + \|x_2 - x_1\| \quad &\Rightarrow \inf_{y\in U} \|x_2 - y\| \le \inf_{y\in U} \|x_1 - y\| + \|x_2 - x_1\| \\ &\Rightarrow \forall x_1, x_2 \in \mathbb{R}^n, \quad d_U(x_2) - d_U(x_1) \le \|x_2 - x_1\|. \end{aligned}$$

By interchanging the roles of x_1 and x_2, we get $|d_U(x_2) - d_U(x_1)| \le \|x_2 - x_1\|$. □

Example 5.3 (distance function)**.**
Let U, $\varnothing \neq U \subset \mathbb{R}^n$ (not necessarily closed) and $x \in \mathbb{R}^n$. As in the previous example, define

$$d_U(x) \overset{\text{def}}{=} \inf_{y\in U} \|x-y\|. \tag{5.9}$$

As $U \subset \overline{U}$, then $d_U(x) \ge d_{\overline{U}}(x)$. However, since the function $y \mapsto \|y-x\|$ is continuous, it is usc and by Theorem 5.2, $d_U(x) = d_{\overline{U}}(x)$. □

5.3 Growth Property at Infinity

A simple condition to ensure that a function has a bounded lower section in an unbounded U is that the function goes to $+\infty$ as the norm of $\|x\|$ tends to infinity.

Definition 5.2.
Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be unbounded in $\mathbb{R}^n$. The function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ has the *growth property* in U if

$$\lim_{x \in U,\, \|x\| \to \infty} f(x) = +\infty. \qquad \square$$

Theorem 5.4. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be unbounded. If $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ has the growth property in U, then it has a nonempty lower section in U.*

Proof. We show that there exists a $k \in \mathbb{R}$ such that the lower section $F_k = \{x \in U : f(x) \leq k\}$ is nonempty and bounded. By definition, $U = \cup_{k \in \mathbb{R}} F_k$ and as $U \neq \varnothing$, $\exists k \in \mathbb{R}$ such that $F_k \neq \varnothing$. By the growth property in U,

$$\exists R(k) > 0 \text{ such that } \forall x \in U \text{ and } \|x\| > R(k), \quad f(x) > k.$$

As a result,

$$F_k = \{x \in U : f(x) \leq k\} \subset \{x \in U : \|x\| \leq R(k)\}$$

and F_k is nonempty and bounded. $\square$

Consider a few generic examples.

Example 5.4.
The functions $f(x) = |x|$ and $f(x) = x^2$ have the growth property in $\mathbb{R}$. $\square$

Example 5.5.
The function $f(x) = x - b$ does not have the growth property in $\mathbb{R}$. Pick a sequence $x_n = -n$ of positive integers n going to infinity. $\square$

Example 5.6.
The function $f(x) = \sin x + (1 + x)^2$ has the growth property in $\mathbb{R}$. Indeed

$$f(x) \geq -1 + (1 + x)^2 = x^2 - 2x \to +\infty$$

as $|x| \to \infty$. $\square$

Example 5.7.
Consider the function

$$f(x_1, x_2) = (x_1 + x_2)^2.$$

f does not have the growth property in $\mathbb{R}^2$: pick the sequence $\{(n, -n)\}$, $n \geq 1$,

$$f(n, -n) = (n - n)^2 = 0 \not\to +\infty.$$

However, f has the growth property in

$$U = \{(x_1, 0) : x_1 \in \mathbb{R}\}$$

since

$$f(x) = x_1^2 \to +\infty$$

as $|x_1|$ goes to $+\infty$ in U. □

Theorem 5.5. *Given $a \in \mathbb{R}^n$, the functions $x \mapsto f(x) = \|x-a\|^p$, $p \geq 1$, have the growth property in $\mathbb{R}^n$.*

Proof. For all $x \neq 0$,

$$f(x) = \|x\|^p \left\| \frac{x}{\|x\|} - \frac{a}{\|x\|} \right\|^p.$$

As $\|x\| \to \infty$, the term

$$\left\| \frac{x}{\|x\ |} - \frac{a}{\|x\|} \right\|$$

converges to 1 and its pth power also converges to 1. For $p \geq 1$, $\|x\|^p \to +\infty$ as $\|x\|$ goes to $+\infty$. The limit of $f(x)$ is the product of the two limits. □

In the next example we use the following technical results. Recall that an $n \times n$ matrix is *symmetric* if $A^\top = A$, where $A^\top$ denotes the *matrix transpose*[6] of a matrix A.

Definition 5.3.
A symmetric matrix A is *positive definite* (resp., *positive semidefinite*) if

$$\forall x \in \mathbb{R}^n, x \neq 0, \quad (Ax)\cdot x > 0 \quad (\text{resp., } \forall x \in \mathbb{R}^n, \quad (Ax)\cdot x \geq 0).$$

This property will be denoted $A > 0$ (resp., $A \geq 0$). □

Lemma 5.1. *A symmetric matrix A is* positive definite *if and only if*

$$\exists \alpha > 0, \forall x \in \mathbb{R}^n, \ (Ax)\cdot x \geq \alpha \|x\|^2.$$

If $A > 0$, the inverse A^{-1} exists.

Proof. ($\Leftarrow$) If there exists $\alpha > 0$ such that $(Ax)\cdot x \geq \alpha\|x\|^2$ for all $x \in \mathbb{R}^n$, then

$$\forall x \in \mathbb{R}^n, \ x \neq 0, \quad (Ax)\cdot x \geq \alpha \|x\|^2 > 0$$

and $A > 0$.

($\Rightarrow$) Conversely, if $A > 0$, then

$$\forall x \in \mathbb{R}^n, \quad (Ax)\cdot x \geq 0.$$

Assume that there exists no $\alpha > 0$ such that

$$\forall x \in \mathbb{R}^n, \quad (Ax)\cdot x \geq \alpha \|x\|^2.$$

So for each integer $k > 0$, there exists x_k such that

$$0 \leq Ax_k \cdot x_k < \frac{1}{k}\|x_k\|^2 \quad \Rightarrow x_k \neq 0.$$

[6]The notation A will be used for both the linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ and its associated $m \times n$ matrix. Similarly, the notation $A^\top$ will be used for both the linear map $A^\top : \mathbb{R}^m \to \mathbb{R}^n$ and its associated $n \times m$ matrix.

By dividing by $\|x_k\|^2$ we get

$$0 \le A\frac{x_k}{\|x_k\|}\cdot\frac{x_k}{\|x_k\|} < \frac{1}{k} \quad \Rightarrow \lim_{k\to\infty} A\frac{x_k}{\|x_k\|}\cdot\frac{x_k}{\|x_k\|} = 0.$$

The points $s_k = x_k/\|x_k\|$ belong to the sphere $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ which is compact in $\mathbb{R}^n$. By the theorem of Bolzano–Weierstrass (see Chapter 1, Theorem 4.4), there exists a subsequence $\{s_{k_\ell}\}$ of $\{s_k\}$ that converges to a point s of S:

$$\exists s,\ \ \|s\| = 1 \ \text{ such that } \ 0 = \lim_{\ell\to\infty} As_{k_\ell}\cdot s_{k_\ell} = As\cdot s.$$

Therefore,

$$\exists s \neq 0 \ \text{ such that } \quad As\cdot s = 0 \implies A \not> 0.$$

This contradicts the assumption that $A > 0$ and proves the result. To show that A is invertible, it is sufficient to check that under the assumption $A > 0$, the linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ is injective, that is, $Ax = 0$ implies $x = 0$. From the previous result, it is readily seen that

$$Ax = 0 \quad \Rightarrow 0 = Ax\cdot x \ge \alpha\,\|x\|^2 \quad \Rightarrow x = 0. \qquad \square$$

Example 5.8.
Let $A > 0$ be an $n\times n$ symmetric matrix. Associate with A and $b \in \mathbb{R}^n$ the following real-valued function:[7]

$$f(x) = \frac{1}{2}(Ax)\cdot x + b\cdot x,\ \ x \in \mathbb{R}^n.$$

By Lemma 5.1 applied to A, there exists $\alpha > 0$ such that

$$\forall x \in \mathbb{R}^n, \quad Ax\cdot x \ge \alpha\|x\|^2$$

and

$$f(x) = \frac{1}{2}(Ax)\cdot x + b\cdot x \qquad \Longrightarrow \quad f(x) \ge \frac{1}{2}\alpha\|x\|^2 - \|b\|\,\|x\|.$$

The function f has the growth property in $\mathbb{R}^n$. Since f is polynomial, it is continuous and as $U = \mathbb{R}^n$ is closed, there exists a minimizer. $\square$

5.4 Some Properties of the Set of Minimizers

Recall that, in the proof of Theorem 5.1 for the infimum of a function f with respect to U, the set of minimizers $\operatorname{argmin} f(U)$ is given by

$$\operatorname{argmin} f(U) = \bigcap_{k>m} F_k, \ \text{ where } m = \inf f(U) \text{ and } F_k = \{x \in U : f(x) \le k\}.$$

So a number of properties of $\operatorname{argmin} f(U)$ can be obtained from those of the F_k's.

[7] If a vector $v \in \mathbb{R}^n$ is considered as an $n\times 1$ matrix, the product $(Ax)\cdot x$ can be rewritten as the product of three matrices $x^\top Ax$ and $b\cdot x$ can be written as the product of two matrices $b^\top x$.

Theorem 5.6. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be closed and let $f : U \to \mathbb{R} \cup \{+\infty\}$ be* lsc *on U. Then* $\operatorname{argmin} f(U)$ *is closed (possibly empty). If, in addition, f has a bounded lower section in U, then* $\operatorname{argmin} f(U)$ *is compact and nonempty.*

Proof. Since U is closed and f is lsc, the lower sections $\{F_k\}$ in (5.3) are closed and $\operatorname{argmin} f(U)$ is closed as an intersection of closed sets. If, in addition, f has a bounded lower section (there exists k_0 such that F_{k_0} is nonempty and bounded), then the closed set F_{k_0} is compact. Therefore, the set of minimizers

$$\operatorname{argmin} f(U) = \bigcap_{k>m} F_k \subset F_{k_0}$$

is compact as a closed subset of the compact set F_{k_0}. □

Example 5.9.
Go back to Examples 5.2 and 5.3 for U, $\varnothing \neq U \subset \mathbb{R}^n$. Given $x \in \mathbb{R}^n$, it was shown that $d_U(x) = d_{\overline{U}}(x)$ and that

$$\exists \hat{x} \in \overline{U} \text{ such that } d_{\overline{U}}(x) \stackrel{\text{def}}{=} \inf_{y \in \overline{U}} \|x - y\| = \|x - \hat{x}\|. \tag{5.10}$$

To establish that result we have considered the infimum of the square of the distance

$$\inf_{y \in U} f(y), \quad f(y) \stackrel{\text{def}}{=} \|y - x\|^2$$

for the continuous function f which has a bounded lower section in $\overline{U}$. By Theorem 5.6, $\operatorname{argmin} f(\overline{U})$ is nonempty and compact. Points of $\operatorname{argmin} f(\overline{U})$ are *projections* of x onto $\overline{U}$. Denote by $\Pi_U(x)$ this set. If $x \in \overline{U}$, then $\Pi_U(x) = \{x\}$ is a singleton. □

6 ► Ekeland's Variational Principle

Ekeland's variational principle[8] in 1974 (I. Ekeland [**1**]) is a basic tool to get existence of approximate minimizers in the absence of compactness. Its major impact is in the context of spaces of functions. In this section we provide the finite-dimensional version and some of its ramifications.

Theorem 6.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$,* $\operatorname{dom} f \neq \varnothing$*, be lsc and bounded below. Then, for any $\varepsilon > 0$, there exists x_ε such that*

$$\inf_{x \in \mathbb{R}^n} f(x) \leq f(x_\varepsilon) < \inf_{x \in \mathbb{R}^n} f(x) + \varepsilon \tag{6.1}$$

and, for any $\eta > 0$, there exists y such that

$$\|y - x_\varepsilon\| < \eta, \quad f(y) \leq f(x_\varepsilon) - \frac{\varepsilon}{\eta} \|y - x_\varepsilon\|, \tag{6.2}$$

and

$$\forall x \in \mathbb{R}^n, x \neq y, \quad f(y) < f(x) + \frac{\varepsilon}{\eta} \|y - x\|. \tag{6.3}$$

[8] Ivar Ekeland (1944–).

This implies that, if $f(x_\varepsilon) > \inf f(\mathbb{R}^n)$, then, in any neighborhood of x_ε, there exists a point y such that $f(y) < f(x_\varepsilon)$.

Proof. [9] Under the assumption of the theorem, the infimum is finite and, for any $\varepsilon > 0$, there exists $x_\varepsilon \in \operatorname{dom} f$ such that

$$\inf_{x\in\mathbb{R}^n} f(x) \le f(x_\varepsilon) < \inf_{x\in\mathbb{R}^n} f(x) + \varepsilon$$

so that $f(x_\varepsilon)$ is also finite. Given $\eta > 0$, consider the new lsc function

$$g_{\varepsilon/\eta}(y) \stackrel{\text{def}}{=} f(y) + \frac{\varepsilon}{\eta}\,\|y - x_\varepsilon\|.$$

The lower section

$$S \stackrel{\text{def}}{=} \{z \in \mathbb{R}^n : g_{\varepsilon/\eta}(z) \le f(x_\varepsilon\}$$

is not empty, since $x_\varepsilon \in S$, and bounded since

$$\begin{aligned} &f(z) + \frac{\varepsilon}{\eta}\,\|z - x_\varepsilon\| \le f(x_\varepsilon) \\ \Rightarrow\ &\inf_{x\in\mathbb{R}^n} f(x) + \frac{\varepsilon}{\eta}\,\|z - x_\varepsilon\| \le f(z) + \frac{\varepsilon}{\eta}\,\|z - x_\varepsilon\| \le f(x_\varepsilon) \\ \Rightarrow\ &\frac{\varepsilon}{\eta}\,\|z - x_\varepsilon\| \le f(x_\varepsilon) - \inf_{x\in\mathbb{R}^n} f(x) \end{aligned}$$

and $f(x_\varepsilon) - \inf_{x\in\mathbb{R}^n} f(x) \ge 0$ is finite. By Theorem 5.3(i), the set of minimizers $\operatorname{argmin} g_{\varepsilon/\eta}$ is not empty and by Theorem 5.6, it is compact. Since f is lsc, by Theorem 5.1, there exists $y \in \operatorname{argmin} g_{\varepsilon/\eta}$ such that

$$\begin{aligned} f(y) &= \inf_{z\in\operatorname{argmin} g_{\varepsilon/\eta}} f(z) \\ \forall x \in \mathbb{R}^n,\quad f(y) + \frac{\varepsilon}{\eta}\,\|y - x_\varepsilon\| &\le f(x) + \frac{\varepsilon}{\eta}\,\|x - x_\varepsilon\|. \end{aligned}$$

For $x = x_\varepsilon$,

$$f(y) + \frac{\varepsilon}{\eta}\,\|y - x_\varepsilon\| \le f(x_\varepsilon) \quad \Rightarrow f(y) \le f(x_\varepsilon) - \frac{\varepsilon}{\eta}\,\|y - x_\varepsilon\|$$

and, from the definition of x_ε,

$$f(y) + \frac{\varepsilon}{\eta}\,\|y - x_\varepsilon\| \le f(x_\varepsilon) < \inf_{x\in\mathbb{R}^n} f(x) + \varepsilon \le f(y) + \varepsilon \quad \Rightarrow \|y - x_\varepsilon\| < \eta.$$

Consider the new function

$$g(z) \stackrel{\text{def}}{=} f(z) + \frac{\varepsilon}{\eta}\,\|z - y\|. \tag{6.4}$$

[9]From the proof of J. M. Borwein and A. S. Lewis [**1**, Thm. 7.1.2, pp. 153–154].

By definition, $\inf_{z\in\mathbb{R}^n} g(z) \le f(y)$ and

$$\forall z \in \mathbb{R}^n, \quad f(z)+\frac{\varepsilon}{\eta}\|z-x_\varepsilon\| \le f(z)+\frac{\varepsilon}{\eta}\|z-y\|+\frac{\varepsilon}{\eta}\|y-x_\varepsilon\|$$

$$\Rightarrow f(y)++\frac{\varepsilon}{\eta}\|y-x_\varepsilon\| = \inf_{z\in\mathbb{R}^n}\left[f(z)+\frac{\varepsilon}{\eta}\|z-x_\varepsilon\|\right]$$

$$\le \inf_{z\in\mathbb{R}^n}\left[f(z)+\frac{\varepsilon}{\eta}\|z-y\|\right]+\frac{\varepsilon}{\eta}\|y-x_\varepsilon\|$$

$$\Rightarrow f(y) \le \inf_{z\in\mathbb{R}^n}\left[f(z)+\frac{\varepsilon}{\eta}\|z-y\|\right] \le f(y) \quad \Rightarrow f(y) = \inf_{z\in\mathbb{R}^n}\left[f(z)+\frac{\varepsilon}{\eta}\|z-y\|\right].$$

To complete the proof, it remains to show that y is the unique minimizer of g. By construction, y is a global minimizer of f on $\operatorname{argmin} g_{\varepsilon/\eta}$. So, for all $y' \in \operatorname{argmin} g_{\varepsilon/\eta}$, $y' \neq y$,

$$f(y) \le f(y') < f(y')+\frac{\varepsilon}{\eta}\|y'-y\|$$

and for $y' \notin \operatorname{argmin} g_{\varepsilon/\eta}$,

$$f(y)+\frac{\varepsilon}{\eta}\|y-x_\varepsilon\| < f(y')+\frac{\varepsilon}{\eta}\|y'-x_\varepsilon\| \le f(y')+\frac{\varepsilon}{\eta}\|y'-y\|+\frac{\varepsilon}{\eta}\|y-x_\varepsilon\|$$

$$\Rightarrow \forall y' \notin \operatorname{argmin} g_{\varepsilon/\eta}, \quad f(y) < f(y')+\frac{\varepsilon}{\eta}\|y'-y\|.$$

Combining the two inequalities,

$$\forall y' \in \mathbb{R}^n, y' \neq y, \quad f(y) < f(y')+\frac{\varepsilon}{\eta}\|y'-y\|$$

and y is the unique minimizer of g. □

Ekeland's variational principle has many interesting ramifications.

Theorem 6.2 (W. TAKAHASHI [**1**]'s existence theorem)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, $\operatorname{dom} f \neq \varnothing$*, be lsc and bounded below. Assume that there exists* $c > 0$ *such that for each* x *for which* $f(x) > \inf f(\mathbb{R}^n)$*,*

$$\exists z \neq x \text{ such that } f(z) \le f(x) - c\|z-x\|. \tag{6.5}$$

Then f *has a minimizer in* $\mathbb{R}^n$*.*

Proof. It is sufficient to prove the theorem for the function $f(x)/c$ that amounts to set $c = 1$. By contradiction, assume that for all x, $f(x) > \inf f(\mathbb{R}^n)$. By the hypothesis of the theorem,

$$\exists z_x \neq x \text{ such that } f(z_x) < f(x) - \|z_x - x\|.$$

But, by Ekeland's variational principle, there exists y such that

$$\forall x \in \mathbb{R}^n, x \neq y, \quad f(y) < f(x)+\|y-x\|$$

(choose $\varepsilon = 1/n$ and $\eta = 1$ in the theorem). Since $z_y \neq y$,

$$f(y) < f(z_y)+\|y-z_y\| \text{ and } f(z_y) \le f(y)-\|z_y-y\|,$$

which yields a contradiction. □

The above two theorems are intimately related to each other and also related to the following fixed point theorem.

Theorem 6.3 (J. CARISTI [**1**]'s fixed point theorem).[10] *Let $F : \mathbb{R}^n \to \mathbb{R}^n$. If there exists an lsc function $f : \mathbb{R}^n \to \mathbb{R}$ that is bounded below such that*

$$\forall x \in \mathbb{R}^n, \quad \|x - F(x)\| \le f(x) - f(F(x)), \tag{6.6}$$

then F has a fixed point.

Proof. By Theorem 6.1 with $\varepsilon/\eta = 1$, there exists y such that

$$\forall x \in \mathbb{R}^n \, x \neq y, \quad f(y) < f(x) + \|x - y\|.$$

If $F(y) \neq y$, then

$$f(y) < f(F(y)) + \|F(y) - y\| \quad \Rightarrow f(y) - f(F(y)) < \|F(y) - y\|.$$

But, by assumption, $f(y) - f(F(y)) \ge \|y - F(y)\|$, which yields a contradiction. Hence $F(y) = y$ and y is a fixed point of F. □

Example 6.1.
For a contracting mapping F,

$$\exists k, 0 < k < 1, \text{ such that } \forall x, y \in \mathbb{R}^n, \quad \|F(y) - F(x)\| \le k\,\|y - x\|,$$

choose the function

$$f(x) \stackrel{\text{def}}{=} \|x - F(x)\|/(1-k)$$

which is continuous and bounded below by 0. Then

$$\begin{aligned}
&\forall x \in \mathbb{R}^n, \quad \|x - F(x)\| = (1-k)\,f(x) = f(x) - k\,f(x)\\
&\forall x \in \mathbb{R}^n, \quad f(F(x)) = \frac{\|F(x) - F(F(x))\|}{1-k} \le \frac{k}{1-k}\|x - F(x)\| = k\,f(x)\\
&\Rightarrow \forall x \in \mathbb{R}^n, \quad \|x - F(x)\| = (1-k)\,f(x) = f(x) - k\,f(x) \le f(x) - f(F(x)).
\end{aligned}$$

The assumptions of Caristi's theorem are verified. This is the Banach fixed point theorem. □

7 Convexity, Quasiconvexity, Strict Convexity, and Uniqueness

7.1 Convexity and Concavity

Different notions of convexity will be introduced to discuss the convexity of $\operatorname{argmin} f(U)$ and the uniqueness of minimizers. For the supremum, the dual notion is concavity. As for

[10] Caristi's fixed point theorem (also known as the Caristi–Kirk fixed point theorem) generalizes the Banach fixed point theorem for maps of a complete metric space into itself. Caristi's fixed point theorem is a variation of the variational principle of Ekeland. Moreover, the conclusion of Caristi's theorem is equivalent to metric completeness, as proved by J. D. WESTON [**1**] (1977). The original result is due to the mathematicians James Caristi and William Arthur Kirk (see J. CARISTI and W. A. KIRK [**1**]).

lsc functions, it is advantageous to extend the notion of convexity to functions with values in $\mathbb{R}\cup\{+\infty\}$.

Definition 7.1. (i) A subset U of $\mathbb{R}^n$ is said to be *convex* if

$$\forall\lambda \in [0,1],\ \forall x,y \in U,\ \lambda x+(1-\lambda)y \in U.$$

By convention, $\varnothing$ is convex.

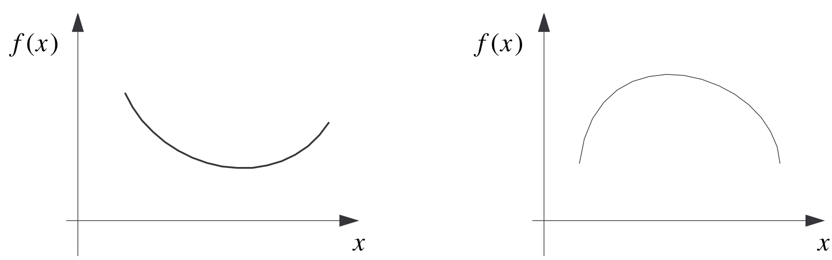
(ii) Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be convex. The function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is said to be *convex* on U if

$$\forall\lambda \in]0,1[,\quad \forall x,y \in U,\quad f(\lambda x+(1-\lambda)y) \le \lambda f(x)+(1-\lambda)f(y),$$

with the convention $(+\infty)+(+\infty) = +\infty$, $(+\infty)+a = +\infty$ for all $a \in \mathbb{R}$, and $(+\infty)a = (a/\|a\|)\infty$ for all $a \in \mathbb{R}$ not equal to 0.

(iii) Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be convex. The function $f : \mathbb{R}^n \to \mathbb{R}\cup\{-\infty\}$ is said to be *concave* on U if $-f$ is convex on U (see Figure 2.4).

(iv) By convention, the function identically equal to $-\infty$ or $+\infty$ is both convex and concave. □



Convex function Concave function

Figure 2.4. *Convex function and concave function.*

The indicator function I_U and the distance function d_U for which $(d_U)_U = I_U$ are, respectively, related to the convexity of U and its closure.

Theorem 7.1. *Let $U \subset \mathbb{R}^n$.*

(i) *The interior,* $\operatorname{int} U$, *and the closure,* $\overline{U}$, *of a convex set U are convex.*

(ii) *The set U is convex if and only if I_U is convex.*

(iii) *The set $\overline{U}$ is convex if and only if d_U is convex.*

(iv) *Let $\{U_\alpha\}_{\alpha\in A}$ be a family of convex subsets of $\mathbb{R}^n$, where the set A of indices is arbitrary and not necessarily finite. Then $U = \bigcap_{\alpha\in A} U_\alpha$ is convex.*

(v) *Let U be convex such that* $\operatorname{int} U \neq \varnothing$. *Then* $\overline{\operatorname{int} U} = \overline{U}$.

Proof. (i) If $U = \varnothing$, then $\overline{\varnothing} = \varnothing$ is convex by convention. If $x, y \in \overline{U}$, then there exist sequences $\{x_n\} \subset U$ and $\{y_n\} \subset U$ such that $x_n \to x$ and $y_n \to y$. By convexity of U, for all $\lambda \in [0,1]$,

$$U \ni \lambda x_n + (1-\lambda) y_n \to \lambda x + (1-\lambda) y.$$

Thus $\lambda x + (1-\lambda) y \in \overline{U}$ and $\overline{U}$ is convex.

If $\operatorname{int} U = \varnothing$, then it is convex by convention. For $x, y \in \operatorname{int} U$, there exists $B_\varepsilon(x)$ and $B_\eta(y)$ such that $B_\varepsilon(x) \subset U$ and $B_\eta(y) \subset U$ and, for all $\lambda \in [0,1]$,

$$\lambda x + (1-\lambda) y \in \lambda B_\varepsilon(x) + (1-\lambda) B_\eta(y).$$

But,

$$\lambda x + (1-\lambda) y \in B_{\min\{\varepsilon,\eta\}}(\lambda x + (1-\lambda) y) \subset \lambda B_\varepsilon(x) + (1-\lambda) B_\eta(y) \subset U$$

and $\lambda x + (1-\lambda) y \in \operatorname{int} U$.

(ii) If $U = \varnothing$, then I_U is identically $+\infty$ which is convex by convention. If $U \neq \varnothing$ is convex, then for all $x, y \in U$ and $\lambda \in]0,1[$, $\lambda x + (1-\lambda) y \in U$ and

$$I_U(\lambda x + (1-\lambda) y) = 0 = \lambda\, I_U(x) + (1-\lambda)\, I_U(y) = 0.$$

If either x or y is not in U, then either $I_U(x)$ or $I_U(y)$ is $+\infty$ and $\lambda\, I_U(x) + (1-\lambda)\, I_U(y) = +\infty$, so that

$$I_U(\lambda x + (1-\lambda) y) \le +\infty = \lambda\, I_U(x) + (1-\lambda)\, I_U(y).$$

Hence I_U is convex.

Conversely, if I_U is identically $+\infty$, then $U = \varnothing$ which is convex by convention. Otherwise, for all $x, y \in U$ and $\lambda \in [0,1]$,

$$0 \le I_U(\lambda x + (1-\lambda) y) \le \lambda\, I_U(x) + (1-\lambda)\, I_U(y) = 0$$

and $\lambda x + (1-\lambda) y \in U$. Hence the convexity of U.

(iii) By convention, $U = \varnothing$ implies $d_U(x) = -\infty$ that is convex, also by convention. For $U \neq \varnothing$. Given x and y in $\mathbb{R}^n$, there exist $\overline{x}$ and $\overline{y}$ in $\overline{U}$ such that $d_U(x) = |x - \overline{x}|$ and $d_U(y) = |y - \overline{y}|$. By convexity of $\overline{U}$, for all λ, $0 \le \lambda \le 1$, $\lambda \overline{x} + (1-\lambda) \overline{y} \in \overline{U}$ and

$$\begin{aligned} d_U(\lambda x + (1-\lambda) y) \le &\|\lambda x + (1-\lambda) y - (\lambda \overline{x} + (1-\lambda) \overline{y})\| \\ \le &\lambda \|x - \overline{x}\| + (1-\lambda) \|y - \overline{y}\| = \lambda\, d_U(x) + (1-\lambda)\, d_U(y) \end{aligned}$$

and d_U is convex in $\mathbb{R}^n$.

Conversely, if $d_U(x) = -\infty$ for some x, then, by convention, $U = \varnothing$ which is convex, also by convention. If d_U is finite and d_U is convex, then

$$\forall \lambda \in [0,1], \forall x, y \in \overline{U}, \quad d_U(\lambda x + (1-\lambda) y) \le \lambda d_U(x) + (1-\lambda) d_U(y).$$

But x and y in $\overline{U}$ imply that $d_U(x) = d_U(y) = 0$ and hence

$$\forall \lambda \in [0,1], \quad d_U(\lambda x + (1-\lambda)y) = 0.$$

Thus $\lambda x + (1-\lambda)y \in \overline{U}$ and $\overline{U}$ is convex.

(iv) If U is empty, U is convex by definition. If U is not empty, choose

$$x \text{ and } y \in U = \cap_{\alpha \in A} U_\alpha \quad \Rightarrow \forall \alpha \in A, \quad x \in U_\alpha, y \in U_\alpha.$$

For all $\lambda \in [0,1]$ and by convexity of U_α,

$$\forall \alpha \in A, \quad \lambda x + (1-\lambda)y \in U_\alpha \quad \Rightarrow \lambda x + (1-\lambda)y \in \cap_{\alpha \in A} U_\alpha = U.$$

(v) As $\operatorname{int} U \neq \varnothing$, pick a point $x \in \operatorname{int} U$. By convexity, for all $y \in \partial U$, the segment $[x,y] = \{\lambda x + (1-\lambda)y : 0 \leq \lambda \leq 1\}$ belongs to $\overline{U}$ and $[x,y[= \{\lambda x + (1-\lambda)y : 0 < \lambda \leq 1\} \subset \operatorname{int} U$. So there exists a sequence $y_n = x + (y-x)/(n+1)$ in $\operatorname{int} U$ that converges to y. Hence the result. □

The following definitions will also be useful.

Definition 7.2.
Let U, $\varnothing \neq U \subset \mathbb{R}^n$.

(i) The *convex hull* of U is the intersection of all convex subsets of $\mathbb{R}^n$ that contain U. It is denoted $\operatorname{co} U$.

(ii) The *closed convex hull* of U is the intersection of all closed convex subsets of $\mathbb{R}^n$ that contains U. It is denoted $\overline{\operatorname{co}}\, U$. □

Theorem 7.2. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$.*

(i) $\operatorname{co} U$ *is convex and*

$$\operatorname{co} U = \left\{ \sum_{i=1}^{k} \lambda_i x_i : \sum_{i=1}^{k} \lambda_i = 1, x_i \in U, 0 \leq \lambda_i \leq 1, k \geq 1 \right\}. \tag{7.1}$$

(ii) $\overline{\operatorname{co}}\, U$ *is closed and convex and*

$$\overline{\operatorname{co}}\, U = \overline{\operatorname{co} U} = \operatorname{co} \overline{U}. \tag{7.2}$$

Proof. (i) By Theorem 7.1(iv), $\operatorname{co} U$ is convex as an intersection of convex sets. Denote by C the right-hand side of (7.1). We have $U \subset \operatorname{co} U \subset C$ since C is a convex that contains U. Taking all convex combinations of elements of U, we get C that is entirely contained in $\operatorname{co} U$ and $\operatorname{co} U = C$.

(ii) By Theorem 7.1(iv) and the fact that the intersection of a family of closed sets is closed. □

The convexity of f on a convex U can also be characterized in terms of the convexity of f_U or of its epigraph epi f_U as in the case when f is lsc (see Lemma 4.1).

Lemma 7.1. (i) *If* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex on* $\mathbb{R}^n$*, then* dom f *is convex and* f *is convex on* dom f.

(ii) *If* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex on a convex subset* U *of* $\mathbb{R}^n$*, then* f_U *is convex on* $\mathbb{R}^n$, dom $f_U = U \cap$ dom f *is convex, and* f_U *is convex on* dom f_U.

Proof. (i) For any convex combination of $x, y \in \text{dom}\, f$, $f(x) < +\infty$ and $f(y) < +\infty$, $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) < +\infty$, and $\lambda x + (1-\lambda)y \in \text{dom}\, f$. Thence, dom f is convex and, by definition, $f : \text{dom}\, f \to \mathbb{R}$ is convex.

(ii) By definition, dom $f_U = U \cap \text{dom}\, f$ and as $f_U(x) = +\infty$ outside of U, f_U is convex on $\mathbb{R}^n$. Finally, from part (i), dom f_U is convex. □

Theorem 7.3. *Let* U, $\varnothing \ne U \subset \mathbb{R}^n$, *be convex and* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. *The following conditions are equivalent:*

(i) f *is convex on* U*;*

(ii) f_U *is convex on* $\mathbb{R}^n$*;*

(iii) epi f_U *is convex on* $\mathbb{R}^n \times \mathbb{R}$.

Proof. (i) $\Rightarrow$ (ii) From Lemma 7.1(ii).

(ii) $\Rightarrow$ (iii) From Lemma 7.1(i), dom f_U is convex and $f_U : \text{dom}\, f_U \to \mathbb{R}$ is convex. For all (x, μ_x) and (y, μ_y) in epi f_U, $\mu_x \ge f(x)$ and $\mu_y \ge f(y)$ and hence $x, y \in \text{dom}\, f_U$. Given $\lambda \in [0,1]$, consider their convex combination

$$\lambda(x, \mu_x) + (1-\lambda)(y, \mu_y) = (\lambda x + (1-\lambda) y, \lambda \mu_x + (1-\lambda)\mu_y).$$

By construction and convexity of f_U, $\lambda \mu_x + (1-\lambda)\mu_y \ge \lambda f(x) + (1-\lambda) f(y) \ge f(\lambda x + (1-\lambda) y)$. Therefore, $f(\lambda x + (1-\lambda) y) < +\infty$, $\lambda x + (1-\lambda) y \in \text{dom}\, f_U \subset U$, and $\lambda(x, \mu_x) + (1-\lambda)(y, \mu_y) \in \text{epi}\, f_U$. So epi f_U is convex on $\mathbb{R}^n \times \mathbb{R}$.

(iii) $\Rightarrow$ (i) By definition of f_U, dom $f_U \subset U$ and $f_U = f$ in dom f_U. Therefore, for all $x, y \in \text{dom}\, f_U$, the pairs $(x, f(x))$ and $(y, f(y))$ belong to epi f_U. As epi f_U is convex,

$$\begin{aligned} &\forall \lambda \in [0,1], \quad \lambda(x, f(x)) + (1-\lambda)(y, f(y)) \in \text{epi}\, f_U \\ &\Rightarrow (\lambda x + (1-\lambda) y, \lambda f(x) + (1-\lambda) f(y)) \in \text{epi}\, f_U \\ &\Rightarrow \lambda f(x) + (1-\lambda) f(y) \ge f_U(x + (1-\lambda) y). \end{aligned}$$

Thence, as $f_U(x + (1-\lambda)y) < +\infty$, we get $\lambda x + (1-\lambda) y \in \text{dom}\, f_U$ and the convexity of dom f_U. Finally, $f_U(x + (1-\lambda)y) = f(x + (1-\lambda)y)$ yields $\lambda f(x) + (1-\lambda) f(y) \ge f(x + (1-\lambda)y)$ and the convexity of the function f in dom f_U. The convexity of f_U on the convex subset $U \supset \text{dom}\, f_U$ is a consequence of Lemma 7.1(i). Finally, as $f_U = f$ on U, f is convex on U. □

Except for functions whose domain is a singleton, the convexity of a function forces its liminf to be less than or equal to its value at each point of its effective domain.

Lemma 7.2. *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, dom $f \ne \varnothing$, *be a convex function.*

(i) *If* dom $f = \{x_0\}$ *is a singleton,*

$$\liminf_{y \to x} f(y) = +\infty = f(x) \quad \textit{if } x \ne x_0 \quad \textit{and} \quad \liminf_{y \to x} f(y) = +\infty > f(x_0). \tag{7.3}$$

(ii) *If* dom f *is not a singleton,*

$$\forall x \in \mathbb{R}^n, \quad \liminf_{y \in \mathrm{dom}\, f \to x} f(y) = \liminf_{y \to x} f(y) \leq f(x), \tag{7.4}$$

where the lim inf *is with respect to elements* $y \neq x$ *in* dom f *that go to* x.

Proof. If $x \notin \mathrm{dom}\, f$, $f(x) = +\infty$ and $\liminf_{y \to x} f(y) \leq +\infty = f(x)$.

(i) If $\mathrm{dom}\, f = \{x_0\}$ is a singleton, by definition of the liminf,

$$\lim_{\varepsilon \searrow 0} \inf_{y \in B'_\varepsilon(x_0)} f(y) = +\infty$$

since $f(y) = +\infty$ for $y \neq x_0$.

(ii) If dom f is not a singleton and $x \in \mathrm{dom}\, f$, there exists $y_0 \in \mathrm{dom}\, f$, $y_0 \neq x$, and, by convexity of f and dom f, for all $0 < t < 1$, $x + t(y_0 - x) \in \mathrm{dom}\, f$, $x + t(y_0 - x) \neq x$, and

$$\begin{aligned} f(x + t(y_0 - x)) \leq f(x) + t\,[f(y_0) - f(x)] \quad &\Rightarrow \liminf_{t \searrow 0} f(x + t(y_0 - x)) \leq f(x) \\ \Rightarrow \liminf_{y \to x} f(y) \leq \liminf_{t \searrow 0} f(x + t(y_0 - x)) &\leq f(x). \end{aligned}$$

Moreover,

$$\begin{aligned} \inf_{y \in B'_\varepsilon(x)} f(y) &= \inf_{y \in B'_\varepsilon(x) \cap \mathrm{dom}\, f} f(y) \\ \Rightarrow \liminf_{y \to x} f(y) = \lim_{\varepsilon \searrow 0} \inf_{y \in B'_\varepsilon(x) \cap \mathrm{dom}\, f} f(y) &= \liminf_{y \in \mathrm{dom}\, f \to x} f(y). \end{aligned}$$ □

As for lsc functions (see Theorem 4.2), convex functions remain convex under several functional operations.

Theorem 7.4. *Let* U, $\varnothing \neq U \subset \mathbb{R}^n$, *be convex.*

(i) *(Jensen's inequality*[11]*) If* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex on* U, *then*

$$\forall m \geq 1, \quad f\left(\sum_{i=1}^{m} \lambda_i\, x_i\right) \leq \sum_{i=1}^{m} \lambda_i\, f(x_i) \tag{7.5}$$

for all families of scalars $\{\lambda_i \in]0,1] : 1 \leq i \leq m, \sum_{i=1}^m \lambda_i = 1\}$ *and points* $\{x_i : 1 \leq i \leq m\}$ *in* U.

(ii) *For all* $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *convex on* U, *the function*

$$(f + g)(x) \stackrel{\text{def}}{=} f(x) + g(x), \quad \forall x \in U,$$

is convex on U.

[11] Johan Ludwig William Valdemar Jensen, mostly known as Johan Jensen (1859–1925).

(iii) *For all* $\lambda \geq 0$ *and all* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *convex on* U*, the function*

$$(\lambda f)(x) \stackrel{\text{def}}{=} \left\{ \begin{matrix} \lambda f(x), & \text{if } \lambda > 0 \\ 0, & \text{if } \lambda = 0 \end{matrix} \right\}, \qquad \forall x \in \mathbb{R}^n,$$

is convex on U.

(iv) *Given a family* $\{f_\alpha\}_{\alpha \in A}$ *(where* A *is a possibly infinite index set) of functions* $f_\alpha : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *convex on* U*, the* upper envelope

$$\left\{ \sup_{\alpha \in A} f_\alpha \right\}(x) \stackrel{\text{def}}{=} \sup_{\alpha \in A} f_\alpha(x), \quad x \in \mathbb{R}^n,$$

is convex on U.

(v) *Given* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *convex on* U*, the function*

$$f_U(x) \stackrel{\text{def}}{=} \begin{cases} f(x), & \text{if } x \in U \\ +\infty, & \text{if } x \in \mathbb{R}^n \backslash U \end{cases}$$

already defined in (3.5) *is convex on* $\mathbb{R}^n$.

Remark 7.1.
The above operations can also be applied to convex functions defined on different convex sets by using their extension f_U to $\mathbb{R}^n$. For instance, for $f^i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convex on the convex U_i, $i = 1, 2$, the function

$$\left[f^1_{U_1} + f^2_{U_2} \right](x) \stackrel{\text{def}}{=} \left\{ \begin{matrix} f^1(x) + f^2(x), & \text{if } x \in U_1 \cap U_2 \\ +\infty, & \text{if } x \in \mathbb{R}^n \backslash (U_1 \cap U_2) \end{matrix} \right\} = (f^1 + f^2)_{U_1 \cap U_2}(x)$$

is convex on $\mathbb{R}^n$. □

Proof of Theorem 7.4. (i) By induction. It is true for $m = 1$ and $m = 2$ (by definition of convexity). Assume that it is true for $m = k$. For $m = k + 1$ consider

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i \right), \quad x_i \in U \text{ and } \lambda_i \in]0, 1], 1 \leq i \leq k+1, \quad \sum_{i=1}^{k+1} \lambda_i = 1.$$

Let

$$\lambda = \sum_{j=0}^{k} \lambda_i > 0 \implies \lambda_{k+1} = 1 - \lambda \geq 0.$$

Choose

$$\mu_j \stackrel{\text{def}}{=} \lambda_j / \lambda, \;\; 0 \leq j \leq k.$$

Then

$$\sum_{j=0}^{k}\mu_j = 1,\ \ \mu_j \ge 0 \implies f\left(\sum_{j=0}^{k}\mu_j x_j\right) \le \sum_{j=0}^{k}\mu_j f(x_j).$$

Moreover

$$\sum_{j=0}^{k+1}\lambda_j x_j = \lambda\sum_{j=0}^{k}\mu_j x_j + (1-\lambda)x_{k+1}$$

and by convexity of f

$$\begin{aligned} f\left(\sum_{j=0}^{k+1}\lambda_j x_j\right) &\le \lambda f\left(\sum_{j=0}^{k}\mu_j x_j\right) + (1-\lambda)f(x_{k+1}) \\ &\le \lambda\sum_{j=0}^{k}\mu_j f(x_j) + (1-\lambda)f(x_{k+1}) \le \sum_{j=0}^{k+1}\lambda_j f(x_j). \end{aligned}$$

This completes the proof.

(ii) to (iv) Exercises. □

When U is convex and f is convex on U, it is readily seen that the lower sections F_k, $k > \inf f(U)$, are convex: for all $\lambda \in [0,1]$ and all pairs x and y in F_k,

$$f(x) \le k \text{ and } f(y) \le k \quad \Rightarrow f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y) \le k. \tag{7.6}$$

As the intersection of an arbitrary family of convex sets is convex by Theorem 7.1(iv), $\operatorname{argmin} f(U)$ is convex.

Theorem 7.5. *If $U \subset \mathbb{R}^n$ is convex and $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is convex on U, the set $\operatorname{argmin} f(U)$ is convex.*

7.2 Quasiconvexity

Going back to what we have done in (7.6), it appears that the notion of convexity might be too strong. Indeed, it is sufficient to establish that for all $k \in \mathbb{R}$, the set

$$\{x \in U : f(x) \le k\}$$

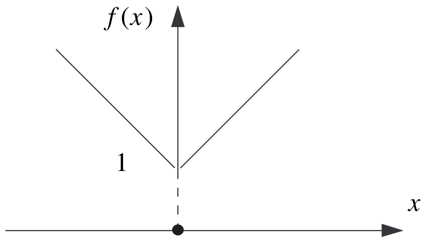
be convex. But to get that property, it is sufficient that

$$\forall x, y \in U \ \text{ such that } \ f(x) \le k \ \text{ and } \ f(y) \le k$$

the property

$$\forall \lambda \in [0,1], \quad f(\lambda x + (1-\lambda)y) \le k$$

be verified. This leads to the less restrictive notion of quasiconvexity.



$$f(x) \stackrel{\text{def}}{=} \begin{Bmatrix} \|x\|+1, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{Bmatrix}$$

Figure 2.5. *Example of a quasiconvex function that is not convex.*

Definition 7.3.
$f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is *quasi-convex*[12] in a convex $U \subset \mathbb{R}^n$ if

$$\forall \lambda \in]0,1[, \ \ \forall x, y \in U, \ \ f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}. \qquad \square$$

Remark 7.2.
A convex function on a convex set is quasiconvex but the converse is not true as illustrated by the quasiconvex function in Figure 2.5. $\square$

Theorem 7.5 generalizes as follows.

Theorem 7.6. *If $U \subset \mathbb{R}^n$ is convex and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is quasi-convex on U,* argmin $f(U)$ *is convex.*

7.3 Strict Convexity and Uniqueness

Definition 7.4.
Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be convex.

(i) A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strictly convex* on U if

$$\forall \lambda \in]0,1[, \ \ \forall x, y \in U, \ \ x \neq y, \ \ f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y).$$

(ii) A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strongly convex* on U if it is convex on U and

$$\exists \delta > 0 \ \text{ such that } \ \forall x, y \in U, \ \ f\left(\frac{x+y}{2}\right) \leq \frac{1}{2} f(x) + \frac{1}{2} f(y) - \frac{1}{4} \delta \|x-y\|^2. \qquad \square$$

The definition is restricted to functions $f : \mathbb{R}^n \to \mathbb{R}$ since the notion of strict convexity of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ would imply $U \subset \operatorname{dom} f$.

[12] "Quasiconvexity was originally introduced as a condition to understand the closely related questions of lower semicontinuity and existence in the calculus of variations. In the last ten years it has become clear that quasiconvexity is also at the root of many fundamental problems in applications, in particular when a change of scale is involved. Examples include the analysis of microstructure, variational models of dislocation structures, the passage from atomistic to continuum models and hybrid analytical computational approaches to multiscale problems." Quasiconvexity and its applications: Perspectives 50 years after C. B. Morrey's seminal paper, Princeton University, November 14–16, 2002. Quoted from http://www.mis.mpg.de/calendar/conferences/2002/quasiconvexity2002.html.

Remark 7.3.
In part (ii) of Definition 7.4 the weaker notion of *midpoint convexity* is implicitly introduced.[13] □

It is easy to check the following implications:

$$\text{strongly convex} \quad \Rightarrow \quad \text{strictly convex} \quad \Rightarrow \quad \text{convex} \quad \Rightarrow \quad \text{quasi-convex}$$

but the converses are not true.

Example 7.1.
The function $f(x) = x$ is convex but not strictly convex on $\mathbb{R}$. □

Example 7.2.
The function $f(x) = e^{-x}$ is strictly convex but not strongly convex on $\mathbb{R}_+$. Consider the expression

$$F(x,y) \stackrel{\text{def}}{=} \frac{\frac{1}{2}f(x) + \frac{1}{2}f(y) - f(\frac{x+y}{2})}{\frac{1}{4}\|x-y\|^2}.$$

By definition, f is strongly convex if there exists $\delta > 0$ such that

$$\forall x, y \in \mathbb{R}_+, \quad F(x,y) \geq \delta.$$

To show that $f(x) = e^{-x}$ is not strongly convex in $\mathbb{R}_+$, choose $(x,y) = (a, a+1)$, $a > 0$. This yields

$$\begin{aligned} F(a,a+1) &= \frac{\frac{1}{2}[e^{-a} + e^{-(a+1)}] - e^{-(\frac{2a+1}{2})}}{\frac{1}{4}} = 2\{e^{-a} + e^{-(a+1)} - 2e^{-(a+\frac{1}{2})}\} \\ &= 2e^{-a}\{1 + e^{-1} - 2e^{-\frac{1}{2}}\} = 2e^{-a}(1 - e^{-\frac{1}{2}})^2, \end{aligned}$$

and $\lim_{a\to 0} F(a,a+1) = 0$. So, for all $\delta > 0$

$$\exists a > 0, \text{ such that } F(a,a+1) = 2e^{-a}(1 - e^{-\frac{1}{2}})^2 < \delta.$$

Therefore there is no $\delta > 0$ for the strong convexity on $\mathbb{R}_+$. □

We have the following condition for the uniqueness of the minimizers.

Theorem 7.7. *Given a strictly convex function $f : \mathbb{R}^n \to \mathbb{R}$ in a nonempty convex $U \subset \mathbb{R}^n$, then* argmin $f(U)$ *is either empty or a singleton.*

[13] A function $f : \mathbb{R}^n \to \mathbb{R}$ is *midpoint convex* on a convex U if

$$\forall x, y \in U, \quad f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}.$$

This condition is only slightly weaker than convexity. For example, a real-valued Lebesgue measurable function that is midpoint convex will be convex (Sierpinski's theorem, p. 12 in William F. Donoghue, *Distributions and Fourier transforms*, Pure and Applied Mathematics vol 32, Elsevier 1969). In particular, a continuous function that is midpoint convex will be convex.

Proof. Assume that $\operatorname{argmin} f(U) \neq \varnothing$ is not a singleton. Let a_1 and a_2 in U be two minimizers of f in U. By convexity of U and strict convexity of f on U,

$$a = \frac{1}{2}(a_1 + a_2) \in U \quad \Rightarrow f(a) < \frac{1}{2} f(a_1) + \frac{1}{2} f(a_2) = \inf_{x \in U} f(x).$$

This contradicts the definition of the infimum. □

Example 7.3.
Go back to Examples 5.2 and 5.9. Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be convex and closed. Given $x \in \mathbb{R}^n$, it has been shown that there exists $\hat{x} \in U$ such that

$$d_U(x) \stackrel{\text{def}}{=} \inf_{y \in U} \|x - y\| = \|x - \hat{x}\|. \tag{7.7}$$

To establish that, we have considered the minimization problem

$$\inf_{y \in U} f(y), \quad f(y) \stackrel{\text{def}}{=} \|y - x\|^2.$$

The function f is strictly convex and even strongly convex in $\mathbb{R}^n$ and U is convex. By Theorem 7.7, $\Pi_U(x) = \operatorname{argmin} f(U)$ is a singleton. The projection of x onto a closed convex set U being unique, it will be denoted $p_U(x)$. □

From this example, we get the following *strict separation theorem*.[14]

Theorem 7.8. *Let U, $\varnothing \neq U \subset \mathbb{R}^n$, be closed and convex.*

(i) *The projection $p_U(x)$ of x onto U is unique and the map*

$$x \mapsto p_U(x) : \mathbb{R}^n \to U \subset \mathbb{R}^n \tag{7.8}$$

is well defined.

(ii) *If $x \notin U$, then*

$$\forall y \in U, \quad -\frac{p_U(x) - x}{\|p_U(x) - x\|} \cdot x \geq -\frac{p_U(x) - x}{\|p_U(x) - x\|} \cdot y + \|p_U(x) - x\|$$

and the hyperplane with unitary normal $n = -(p_U(x) - x)/\|p_U(x) - x\|$ strictly separates x and U.

Proof. (i) Let $x_1, x_2 \in U$ and $\lambda \in [0,1]$. Computing

$$\begin{aligned}
&\|\lambda x_1 + (1-\lambda) x_2 - x\|^2 - \lambda \|x_1 - x\|^2 - (1-\lambda) \|x_2 - x\|^2 \\
=&\|\lambda (x_1 - x)\|^2 + \|(1-\lambda)(x_2 - x)\|^2 + 2\lambda(1-\lambda)(x_1 - x) \cdot (x_2 - x) \\
&- \lambda \|x_1 - x\|^2 - (1-\lambda) \|x_2 - x\|^2 \\
=&(\lambda^2 - \lambda) \|x_1 - x\|^2 + ((1-\lambda)^2 - (1-\lambda)) \|x_2 - x\|^2 + 2\lambda(1-\lambda)(x_1 - x) \cdot (x_2 - x) \\
=&-\lambda(1-\lambda) \|x_1 - x_2\|^2.
\end{aligned}$$

[14] The problem of separating sets by a hyperplane goes back at least to Paul Kirchberger (1878–1945) in 1902. The problem of replacing the "separating hyperplanes" by "separating spherical surfaces" was first posed by Frederick Albert Valentine in 1964. This issue and related topics, including the separation of stationary sheep from goats by a straight fence, are very nicely covered in the chapter on Kirchberger-type theorems of S. R. Lay [**1**].

For $\lambda \in]0,1[$ and $x_1 \neq x_2$, the last line is strictly positive and the function $\|x\|^2$ is indeed strictly convex. It is also strongly convex by setting $\lambda = 1/2$:

$$\left\| \frac{x_1+x_2}{2} - x \right\|^2 = \frac{1}{2}\|x_1-x\|^2 + \frac{1}{2}\|x_2-x\|^2 - \frac{1}{4}\|x_1-x_2\|^2,$$

which corresponds to choosing $\delta = 1$ in the definition.

(ii) From part (i), there exists a unique $\hat{x} \in U$ such that $\|y-x\|^2 \geq \|\hat{x}-x\|^2$ for all $y \in U$. As U is convex, for all y and $\varepsilon \in]0,1]$, $\varepsilon y + (1-\varepsilon)\hat{x} \in U$, and

$$\begin{aligned}
&\|\varepsilon y + (1-\varepsilon)\hat{x} - x\|^2 - \|\hat{x}-x\|^2 \\
&= \|\hat{x}-x+\varepsilon(y-\hat{x})\|^2 - \|\hat{x}-x\|^2 = 2\varepsilon(y-\hat{x})\cdot(\hat{x}-x) + \varepsilon^2\|y-\hat{x}\|^2 \\
&\qquad \Rightarrow \frac{2\varepsilon(y-\hat{x})\cdot(\hat{x}-x)+\varepsilon^2\|y-\hat{x}\|^2}{2\varepsilon} \geq 0.
\end{aligned}$$

By going to the limit as ε goes to 0,

$$\forall y \in U, \quad (y-\hat{x})\cdot(\hat{x}-x) \geq 0$$

or

$$\forall y \in U, \quad (y-x)\cdot(\hat{x}-x) - \|\hat{x}-x\|^2 \geq 0.$$

If $x \notin U$, then $\|\hat{x}-x\| > 0$. So we can choose $n = -(\hat{x}-x)/\|\hat{x}-x\|$,

$$\forall y \in U, \quad -(y-x)\cdot n - \|\hat{x}-x\| \geq 0 \quad \Rightarrow \forall y \in U, \quad n\cdot x \geq n\cdot y + \|\hat{x}-x\|$$

and the hyperplane with unitary normal n strictly separates x and U. □

8 Linear and Affine Subspace and Relative Interior

8.1 Definitions

First recall the following definitions.

Definition 8.1. (i) A subset S of $\mathbb{R}^n$ is a *linear subspace* of $\mathbb{R}^n$ if

$$\forall \alpha, \beta \in \mathbb{R}, \ \forall x, y \in S, \ \alpha x + \beta y \in S. \tag{8.1}$$

(ii) A subset A of $\mathbb{R}^n$ is an *affine subspace* of $\mathbb{R}^n$ if

$$\forall \alpha \in \mathbb{R}, \ \forall x, y \in A, \ \alpha x + (1-\alpha) y \in A. \tag{8.2}$$

□

A linear subspace S of $\mathbb{R}^n$ is closed and contains the origin since $0 = x + (-1)x \in S$. S is said to be of *dimension* k if it is generated by k vectors $v_1, v_2, \ldots, v_k$ of S that are linearly independent in $\mathbb{R}^n$; that is,

$$\forall x \in S, \ \exists \text{ a unique sequence } \alpha_1, \ldots, \alpha_k \in \mathbb{R}, \ x = \alpha_1 v_1 + \cdots + \alpha_k v_k.$$

The *translation* of a subset U of $\mathbb{R}^n$ by a vector a of $\mathbb{R}^n$ is defined as

$$U+a \overset{\text{def}}{=} \{x+a \,:\, x\in U\}. \tag{8.3}$$

By commutativity of the addition, we also get the commutativity of the notation

$$a+U \overset{\text{def}}{=} \{a+x \,:\, x\in U\} = U+a$$
$$U-a \overset{\text{def}}{=} \{x-a \,:\, x\in U\} = \{-a+x \,:\, x\in U\} = -a+U.$$

The *orthogonal set* to a subset U of $\mathbb{R}^n$ is

$$U^\perp \overset{\text{def}}{=} \{y\in\mathbb{R}^n \,:\, y\cdot x=0,\ \ \forall x\in U\}. \tag{8.4}$$

It is a linear (closed) subspace of $\mathbb{R}^n$. When $U=S$, where S is a linear subspace of $\mathbb{R}^n$, we say that $S^\perp$ is the *orthogonal complement* of S.

An affine subspace is the translation of a unique linear subspace.

Theorem 8.1. (i) *Given an affine subspace A of $\mathbb{R}^n$, there exists a unique linear subspace S of $\mathbb{R}^n$ such that*

$$\forall a\in A,\quad A=a+S.$$

In particular, A is closed and convex.

(ii) *Let $\{A_\beta\}_{\beta\in B}$ be a family of affine subspaces of $\mathbb{R}^n$, where the set B of indices is arbitrary and not necessarily finite. Then $\cap_{\beta\in B} A_\beta$ is an affine subspace.*

Proof. (i) Associate with $a\in A$ the set $S=A-a$. For any $\alpha\in\mathbb{R}$ and $x\in S$, there exists x_A such that $x=x_A-a$ and $\alpha x=\alpha(x_A-a)+(1-\alpha)(a-a)=\alpha x_A+(1-\alpha)a-a\in S$. For any $x,y\in S$, there exist x_A and y_A in A such that $x=x_A-a$ and $y=y_A-a$. Then

$$\frac{1}{2}(x+y)=\left(\frac{1}{2}x_A+\left(1-\frac{1}{2}\right)y_A\right)-a\in S\quad\Rightarrow x+y=2\left[\frac{1}{2}(x+y)\right]\in S.$$

So for α and β in $\mathbb{R}$ and x and y in S, $\alpha x\in S$ and $\beta y\in S$, and their sum $\alpha x+\beta y\in S$. Therefore, S is a linear subspace.

It remains to show that the subspace $S=A-a$ is independent of the choice of $a\in A$. Consider another point $a'\in A$ and its linear subspace $S'\overset{\text{def}}{=}A-\{a'\}$. Then

$$\forall x'\in S'=A-\{a'\},\ \exists x'_A\in A \text{ such that } x'=x'_A-a'$$
$$\Rightarrow\ x'=\underbrace{x'_A-a}_{\in S}+\underbrace{(-1)(a'-a)}_{\in S}\in S\quad\Rightarrow\ S'\subset S$$

and conversely

$$\forall x\in S=A-\{a\},\ \exists x_A\in A \text{ such that } x=x_A-a$$
$$\Rightarrow\ x=\underbrace{x_A-a'}_{\in S}+\underbrace{(-1)(a-a')}_{\in S}\in S\quad\Rightarrow\ S\subset S'.$$

So the linear subspace is unique and for all $a\in A$, $S=A-a$.

(ii) Same proof as the one of Theorem 7.1(iv) for convex sets. $\square$

Definition 8.2.
Let $U \subset \mathbb{R}^n$, $U \neq \varnothing$.

(i) $\text{aff}\,(U)$ is the intersection of all affine subspaces of $\mathbb{R}^n$ that contain U.

(ii) The *relative interior* of U is

$$\text{ri}\,(U) \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : \exists \varepsilon > 0 \text{ such that } B_\varepsilon(x) \cap \text{aff}\,(U) \subset U\right\}.$$

This implies that $\text{ri}\,(U) \subset U$. □

If $U = \{x\}$ is a singleton, then $U = \text{aff}\,(U) = \text{ri}\,(U)$.

8.2 Domain of Convex Functions

By Lemma 7.1, the domain $\text{dom}\, f$ of a convex function is convex and is contained in the smallest affine subspace of $\mathbb{R}^n$ that contains it. The following lemmas will be used later on.

Lemma 8.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a convex function.*

(i) $\text{dom}\, f$ *and* $\text{ri}\,(\text{dom}\, f)$ *are convex.*

(ii) $\text{ri}\,(\text{dom}\, f) \neq \varnothing$ *if and only if* $\text{dom}\, f \neq \varnothing$.

(iii) *If* $\text{dom}\, f \neq \varnothing$*, then* $\overline{\text{ri}\,(\text{dom}\, f)} = \overline{\text{dom}\, f}$.

Proof. (i) From Lemma 7.1(i), $\text{dom}\, f$ is convex. By Theorem 7.1(i), applied to $\text{aff}\,(\text{dom}\, f)$, $\text{ri}\,(\text{dom}\, f)$ is also convex.

(ii) If $\text{dom}\, f$ is a singleton, then $\text{ri}\,(\text{dom}\, f) = \text{dom}\, f \neq \varnothing$. If the dimension of $\text{aff}\,(\text{dom}\, f)$ is greater than or equal to one, then, by convexity of $\text{dom}\, f$, U contains an open nonempty segment that is contained in $\text{ri}\,(\text{dom}\, f)$.

(iii) From part (ii), $\varnothing \neq \text{ri}\,(\text{dom}\, f) \subset \text{aff}\,(\text{dom}\, f)$. Hence, by convexity of $\text{dom}\, f$ and Theorem 7.1(v), $\overline{\text{ri}\,(\text{dom}\, f)} = \overline{\text{dom}\, f}$. □

Lemma 8.2. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$,* $\text{dom}\, f \neq \varnothing$*, be a convex function. Denote by $S(\text{dom}\, f)$ the unique linear subspace of $\mathbb{R}^n$ associated with* $\text{aff}\,(\text{dom}\, f)$ *and by m the dimension of $S(\text{dom}\, f)$.*

(i) *Given $x \in \text{dom}\, f$, there exists $\{x_1, \ldots, x_m\} \subset \text{dom}\, f$ such that the set of vectors $\{x_i - x : i = 1, \ldots, m\}$ is linearly independent and it forms a basis of $S(\text{dom}\, f)$.*

(ii) *The function*

$$y \mapsto f_x(y) \stackrel{\text{def}}{=} f(y + x) - f(x) : S(\text{dom}\, f) \to \mathbb{R} \tag{8.5}$$

is convex with domain $\text{dom}\, f_x = \text{dom}\, f - x$ *and* $\text{ri}\,(\text{dom}\, f) \neq \varnothing$.

Proof. (i) In view of Theorem 8.1, $\text{aff}\,(\text{dom}\, f) = x + S(\text{dom}\, f)$. Let m be the dimension of the linear subspace $S(\text{dom}\, f)$. Therefore,

$$\begin{aligned} S(\text{dom}\, f) &= \text{aff}\,(\text{dom}\, f) - x \\ &= \{\alpha(a - x) + (1 - \alpha)(b - x) : \forall a, b \in \text{dom}\, f \text{ and } \forall \alpha \in \mathbb{R}\}, \end{aligned}$$

and there exists a set of m linearly independent vectors of the form $a-x$, $a \in \operatorname{dom} f$, that generate $S(\operatorname{dom} f)$.

(ii) By definition and by observing that if $\operatorname{dom} f \neq \varnothing$, $\operatorname{ri}(\operatorname{dom} f) \neq \varnothing$. □

In the study of convex functions with nonempty domain, it can be advantageous to work with f_x since $\operatorname{dom} f_x$ has a nonempty interior in the subspace $S(\operatorname{dom} f)$.

9 Convexification and Fenchel–Legendre Transform

9.1 Convex lsc Functions as Upper Envelopes of Affine Functions

Convex lsc functions are nice functions that naturally occur in conjunction with the infimum of an arbitrary objective function as we shall later see in section 9.4.

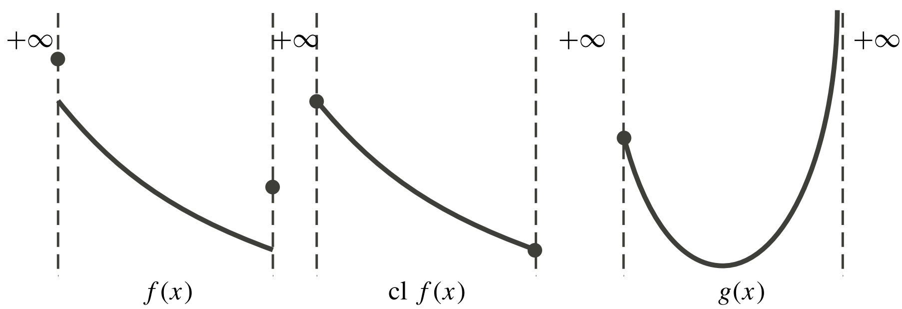


Figure 2.6. *Examples of convex functions: f (not lsc),* cl f*, and g (lsc).*

Theorem 9.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex and lsc.*

(i) $\operatorname{dom} f$ *and* $\operatorname{ri}(\operatorname{dom} f)$ *are convex subsets of* $\mathbb{R}^n$ *and* $\operatorname{epi} f$ *is a closed convex subset of* $\mathbb{R}^n \times \mathbb{R}$.

(ii) *If* $\operatorname{dom} f$ *is not a singleton, then*

$$\forall x \in \mathbb{R}^n, \quad \liminf_{y \to x} f(y) = f(x). \tag{9.1}$$

Caution! The second one-dimensional example of Figure 2.6 is misleading. In dimension greater than or equal to 2, the liminf cannot be replaced by a simple limit in (9.1) as can be seen from the following example.

Example 9.1.
Consider the following convex function defined on $\mathbb{R}^2$ (Figure 2.7):

$$f(x,y) \overset{\text{def}}{=} \begin{cases} x^2/y, & y > 0 \\ 0, & (x,y) = (0,0) \\ +\infty, & (x,0),\ x \neq 0 \\ +\infty, & y < 0. \end{cases} \tag{9.2}$$

The domain is $\text{dom}\, f = \{(x,y) : y > 0\} \cup \{(0,0)\}$, $f(x,y) \geq 0$ for all (x,y), and

$$\lim_{n\to\infty} f(1/n, 1/n) = \lim_{n\to\infty} 1/n \to 0 \qquad \Rightarrow \liminf_{(x,y)\to(0,0)} = \liminf_{(x,y)\in \text{dom}\, f\to(0,0)} = 0$$

$$\lim_{n\to\infty} f(1/n, 1/n^3) = \lim_{n\to\infty} n \to +\infty \qquad \Rightarrow \limsup_{(x,y)\to(0,0)} = \limsup_{(x,y)\in \text{dom}\, f\to(0,0)} = +\infty.$$

Figure 2.7 shows the function for $y \geq \varepsilon > 0$, but the parabola closes around the z axis as ε goes to zero. □

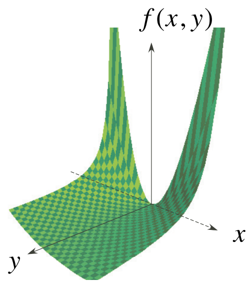


Figure 2.7. *The function $f(x,y) = x^2/y$ for $y \geq \varepsilon > 0$ and some small ε.*

Proof of Theorem 9.1. (i) By Lemma 8.1(i), dom f and ri(dom f) are convex. By Theorem 7.3 and Lemma 4.1, epi f is convex and closed.

(ii) For an lsc function, we already know by Theorem 4.3 that

$$\forall x \in \mathbb{R}^n, \quad \liminf_{y\to x} f(y) \geq f(x). \tag{9.3}$$

If dom f is not a singleton, from Lemma 7.2,

$$\liminf_{y\to x} f(y) \leq f(x).$$

and we get (9.1) from (9.3). □

The next theorem establishes the intimate connection between the lsc convexity of a function f and the upper envelope of affine functions dominated by f.

Theorem 9.2 (R. T. Rockafellar [**1**, Thm. 12.1, p. 102])**.** *Let $\mathcal{A}$ be the set of all affine functions $\ell : \mathbb{R}^n \to \mathbb{R}$ of the form*

$$\ell(x) = a \cdot x + b, \quad a \in \mathbb{R}^n \text{ and } b \in \mathbb{R}. \tag{9.4}$$

If $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, dom $f \neq \varnothing$, is convex and lsc, then

$$f(x) = \sup_{\substack{\ell\in\mathcal{A} \\ \ell \leq f \text{ on } \mathbb{R}^n}} \ell(x) = \sup_{a\in\mathbb{R}^n} \left\{ a\cdot x - \sup_{y\in\mathbb{R}^n} [y\cdot a - f(y)] \right\}, \quad x \in \mathbb{R}^n. \tag{9.5}$$

The underlying geometric interpretation is that a closed convex set is the intersection of all the half spaces containing it which is a direct consequence of the Hahn–Banach theorem. A direct proof is provided here that will necessitate the following general result (Theorem 9.3) that will be proved right after the proof of Theorem 9.2. Its proof using the *distance function* to a closed convex set (see Example 7.3) is straightforward for finite-valued functions. However, when the function takes the value $+\infty$, a "tangent vertical hyperplane" to the epigraph is not sufficient. It is necessary to show that a hyperplane slightly slanted off the vertical can be constructed (see Figure 2.8) in order to generate an affine function under the epigraph.[15]

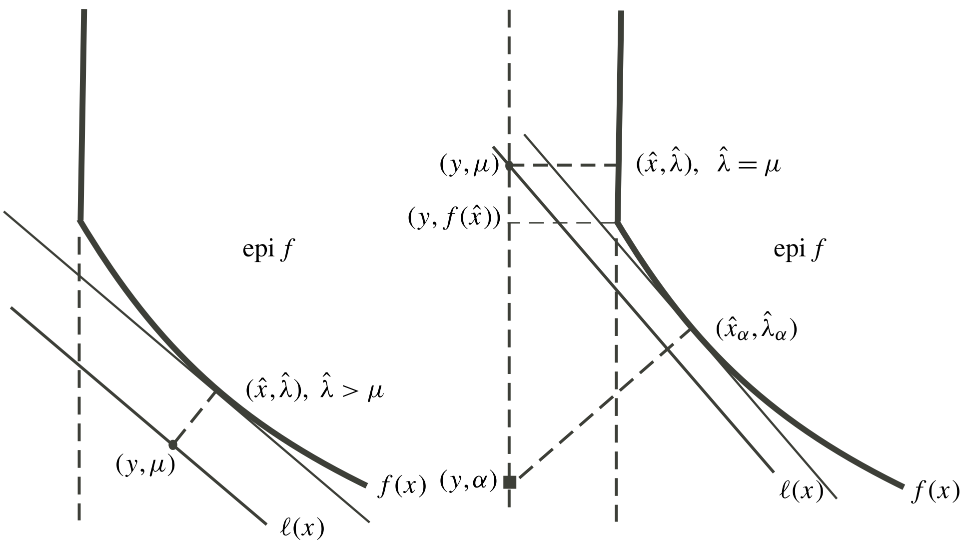


Figure 2.8. *Cases $y \in \operatorname{dom} f$ (left) and $y \notin \operatorname{dom} f$ (right).*

Theorem 9.3. *Let $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, $\operatorname{dom} f \neq \varnothing$, be convex and lsc. Then, for all $(y,\mu) \notin \operatorname{epi} f$, there exists an affine function of the form $\ell(x) = \mu + a \cdot (y - x)$, $a \in \mathbb{R}^n$, such that for all $x \in \mathbb{R}^n$, $f(x) \geq \ell(x)$.*

Proof of Theorem 9.2. Define

$$\hat{f}(x) \stackrel{\text{def}}{=} \sup_{\substack{\ell \in \mathcal{A} \\ \ell \leq f \text{ on } \mathbb{R}^n}} \ell(x), \quad x \in \mathbb{R}^n. \tag{9.6}$$

As an upper envelope of convex lsc functions, $\hat{f}$ is convex and lsc and $\hat{f}(x) \leq f(x)$. For each $y \in \mathbb{R}^n$, there exists a monotone increasing sequence $\{\mu_n\}$, $\mu_n < f(y)$, such that $\mu_n \to f(y)$. As $(y,\mu_n) \notin \operatorname{epi} f$, by Theorem 9.3, there exists $\ell_n \in \mathcal{A}$ such that $\ell_n(y) = \mu_n$ and $f(x) \geq \ell_n(x)$ for all x. By going to the limit

$$\ell_n(y) \leq \hat{f}(y) \leq f(y) \quad \Rightarrow f(y) = \lim_{n\to\infty} \ell_n(y) \leq \hat{f}(y) \leq f(y)$$

[15]Cf. R. T. ROCKAFELLAR [**1**, Cor. 12.1.1, p. 103] for another proof.

and $\hat{f}(y) = f(y)$. Since each $\ell \in \mathcal{A}$ is of the form $\ell(x) = a \cdot x + b$, we get

$$\hat{f}(x) = \sup_{\substack{a\in\mathbb{R}^n \text{ and } b\in\mathbb{R} \\ \forall y\in\mathbb{R}^n, a\cdot y+b\le f(y)}} a\cdot x + b = \sup_{a\in\mathbb{R}^n} \left\{ a\cdot x + \inf_{y\in\mathbb{R}^n} [f(y) - a\cdot y] \right\}$$

since

$$\forall y \in \mathbb{R}^n, \quad a\cdot y + b \le f(y) \quad \Rightarrow b \le \inf_{y\in\mathbb{R}^n} [f(y) - a\cdot y] = -\sup_{y\in\mathbb{R}^n} [a\cdot y - f(y)].$$

This yields expression (9.5). □

Proof of Theorem 9.3. (i) If f is identically equal to $+\infty$, choose the constant function $\ell(x) = \mu$.

(ii) If f is not identically equal to $+\infty$, then $\text{epi}\, f \neq \varnothing$. As, by assumption, f is convex and lsc, $\text{epi}\, f$ is convex, closed, and nonempty. The objective function

$$(x,\lambda) \mapsto J(x,\lambda) \overset{\text{def}}{=} \|x-y\|^2 + |\lambda - \mu|^2 : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \tag{9.7}$$

is continuous, has the growth property, and is strictly convex. So there exists a unique point $(\hat{x},\hat{\lambda}) \in \text{epi}\, f$ such that $J(\hat{x},\hat{\lambda}) = \inf J(\text{epi}\, f)$. By convexity, for $\varepsilon \in]0,1[$ and $(x,\lambda) \in \text{epi}\, f$, we have $(\hat{x},\hat{\lambda}) + \varepsilon(x-\hat{x},\lambda-\hat{\lambda}) \in \text{epi}\, f$ and

$$\frac{J((\hat{x},\hat{\lambda}) + \varepsilon(x-\hat{x},\lambda-\hat{\lambda})) - J(\hat{x},\hat{\lambda})}{\varepsilon} \ge 0.$$

From expression of J this quotient reduces to

$$\begin{aligned}&\frac{\|\hat{x}+\varepsilon(x-\hat{x})-y\|^2 + |\hat{\lambda}+\varepsilon(\lambda-\hat{\lambda})-\mu|^2 - \|\hat{x}-y\|^2 - |\hat{\lambda}-\mu|^2}{\varepsilon} \\ &= (2(\hat{x}-y)+\varepsilon(x-\hat{x})\cdot(x-\hat{x}) + (2(\hat{\lambda}-\mu)+\varepsilon(\lambda-\hat{\lambda}))(\lambda-\hat{\lambda}).\end{aligned}$$

Let ε go to 0. Therefore, there exists $\hat{x} \in \text{dom}\, f$ such that $\hat{\lambda} \ge f(\hat{x})$ and

$$(\hat{x}-y)\cdot(x-\hat{x}) + (\hat{\lambda}-\mu)(\lambda-\hat{\lambda}) \ge 0, \quad \forall (x,\lambda) \in \text{epi}\, f. \tag{9.8}$$

Choosing $(x,\lambda) = (\hat{x},\hat{\lambda}+1)$, we get $\hat{\lambda} \ge \mu$. We get two cases: $\hat{\lambda} > \mu$ and $\hat{\lambda} = \mu$.

If $\hat{\lambda} > \mu$, by choosing $(x,\lambda) = (\hat{x}, f(\hat{x}))$, we get the inequality $(\hat{\lambda}-\mu)(f(\hat{x})-\hat{\lambda}) \ge 0$ that implies $f(\hat{x}) \ge \hat{\lambda} \ge f(\hat{x})$ and $\hat{\lambda} = f(\hat{x})$. Finally, for all $x \in \text{dom}\, f$,

$$f(x) \ge f(\hat{x}) + \frac{1}{f(\hat{x})-\mu}(\hat{x}-y)\cdot(\hat{x}-x). \tag{9.9}$$

But

$$\begin{aligned}
&f(\hat{x})+\frac{1}{f(\hat{x})-\mu}(\hat{x}-y)\cdot(\hat{x}-x) \\
&=\mu+\frac{|f(\hat{x})-\mu|^2+\|\hat{x}-y\|^2}{f(\hat{x})-\mu}+\frac{1}{f(\hat{x})-\mu}(\hat{x}-y)\cdot(\hat{x}-x).
\end{aligned} \tag{9.10}$$

By choosing

$$\ell(x) \overset{\text{def}}{=} \mu+\frac{1}{f(\hat{x})-\mu}(\hat{x}-y)\cdot(\hat{x}-x) \tag{9.11}$$

we get $f(y)=\mu$, $f(x)\geq \ell(x)$ for all $x\in \operatorname{dom} f$, and, a fortiori, for all $x\in\mathbb{R}^n\setminus\operatorname{dom} f$ since $f(x)=+\infty$.

The second case is $\hat{\lambda}=\mu$. If $y\in\operatorname{dom} f$, then $f(y)<+\infty$ and, from (9.8), we get

$$(\hat{x}-y)\cdot(y-\hat{x})\geq 0 \tag{9.12}$$

and hence $\hat{x}=y$ that yields the contradiction $(\hat{x},\hat{\lambda})=(y,\mu)$. Therefore $f(y)=+\infty$. In this case the variational inequality (9.12) does not directly give the function ℓ. Getting it requires the auxiliary construction on the right-hand side of Figure 2.8. Associate with each $\alpha<f(\hat{x})$, the point $(y,\alpha)\notin \operatorname{epi} f$ for which there exists a unique point $(\hat{x}_\alpha,\hat{\lambda}_\alpha)\in\operatorname{epi} f$ such that $\hat{\lambda}_\alpha\geq f(\hat{x}_\alpha)$ and

$$(\hat{x}_\alpha-y)\cdot(x-\hat{x}_\alpha)+(\hat{\lambda}_\alpha-\alpha)(\lambda-\hat{\lambda}_\alpha)\geq 0,\quad \forall (x,\lambda)\in\operatorname{epi} f. \tag{9.13}$$

As before, choosing $(x,\lambda)=(\hat{x}_\alpha,\lambda_\alpha+1)$ yields $\hat{\lambda}_\alpha\geq\alpha$. If $\hat{\lambda}_\alpha=\alpha$, then from (9.13) and inequality (9.12) with $x=\hat{x}_\alpha$, we get

$$\begin{aligned}
0\leq(\hat{x}_\alpha-y)\cdot(\hat{x}-\hat{x}_\alpha)&=-\|\hat{x}_\alpha-\hat{x}\|^2+(\hat{x}-y)\cdot(\hat{x}-\hat{x}_\alpha)\\
\Rightarrow \|\hat{x}_\alpha-\hat{x}\|^2&\leq(\hat{x}-y)\cdot(\hat{x}-\hat{x}_\alpha)\leq 0 \quad\Rightarrow \hat{x}_\alpha=\hat{x}.
\end{aligned}$$

We then get the contradiction $(\hat{x},\alpha)=(\hat{x}_\alpha,\hat{\lambda}_\alpha)\in\operatorname{epi} f$ since $\alpha<f(\hat{x})$.

In summary, for $\hat{\lambda}=\mu$, we have $f(y)=+\infty$ and for all $\alpha<f(\hat{x})$ we get the strict inequality $\hat{\lambda}_\alpha>\alpha$. By an argument similar to the previous one, $\hat{\lambda}_\alpha=f(\hat{x}_\alpha)$. Moreover, we get $f(\hat{x})\geq f(\hat{x}_\alpha)$ by choosing $(x,\lambda)=(\hat{x},f(\hat{x}))$ in (9.13) since

$$\begin{aligned}
0&\leq(\hat{x}_\alpha-y)\cdot(\hat{x}-\hat{x}_\alpha)+(f(\hat{x}_\alpha)-\alpha)(f(\hat{x})-f(\hat{x}_\alpha))\\
&=-\|\hat{x}-\hat{x}_\alpha\|^2+(\hat{x}-y)\cdot(\hat{x}-\hat{x}_\alpha)+(f(\hat{x}_\alpha)-\alpha)(f(\hat{x})-f(\hat{x}_\alpha))\\
&\Rightarrow (f(\hat{x}_\alpha)-\alpha)(f(\hat{x})-f(\hat{x}_\alpha))\geq\|\hat{x}-\hat{x}_\alpha\|^2+(\hat{x}-y)\cdot(\hat{x}_\alpha-\hat{x})\geq 0\\
&\Rightarrow f(\hat{x})-f(\hat{x}_\alpha)\geq 0
\end{aligned}$$

from inequality (9.12) with $x=\hat{x}_\alpha$. Then, from (9.13) for all $x\in\operatorname{dom} f$,

$$f(x)\geq f(\hat{x}_\alpha)+\frac{(\hat{x}_\alpha-y)\cdot(\hat{x}_\alpha-x)}{f(\hat{x}_\alpha)-\alpha}=f(\hat{x}_\alpha)+\frac{\|\hat{x}_\alpha-y\|^2}{f(\hat{x}_\alpha)-\alpha}+\frac{(\hat{x}_\alpha-y)\cdot(y-x)}{f(\hat{x}_\alpha)-\alpha}.$$

But

$$\begin{aligned} f(\hat{x}_\alpha) + \frac{\|\hat{x}_\alpha - y\|^2}{f(\hat{x}_\alpha)-\alpha} + \frac{(\hat{x}_\alpha - y)\cdot(y-x)}{f(\hat{x}_\alpha)-\alpha} \\ = \alpha + \frac{|f(\hat{x}_\alpha)-\alpha|^2 + \|\hat{x}_\alpha - y\|^2}{f(\hat{x}_\alpha)-\alpha} + \frac{(\hat{x}_\alpha - y)\cdot(y-x)}{f(\hat{x}_\alpha)-\alpha}. \end{aligned} \tag{9.14}$$

If there exists $\alpha < f(\hat{x})$ such that

$$\alpha + \frac{|f(\hat{x}_\alpha)-\alpha|^2 + \|\hat{x}_\alpha - y\|^2}{f(\hat{x}_\alpha)-\alpha} \geq \mu, \tag{9.15}$$

choose

$$\ell(x) \stackrel{\text{def}}{=} \mu + \frac{(\hat{x}_\alpha - y)\cdot y - x)}{f(\hat{x}_\alpha)-\alpha} \tag{9.16}$$

to get $f(x) \geq \ell(x)$ for all x and $\ell(y) = \mu$.

Finally, proceed by contradiction. Assume that for all $\alpha < f(\hat{x})$, we have

$$\alpha + \frac{|f(\hat{x}_\alpha)-\alpha|^2 + \|\hat{x}_\alpha - y\|^2}{f(\hat{x}_\alpha)-\alpha} < \mu. \tag{9.17}$$

Recalling that $f(\hat{x}) \geq f(\hat{x}_\alpha)$, we get the following inequality:

$$0 \leq |f(\hat{x}_\alpha)-\alpha|^2 + \|\hat{x}_\alpha - y\|^2 < (\mu-\alpha)(f(\hat{x}_\alpha)-\alpha) \leq (\mu-\alpha)(f(\hat{x})-\alpha)$$

since $\mu - \alpha = (\mu - \hat{\lambda}) + (\hat{\lambda} - f(\hat{x})) + (f(\hat{x}) - \alpha) > 0$. As $\alpha \to f(\hat{x})$, this yields

$$(\hat{x}_\alpha, f(\hat{x}_\alpha)) \to (y, f(\hat{x})) \notin \operatorname{epi} f. \tag{9.18}$$

Since $(\hat{x}_\alpha, f(\hat{x}_\alpha)) \in \operatorname{epi} f$, this contradicts the fact that $\operatorname{epi} f$ is closed. □

9.2 Fenchel–Legendre Transform

Expression (9.5) in Theorem 9.2 exhibits a famous transform in physics and mechanics that is related to *duality* in optimization.

Definition 9.1.
The *Fenchel–Legendre transform*[16] [17] of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is the function

$$x^* \mapsto f^*(x^*) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}^n} \left\{ y \cdot x^* - f(y) \right\} : \mathbb{R}^n \to \overline{\mathbb{R}} \tag{9.19}$$

[16] Adrien-Marie Legendre (1752–1833). He is credited for the first definition of the *transform* in physics and mechanics for differentiable functions that Fenchel generalized to nondifferentiable functions.

[17] Moritz Werner Fenchel (1905–1988) is known for his contributions to geometry and optimization (see W. FENCHEL and T. BONNESEN [**1**, **2**, **3**, **4**]). He established the basic results of convex analysis and nonlinear programming.

and the *Fenchel–Legendre bitransform* is the function

$$x^{**} \mapsto f^{**}(x^{**}) \stackrel{\text{def}}{=} (f^*)^*(x^{**}) = \sup_{x^* \in \mathbb{R}^n} \left\{ x^* \cdot x^{**} - f^*(x^*) \right\} : \mathbb{R}^n \to \overline{\mathbb{R}}. \tag{9.20}$$

R. T. ROCKAFELLAR [**1**, Chapter 5] uses the terminology *conjugate function* for the transform f^* of a convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. □

By Theorem 9.2, it is readily seen that $f = f^{**}$ when f is convex lsc.

It is useful to review some of the main properties of the transform and the bitransform of functions that are not necessarily convex and lsc.

Theorem 9.4. *Let* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.

(i) *If* $f \equiv +\infty$*, then* $f^* \equiv -\infty$[18] *and* $f^{**} \equiv +\infty$.

(ii) *If there exists* $\bar{x} \in \mathbb{R}^n$ *such that* $f(\bar{x}) < +\infty$*, then* $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *and* f^* *is convex lsc.*

(iii) *If there exists* $\bar{y} \in \mathbb{R}^n$ *such that* $f(\bar{y}) = -\infty$*, then* $f^* \equiv +\infty$ *and* $f^{**} \equiv -\infty$.

(iv) *If* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex* lsc*, then* $f^{**} = f$*,* $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex* lsc*, and*

$$\text{ri}\,(\text{dom}\, f) \neq \varnothing \;\Leftrightarrow\; \text{dom}\, f \neq \varnothing \;\Leftrightarrow\; \text{dom}\, f^* \neq \varnothing \;\Leftrightarrow\; \text{ri}\,(\text{dom}\, f^*) \neq \varnothing.$$

Proof. If f is convex lsc, then, by Theorem 9.2, $f^{**} = f$ and $\text{dom}\, f^{**} = \text{dom}\, f$.

(i) By definition.

(ii) Since there exists $\bar{x} \in \mathbb{R}^n$ such that $f(\bar{x}) < +\infty$, we have

$$\forall p \in \mathbb{R}^n, \quad f^*(p) = \sup_{x \in \mathbb{R}^n} p \cdot x - f(x) \geq p \cdot \bar{x} - f(\bar{x}) > -\infty$$

and $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. As for the convexity, by definition, for all $\lambda \in]0,1[$

$$\begin{aligned}
(\lambda x_1^* + (1-\lambda) x_2^*) \cdot x - f(x) &= \lambda \left(x_1^* \cdot x - f(x) \right) + (1-\lambda) \left(x_2^* \cdot x - f(x) \right) \\
&\leq \lambda f^*(x_1^*) + (1-\lambda) f^*(x_2^*) \\
\Rightarrow\; f^*(\lambda x_1^* + (1-\lambda) x_2^*) &\leq \lambda f^*(x_1^*) + (1-\lambda) f^*(x_2^*)
\end{aligned}$$

since $f^*(x^*)$ either is finite or $+\infty$. Therefore f^* is convex.

To establish the lower semicontinuity at $p \in \mathbb{R}^n$ let $h < f^*(p)$. As $f^*(p) > -\infty$, by definition of the supremum, there exists $\hat{x} \in \mathbb{R}^n$ such that

$$f^*(p) \geq \hat{x} \cdot p - f(\hat{x}) > h.$$

If $\hat{x} \neq 0$, then

$$f^*(q) \geq \hat{x} \cdot q - f(\hat{x}) = \hat{x} \cdot (q-p) + \hat{x} \cdot p - f(\hat{x}) \geq -\|q-p\|\,\|\hat{x}\| + \hat{x} \cdot p - f(\hat{x}) > h$$

[18] Recall that, by convention, a function identically equal to $-\infty$ is convex lsc.

for all $q \in B_\varepsilon(p)$ with

$$\varepsilon = \frac{\hat{x}\cdot p - f(\hat{x}) - h.}{\|\hat{x}\|}.$$

If $\hat{x} = 0$, we can pick $V(p) = \mathbb{R}^n$ since $-f(0) > h$ and for all $q \in \mathbb{R}^n$, $f^*(q) \geq q\cdot 0 - f(0) = -f(0) > h$. Therefore f^* is lsc on $\mathbb{R}^n$.

(iii) If there exists $\bar{y} \in \mathbb{R}^n$ such that $f(\bar{y}) = -\infty$, then for all p, $f^*(p) \geq p\cdot\bar{y} - f(\bar{y}) = +\infty$ and f^* identically equals to $+\infty$.

(iv) The equivalences with the relative interiors of the domains on the extreme left and right follow from Lemma 8.1(ii). So it is sufficient to prove the equivalence for the domains. If f is convex lsc, then $f^{**} = f$ and $\operatorname{dom} f^{**} = \operatorname{dom} f$. From part (ii), if $\operatorname{dom} f \neq \varnothing$, $f^* : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ and f^* is convex lsc. If $\operatorname{dom} f^* = \varnothing$, then for all $x^* \in \mathbb{R}^n$,

$$y\cdot x^* - f^*(x^*) = -\infty \quad \Rightarrow f(y) = f^{**}(y) = -\infty$$

that contradicts the fact that $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$. The proof of the converse is the same applied to f^* in place of f since $f^* : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ and f^* is convex lsc. So $\operatorname{dom} f^{**} \neq \varnothing$ and as $f^{**} = f$, $\operatorname{dom} f \neq \varnothing$. □

Example 9.2 (Exercise 10.10)**.**
For $f(x) = a\cdot x + b$,

$$f^*(x^*) = \begin{cases} -b, & \text{if } x^* = a \\ +\infty, & \text{if } x^* \neq a \end{cases} \quad \text{and} \quad f^{**}(x) = a\cdot x + b = f(x).$$

For $f(x) = \|x\|$,

$$f^*(x^*) = \begin{cases} 0, & \text{if } \|x^*\| \leq 1 \\ +\infty, & \text{if } \|x^*\| > 1 \end{cases} \quad \text{and} \quad f^{**}(x) = \|x\| = f(x).$$

For the concave function $x \mapsto f(x) = -x^2$, we have $f^* \equiv +\infty$ and $f^{**} \equiv -\infty$. For the function $x \mapsto f(x) = \sin x$, we have $f^*(0) = 1$ and $f^*(p) = +\infty$ for $p \neq 0$, $f^{**} \equiv -1$. □

Example 9.3 (Exercise 10.11)**.**
For a positive semidefinite matrix A and $f(x) = \frac{1}{2}(Ax)\cdot x + b\cdot x$,

$$\begin{cases} f^*(Ay+b) = \dfrac{1}{2}(Ay)\cdot y, & \forall y \in \mathbb{R}^n \\ f^*(x^*) = +\infty, & \forall x^* \notin \operatorname{Im} A + b \end{cases}$$

and $\operatorname{dom} f^* = b + A\mathbb{R}^n$. f^* is well defined since for y_1, y_2 such that $Ay_1 + b = x^* = Ay_2 + b$, $y_2 - y_1 \in \operatorname{Ker} A$ and $Ay_1\cdot y_1 = Ay_2\cdot y_1 = y_2\cdot Ay_1 = Ay_2\cdot y_2$. Equivalently,

$$f^*(x^*) = \begin{cases} \sup\limits_{y\in\mathbb{R}^n,\, Ay+b=x^*} \dfrac{1}{2}(Ay)\cdot y, & \text{if } x^* \in b + A\mathbb{R}^n \\ +\infty, & \text{if } x^* \notin b + A\mathbb{R}^n. \end{cases}$$

If A is a positive definite matrix, then A is invertible, $\operatorname{Im} A + b = \mathbb{R}^n$, and

$$f^*(x^*) = \frac{1}{2}(A^{-1}(x^* - b))\cdot(x^* - b), \quad \forall x^* \in \mathbb{R}^n. \qquad \square$$

Example 9.4.
For $f(x) = \|x\|^p/p$, $1 < p < \infty$,

$$f^*(x^*) = \frac{1}{q}\|x^*\|^q, \quad \frac{1}{q}+\frac{1}{p} = 1, \quad \text{and} \quad f^{**}(x) = \frac{1}{p}\|x\|^p = f(x).$$

For the exponential function $f(x) = \beta e^{fx}$, $\beta > 0$, $f > 0$,

$$f^*(x^*) = \begin{cases} \dfrac{x^*}{f}\left(\ln\dfrac{x^*}{\beta f} - 1\right), & \text{if } x^* > 0 \\ 0, & \text{if } x^* = 0 \\ +\infty, & \text{if } x^* < 0. \end{cases}$$

□

Example 9.5.
For $f(x) = \|x-a\|$, $a \in \mathbb{R}^n$,

$$f^*(x^*) = x^* \cdot a + I_{\overline{B_1(0)}}(x^*). \tag{9.21}$$

Consider the estimate

$$x^* \cdot x - \|x-a\| = x^* \cdot a + x^* \cdot (x-a) - \|x-a\| \le x^* \cdot a + (\|x^*\|-1)\|x-a\|.$$

If $\|x^*\| \le 1$, then the supremum over x achieved in $x = a$ is 0; if $\|x^*\| > 1$, then the supremum over x is $+\infty$. This yields the above expression. □

Example 9.6 (Exercise 10.12)**.**
Given two proper functions for the infimum $f, g : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, their *infimal convolution*[19] is defined as

$$(f \,\square\, g)(x) \stackrel{\text{def}}{=} \inf_{y\in\mathbb{R}^n} [f(x-y)+g(y)] \tag{9.22}$$

and its Fenchel–Legendre transform is given by

$$(f \,\square\, g)^* = f^* + g^*. \tag{9.23}$$

If, in addition, f and g are convex, then $f \,\square\, g$ is convex. For $f(x) = \|x\|_{\mathbb{R}^n}$ and $g = I_U$, the indicator function of a nonempty set $U \subset \mathbb{R}^n$, $f \,\square\, g = d_U$, the distance function to U. Thus, if U is convex, then I_U and d_U are convex. □

Example 9.7 (Exercise 10.13)**.**
(i) Let $B : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, and define the *marginal value* function

$$(Bf)(y) \stackrel{\text{def}}{=} \begin{cases} \inf\limits_{x\in\mathbb{R}^n} \{f(x) : Bx = y\}, & \text{if } y \in \operatorname{Im} B \\ +\infty, & \text{if } y \notin \operatorname{Im} B. \end{cases} \tag{9.24}$$

Its Fenchel–Legendre transform is $(Bf)^* = f^* \circ B^\top$ and, if f is convex, then Bf is convex.

[19] R. T. ROCKAFELLAR [**1**, pp. 34 and 145].

(ii) Let $A:\mathbb{R}^m\to\mathbb{R}^n$ be a linear map and $g:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$. A direct computation yields

$$(g\circ A)^*(y^*)=\sup_{y\in\mathbb{R}^m}\left\{y^*\cdot y-g(Ay)\right\},$$

$$(A^\top g^*)(y^*)=\begin{cases}\inf\limits_{x\in\mathbb{R}^n}\left\{g^*(x):A^\top x=y^*\right\}, & \text{if } y^*\in\operatorname{Im}A^\top\\ +\infty, & \text{if } y^*\notin\operatorname{Im}A^\top.\end{cases}$$

In general, $(g\circ A)^*\neq A^\top g^*$. However, from part (i), we get

$$(g^{**}\circ A)^*=(A^\top g^*)^{**}$$

and, for g convex, we get the formula

$$(\mathrm{cl}\ g\circ A)^*=\mathrm{cl}\,(A^\top g^*)$$

(see R. T. ROCKAFELLAR [**1**, p. 38, Thm. 16.1, p. 142]). □

9.3 Lsc Convexification and Fenchel–Legendre Bitransform

Definition 9.2.
The *lsc convexification* of $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ is defined as follows:

(i) if $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ and there exists a convex lsc function g on $\mathbb{R}^n$ such that $g\le f$ on $\mathbb{R}^n$, define

$$\overline{\mathrm{co}}\,f(x)\overset{\text{def}}{=}\sup_{\substack{g:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}\\ g\text{ convex and lsc on }\mathbb{R}^n\\ g\le f\text{ on }\mathbb{R}^n}} g(x),\quad x\in\mathbb{R}^n; \tag{9.25}$$

(ii) otherwise, we set $\overline{\mathrm{co}}\,f(x)\overset{\text{def}}{=}-\infty$[20] for all $x\in\mathbb{R}^n$.

In case (i), we say that f is *lsc convexifiable*. □

We have the following properties.

Theorem 9.5. *Let* $f:\mathbb{R}^n\to\overline{\mathbb{R}}$.

(i) *If* $f\equiv+\infty$*, then* $f^*\equiv-\infty$ *and*

$$\forall x\in\mathbb{R}^n,\quad \overline{\mathrm{co}}\,f(x)=\sup_{\substack{\ell\in\mathcal{A}\\ \ell\le f\text{ on }\mathbb{R}^n}}\ell(x)=f^{**}(x)=+\infty. \tag{9.26}$$

(ii) *If* $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$, dom $f\neq\varnothing$*, and if there exists* $\ell\in\mathcal{A}$ *such that* $\ell(x)\le f(x)$ *for all* $x\in\mathbb{R}^n$*, then* $f^*:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ *is convex lsc and*

$$\forall x\in\mathbb{R}^n,\quad \overline{\mathrm{co}}\,f(x)=\sup_{\substack{\ell\in\mathcal{A}\\ \ell\le f\text{ on }\mathbb{R}^n}}\ell(x)=f^{**}(x). \tag{9.27}$$

[20] Recall that, by convention, the function identically equal to $-\infty$ is convex lsc.

(iii) *If there exists* $\bar{x} \in \mathbb{R}^n$ *such that* $f(\bar{x}) = -\infty$ *or if for every* $\ell \in \mathcal{A}$, *there exists* $\bar{x} \in \mathbb{R}^n$ *such that* $\ell(\bar{x}) > f(\bar{x})$, *then* $f^*(p) = +\infty$ *for all* $p \in \mathbb{R}^n$ *and*

$$\forall x \in \mathbb{R}^n, \quad f^{**}(x) = -\infty = \overline{\text{co}}\, f(x). \tag{9.28}$$

(iv) *If* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex and* lsc, *then*

$$\forall x \in \mathbb{R}^n, \quad f(x) = \overline{\text{co}}\, f(x) = \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le f \text{ on } \mathbb{R}^n}} \ell(x) = f^{**}(x). \tag{9.29}$$

Corollary 1. (i) *For any function*[21] $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, *we have*

$$\overline{\text{co}}\, f = f^{**} \quad \text{and} \quad f^* = (\overline{\text{co}}\, f)^*. \tag{9.30}$$

(ii) $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is lsc convexifiable if and only if there exists* $\ell \in \mathcal{A}$ *such that* $\ell(x) \le f(x)$ *for all* $x \in \mathbb{R}^n$. *In that case* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *and*

$$\overline{\text{co}}\, f(x) = \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le f \text{ on } \mathbb{R}^n}} \ell(x) = f^{**}(x). \tag{9.31}$$

The *Fenchel–Legendre bitransform* effectively constructs the lsc convexification of a function f. In particular, all functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ that are bounded below are lsc convexifiable. However, the *Fenchel–Legendre transform* does not distinguish between the function f and its lsc convexification $\overline{\text{co}}\, f$.

Proof of Theorem 9.5. (i) If f is identically equal to $+\infty$, then f^* is identically equal to $-\infty$ and f^{**} to $+\infty$. Moreover,

$$\overline{\text{co}}\, f(x) = \sup_{\substack{g:\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \\ g \text{ convex and lsc on } \mathbb{R}^n \\ g \le f \text{ on } \mathbb{R}^n}} g(x) \ge \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le f \text{ on } \mathbb{R}^n}} \ell(x) = f^{**}(x) = +\infty,$$

which yields the identity.

(ii) By Theorem 9.4(ii). It remains to prove the identity. By definition of $\overline{\text{co}}\, f(x)$ and from part (ii) of Theorem 9.5, we know that

$$\overline{\text{co}}\, f(x) = \sup_{\substack{g:\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \\ g \text{ convex and lsc on } \mathbb{R}^n \\ g \le f \text{ on } \mathbb{R}^n}} g(x) \ge \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le f \text{ on } \mathbb{R}^n}} \ell(x) = f^{**}(x)$$

since $\ell \in \mathcal{A}$ is convex lsc. Conversely, from Theorem 9.2, we know that for all convex lsc function g, we have

$$\forall x \in \mathbb{R}^n, \quad g(x) = \overline{\text{co}}\, g(x) = \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le g \text{ on } \mathbb{R}^n}} \ell(x) \tag{9.32}$$

$$\Rightarrow \overline{\text{co}}\, f(x) = \sup_{\substack{g:\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \\ g \text{ convex and lsc on } \mathbb{R}^n \\ g \le f \text{ on } \mathbb{R}^n}} \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le g \text{ on } \mathbb{R}^n}} \ell(x) = \sup_{\substack{\ell \in \mathcal{A} \\ \ell \le f \text{ on } \mathbb{R}^n}} \ell(x). \tag{9.33}$$

[21] Cf. R. T. ROCKAFELLAR [**1**, Thm. 12.2, p. 104] in the case of a convex function f.

(iii) In the first case, if there exists $\bar{x}$ such that $f(\bar{x}) = -\infty$, we have for all p

$$f^*(p) \geq p \cdot \bar{x} - f(\bar{x}) = +\infty.$$

In the second case, for all $\ell \in \mathcal{A}$, there exists $\bar{x} \in \mathbb{R}^n$ such that $\ell(\bar{x}) > f(\bar{x})$. Let $p \in \mathbb{R}^n$. For each $n \to \infty$, consider the function $\ell_n(x) = -n + p \cdot x$. Therefore, there exists x_n such that $\ell_n(x_n) = -n + p \cdot x_n > f(x_n)$. Thence

$$f^*(p) \geq p \cdot x_n - f(x_n) > p \cdot x_n - (-n + p \cdot x_n) = n \to +\infty.$$

(iv) By Theorem 9.3. □

Proof of Corollary 1. (i) The first identity follows from Theorem 9.5. As for the second one, $f^* = (f^*)^{**}$ since f^* is convex lsc and

$$f^* = (f^*)^{**} = (f^{**})^* = (\overline{\text{co}}\, f)^*.$$

(ii) If f is lsc convexifiable, there exists a convex lsc function g such that $g \leq f$. But, by Theorem 9.2, there exists $\ell \in \mathcal{A}$ such that $\ell \leq g$. Therfore, there exists $\ell \in \mathcal{A}$ such that $\ell \leq f$. Conversely, if there exists $\ell \in \mathcal{A}$ such that $\ell \leq f$, then f is lsc convexifiable since ℓ is lsc convex. □

The Fenchel–Legendre bitransform of the indicator function will play an important role later on.

Example 9.8.
Consider the *indicator function* of a nonempty set U:

$$I_U(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } x \in U, \\ +\infty, & \text{if } x \notin U. \end{cases}$$

By definition, the Fenchel–Legendre transform of I_U is

$$(I_U)^*(x^*) = \sup_{x \in \mathbb{R}^n} x^* \cdot x - I_U(x) = \sup_{x \in U} x^* \cdot x,$$

where the function

$$x^* \mapsto \sigma_U(x^*) \stackrel{\text{def}}{=} \sup_{x \in U} x^* \cdot x : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \tag{9.34}$$

is called the *support function* of U.

We now show that $(I_U)^{**} = I_{\overline{\text{co}}\, U}$ and that $\sigma_U = \sigma_{\overline{\text{co}}\, U}$. By definition $\sigma_U \leq \sigma_{\text{co}\, U} \leq \sigma_{\overline{\text{co}}\, U}$. Given x^*, the points $x, y \in U$, and $0 \leq \lambda \leq 1$,

$$\begin{aligned} & x^* \cdot (\lambda x + (1-\lambda) y) = \lambda\, x^* \cdot x + (1-\lambda)\, x^* \cdot y \\ \Rightarrow\ & x^* \cdot (\lambda x + (1-\lambda) y) \leq \lambda \sup_{x \in U} x^* \cdot x + (1-\lambda) \sup_{y \in U} x^* \cdot y = \sigma_U(x^*) \\ \Rightarrow\ & \sup_{z \in \text{co}\, U} x^* \cdot z \leq \sigma_U(x^*) \quad \Rightarrow \sigma_{\text{co}\, U} \leq \sigma_U. \end{aligned}$$

Similarly, given $x \in \overline{\text{co}}\,(U)$, there exists $\{x\} \subset \text{co}\,U$ such that $x_n \to x$ and

$$x^* \cdot x = x^* \cdot x_n + x^* \cdot (x - x_n) \le \sigma_{\text{co}\,U}(x^*) + x^* \cdot (x - x_n).$$

As $n \to \infty$, $x^* \cdot (x - x_n) \to 0$ and

$$\begin{aligned} x^* \cdot x \le \sigma_{\text{co}\,U}(x^*) \quad &\Rightarrow \sigma_{\overline{\text{co}}\,U}(x^*) \le \sigma_{\text{co}\,U}(x^*) \le \sigma_U(x^*) \\ &\Rightarrow \sigma_{\overline{\text{co}}\,U} = \sigma_{\text{co}\,U} = \sigma_U. \end{aligned}$$

Since $\overline{\text{co}}\,U$ is closed and convex by Theorem 7.2, from Lemma 4.4 and Example 4.2, the *indicator function* $I_{\overline{\text{co}}\,U}$ of a closed subset $\overline{\text{co}}\,U$ of $\mathbb{R}^n$ is lsc on $\mathbb{R}^n$ and from Theorem 7.1, it is convex. In view of this, $(I_{\overline{\text{co}}\,U})^{**} = I_{\overline{\text{co}}\,U}$ and

$$(I_U)^* = \sigma_U = \sigma_{\overline{\text{co}}\,U} = (I_{\overline{\text{co}}\,U})^* \quad \Rightarrow (I_U)^{**} = (I_{\overline{\text{co}}\,U})^{**} = I_{\overline{\text{co}}\,U}. \qquad \square$$

Example 9.9 (Fenchel–Legendre transform of d_U)**.**
From the definition of the distance function d_U,

$$\begin{aligned} d_U^*(x^*) &= \sup_{x\in\mathbb{R}^n} x^* \cdot x - d_U(x) = \sup_{x\in\mathbb{R}^n} [x^* \cdot x - \inf_{p\in U} \|x-p\|] \\ &= \sup_{x\in\mathbb{R}^n} \sup_{p\in U} x^* \cdot x - \|x-p\| = \sup_{p\in U} \sup_{x\in\mathbb{R}^n} x^* \cdot x - \|x-p\| \\ &= \sup_{p\in U} \sup_{x\in\mathbb{R}^n} x^* \cdot p + x^* \cdot (x-p) - \|x-p\| \\ &= \sup_{p\in U} x^* \cdot p + \sup_{x\in\mathbb{R}^n} [x^* \cdot (x-p) - \|x-p\|] \\ &= \sup_{p\in U} x^* \cdot p + \sup_{x\in\mathbb{R}^n} [x^* \cdot x - \|x\|] = \sigma_U(x^*) + I_{\overline{B_1(0)}}(x^*). \end{aligned}$$

The functions d_U and $I_{\overline{U}}$ are both zero on $\overline{U}$ and the Fenchel–Legendre transforms coincide on $\overline{B_1(0)}$. In view of the previous example, $\sigma_U = \sigma_{\overline{\text{co}}\,U}$ and hence

$$\begin{aligned} d_U^*(x^*) = \sigma_U(x^*) + I_{\overline{B_1(0)}}(x^*) &= \sigma_{\overline{\text{co}}\,U}(x^*) + I_{\overline{B_1(0)}}(x^*) = d^*_{\overline{\text{co}}\,U}(x^*) \\ \Rightarrow d_U^* = d^*_{\overline{\text{co}}\,U} \quad &\text{and} \quad d_U^{**} = d_{\overline{\text{co}}\,U}. \end{aligned}$$

So we have the same pattern as for the indicator function. The bitransform yields the closed convex hull of U. $\square$

9.4 Infima of the Objective Function and of Its lsc Convexified Function

In general, an objective function that has a global infimum can have several local infima, while a convex function has only one infimum (that is both local and global). If the function is *convexified* and if we consider the infimum of the convexified function, we readily get the global infimum of the initial function. To do that, first associate with $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ and $U \subset \mathbb{R}^n$ the function

$$f_U(x) = \left\{ \begin{array}{ll} f(x), & \text{if } x \in U \\ +\infty, & \text{if } x \in \mathbb{R}^n \backslash U \end{array} \right\} = f(x) + I_U(x) \tag{9.35}$$

as defined in (3.5). From Theorem 3.2 we have

$$\inf f(U) = \inf f_U(\mathbb{R}^n).$$

Then, we *lsc convexify* the function f_U.

Theorem 9.6. *For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and U, $\varnothing \neq U \subset \mathbb{R}^n$,*

$$\inf f(U) = \inf f_U(\mathbb{R}^n) = \inf (f_U)^{**}(\mathbb{R}^n) = -f_U^*(0). \tag{9.36}$$

The function $L(x,x^*) = x^* \cdot x - f_U^*(x^*)$ is lsc convex in x, usc concave in x^*, and

$$\sup_{x^* \in \mathbb{R}^n} \inf_{x \in \mathbb{R}^n} \left[x^* \cdot x - f_U^*(x^*)\right] = -f_U^*(0) = \inf_{x \in \mathbb{R}^n} \sup_{x^* \in \mathbb{R}^n} \left[x^* \cdot x - f_U^*(x^*)\right]. \tag{9.37}$$

Proof. We already know that $\inf f(U) = \inf f_U(\mathbb{R}^n) = -f_U^*(0)$ and, by Corollary 1 to Theorem 9.5, that $\overline{\text{co}}\, f_U = (f_U)^{**}$. It remains to prove that

$$\inf f_U(\mathbb{R}^n) = \inf \overline{\text{co}}\, f_U(\mathbb{R}^n) \tag{9.38}$$

to establish the chain of equalities (9.36). By definition, for all $x \in \mathbb{R}^n$, $\overline{\text{co}}\, f_U(x) \leq f_U(x)$ and $\inf \overline{\text{co}}\, f_U(\mathbb{R}^n) \leq \inf f_U(\mathbb{R}^n)$. If $\inf f_U(\mathbb{R}^n) = +\infty$, $f_U(x) = +\infty$ for all x, $\overline{\text{co}}\, f_U(x) = \text{co}\, f_U(x) = +\infty$ for all x, and (9.38) is verified. If $-\infty < m = \inf f_U(\mathbb{R}^n) < +\infty$, then for all $x \in \mathbb{R}^n$, $m = \inf f_U(\mathbb{R}^n) \leq f_U(x)$ and $f_U : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. As the constant function m belongs to $\mathcal{A}$, $m \leq \overline{\text{co}}\, f_U(x)$ for all x by Theorem 9.5(ii) and we get $\inf f_U(\mathbb{R}^n) \leq \inf \overline{\text{co}}\, f_U(x)$. By combining this inequality with the one from the beginning, we get $\inf f_U(\mathbb{R}^n) \leq \inf \overline{\text{co}}\, f_U(\mathbb{R}^n) \leq \inf f_U(\mathbb{R}^n)$, and (9.38) is verified. If $\inf f_U(\mathbb{R}^n) = -\infty$, then $\inf \overline{\text{co}}\, f_U(\mathbb{R}^n) \leq \inf f_U(\mathbb{R}^n) = -\infty$, and (9.38) is verified.

The properties of $L(x,x^*)$ are obvious. By definition of $(f_U)^{**}$,

$$-f_U^*(0) = \inf (f_U)^{**}(\mathbb{R}^n) = \inf_{x \in \mathbb{R}^n} \sup_{x^* \in \mathbb{R}^n} \left[x^* \cdot x - f_U^*(x^*)\right]$$

which yields the second half of (9.37). For the sup inf we have

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \left[x^* \cdot x - f_U^*(x^*)\right] &= \begin{cases} -\infty, & \text{if } x^* \neq 0, \\ -f_U^*(0), & \text{if } x^* = 0 \end{cases} \\ \Rightarrow \sup_{x^* \in \mathbb{R}^n} \inf_{x \in \mathbb{R}^n} \left[x^* \cdot x - f_U^*(x^*)\right] &= -f_U^*(0) \end{aligned}$$

which yields the first half of (9.37). □

9.5 Primal and Dual Problems and Fenchel Duality Theorem

We have seen in Theorem 9.6 that for any $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and U, $\varnothing \neq U \subset \mathbb{R}^n$,

$$\inf f(U) = \inf f_U(\mathbb{R}^n) = -f_U^*(0).$$

The function f_U is a special case of functions $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ for which

$$\inf f(\mathbb{R}^n) = -f^*(0) \tag{9.39}$$

by definition of the transform. If f_U is written in the form

$$f_U = f + I_U, \quad I_U(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } x \in U \\ +\infty, & \text{if } x \notin U, \end{cases} \tag{9.40}$$

where I_U is the *indicator function* of U, then

$$\inf f_U(\mathbb{R}^n) = \inf(f + I_U)(\mathbb{R}^n) = -(f + I_U)^*(0),$$

but it would be nice to express the right-hand side as a function of f^* and I_U^* instead of $(f + I_U)^*$.

Consider $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\operatorname{dom} f \neq \varnothing$, and a second function $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\operatorname{dom} g \neq \varnothing$. Under these assumptions $f^*, g^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and the sums $f(x) + g(x)$ and $g^*(x^*) + f^*(-x^*)$ are well defined. In particular,

$$\begin{aligned} -f(x) - g(x) = \big[x^* \cdot x - g(x)\big] + \big[-x^* \cdot x - f(x)\big] &\leq g^*(x^*) + f^*(-x^*) \\ \Rightarrow (f+g)^*(0) &\leq g^*(x^*) + f^*(-x^*) \\ \Rightarrow \inf_{x \in \mathbb{R}^n} [f(x) + g(x)] = -(f+g)^*(0) &\geq -g^*(x^*) - f^*(-x^*) \\ \Rightarrow \inf_{x \in \mathbb{R}^n} [f(x) + g(x)] &\geq \sup_{x^* \in \mathbb{R}^n} \big[-g^*(x^*) - f^*(-x^*)\big] \end{aligned} \tag{9.41}$$

by using the fact that $-(f+g)^*(0) = \inf_{x \in \mathbb{R}^n} [f(x) + g(x)]$ from Theorem 9.6.

Theorem 9.7 (Fenchel primal and dual problems). *Let* $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be such that* $\operatorname{dom} f \neq \varnothing$ *and* $\operatorname{dom} g \neq \varnothing$.[22]

(i) *The so-called* primal *and* dual *problems are related as follows:*

$$\underbrace{\inf_{x \in \mathbb{R}^n} [f(x) + g(x)]}_{\textit{primal problem}} \geq \underbrace{\sup_{x^* \in \mathbb{R}^n} \big[-g^*(x^*) - f^*(-x^*)\big]}_{\textit{dual problem}}. \tag{9.42}$$

More precisely,

$$\inf_{x \in \mathbb{R}^n} (f+g)^{**}(x) = \inf_{x \in \mathbb{R}^n} [f(x) + g(x)]$$

$$\inf_{x \in \mathbb{R}^n} [f(x) + g(x)] \geq \inf_{x \in \mathbb{R}^n} \big[g^{**}(x) + f^{**}(x)\big] \geq \sup_{x^* \in \mathbb{R}^n} \big[-g^*(x^*) - f^*(-x^*)\big].$$

(ii) *If equality holds in* (9.42)*, then*

$$\inf_{x \in \mathbb{R}^n} (f+g)^{**}(x) = \inf_{x \in \mathbb{R}^n} [f(x) + g(x)] = \inf_{x \in \mathbb{R}^n} \big[g^{**}(x) + f^{**}(x)\big]. \tag{9.43}$$

The left-hand side of inequality (9.42) is called the *primal value* and the right-hand side the *dual value*. When the inequality is strict, we say that there is a *duality gap*. What is important to retain is that for the *primal problem*,

$$\inf_{x \in \mathbb{R}^n} [f(x) + g(x)], \tag{9.44}$$

[22] In the terminology of Definition 3.2, f and g are proper for the infimum.

the sum, $f+g$, of the two functions is not necessarily lsc convex, while for the dual problem,

$$\sup_{x^*\in\mathbb{R}^n}\left[-g^*(x^*)-f^*(-x^*)\right]=-\inf_{x^*\in\mathbb{R}^n}\big[\underbrace{g^*(x^*)+f^*(-x^*)}_{\text{lsc convex}}\big], \tag{9.45}$$

the sum $g^*(x^*)+f^*(-x^*)$ is lsc convex. So, if g^* and f^* are easy to compute and if the primal value is equal to the dual value, the infimum of the primal problem can be obtained from the infimum of the sum of two lsc convex functions.

Proof. (i) From the previous discussion but applied to g^{**} instead of g,

$$\inf_{x\in\mathbb{R}^n}\left[g^{**}(x)+f^{**}(x)\right]\geq\sup_{x^*\in\mathbb{R}^n}\left[-g^*(x^*)-f^*(-x^*)\right],$$

since $g^*=(g^{**})^*$. Moreover, since $g^{**}\leq g$ and $f^{**}\leq f$,

$$\inf_{x\in\mathbb{R}^n}(f+g)^{**}(x)=\inf_{x\in\mathbb{R}^n}[f(x)+g(x)]\geq\inf_{x\in\mathbb{R}^n}\left[g^{**}(x)+f^{**}(x)\right]. \tag{9.46}$$

(ii) Since $g^*=(g^{**})^*$ and $g^*=(g^{**})^*$, from part (i) applied to g^{**} and g^{**},

$$\begin{aligned}\sup_{x^*\in\mathbb{R}^n}\left[-g^*(x^*)-f^*(-x^*)\right]&=\sup_{x^*\in\mathbb{R}^n}\left[-(g^{**})^*(x^*)-(f^{**})^*(-x^*)\right]\\&\leq\inf_{x\in\mathbb{R}^n}\left[g^{**}(x)+f^{**}(x)\right].\end{aligned}$$

When equality holds in (9.42), combining this last inequality with inequality (9.46), we get the identities (9.43). □

Since the Fenchel–Legendre transform does not distinguish the function g from its lsc convexified g^{**}, the dual problem will at best give solutions of the following lsc convex problem:

$$\inf_{x\in\mathbb{R}^n}\left[g^{**}(x)+f^{**}(x)\right], \tag{9.47}$$

which stands between the primal and dual problems as indicated in part (i). When equality holds in (9.42), they are all equal.

For instance, choose $g\equiv 0$ for which $g^*(x^*)=0$ if $x^*=0$ and $+\infty$ if $x^*\neq 0$. This yields

$$-f^*(-x^*)-g^*(x^*)=\begin{cases}-f^*(0), & \text{if } x^*=0\\ -\infty, & \text{if } x^*\neq 0\end{cases}$$
$$\Rightarrow\sup_{x^*\in\mathbb{R}^n}\left[-f^*(-x^*)-g^*(x^*)\right]=-f^*(0)=\inf_{x\in\mathbb{R}^n}f(x)=\inf_{x\in\mathbb{R}^n}[f(x)+g(x)],$$

since, by definition of f^*, $\inf_{x\in\mathbb{R}^n}f(x)=-f^*(0)$. In that case, we have equality in (9.41) with the function $g\equiv 0$ for which g^* is not identically zero! The function identically zero is nothing but the indicator function $I_{\mathbb{R}^n}$ of $\mathbb{R}^n$.

The pattern is now clear. Given $g=I_U$, the *indicator function* of a nonempty closed convex set U, and an objective function $f:\mathbb{R}^n\to\mathbb{R}$ to be minimized over U, we have

$$\inf f(U)=\inf_{x\in\mathbb{R}^n}[f(x)+I_U(x)].$$

The Fenchel–Legendre transform of I_U is the *support function* of U:

$$\sigma_U(x^*) \stackrel{\text{def}}{=} I_U^*(x^*) = \sup_{x\in\mathbb{R}^n} \left[x^* \cdot x - I_U(x)\right] = \sup_{x\in U} x^* \cdot x \tag{9.48}$$

that yields

$$\inf_{x\in\mathbb{R}^n} [f(x)+I_U(x)] \le \sup_{x^*\in\mathbb{R}^n} \left[-f^*(-x^*)-\sigma_U(x^*)\right] = -\inf_{x^*\in\mathbb{R}^n} \left[f^*(-x^*)+\sigma_U(x^*)\right].$$

The systematic study of functions of the form $f+g$ for f and g convex proper (for the infimum) leads to Fenchel duality theorem that provides sufficient conditions under which equality holds in (9.42). It will clearly be necessary that

$$\text{either} \quad \operatorname{dom} f \cap \operatorname{dom} g \neq \varnothing \quad \text{or} \quad \operatorname{dom} g^* \cap (-\operatorname{dom} f^*) \neq \varnothing \tag{9.49}$$

to get equality since $\operatorname{dom} f \cap \operatorname{dom} g = \varnothing$ and $\operatorname{dom} g^* \cap (-\operatorname{dom} f^*) = \varnothing$ imply

$$+\infty = \inf_{x\in\mathbb{R}^n} [f(x)+g(x)] > -\inf_{x^*\in\mathbb{R}^n} \left[g^*(x^*)+f^*(-x^*)\right] = -\infty.$$

But condition (9.49) can be strengthened to get sufficient conditions.

Theorem 9.8. [23] *Let* $f,g:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ *be convex with domains* $\operatorname{dom} f \neq \varnothing$ *and* $\operatorname{dom} g \neq \varnothing$.[24] *Then*

$$\inf_{x\in\mathbb{R}^n} [f(x)+g(x)] = -\inf_{x^*\in\mathbb{R}^n} \left[g^*(x^*)+f^*(-x^*)\right] \tag{9.50}$$

if

(a) *either* $\operatorname{ri}(\operatorname{dom} f)\cap \operatorname{ri}(\operatorname{dom} g) \neq \varnothing$,

(b) *or* f *and* g *are lsc and* $\operatorname{ri}(\operatorname{dom} g^*)\cap \operatorname{ri}(-\operatorname{dom} f^*) \neq \varnothing$,

where $\operatorname{ri} X$ *is the* relative interior *of a subset* X *of* $\mathbb{R}^n$.

So, for a convex function $g:\mathbb{R}^n \to \mathbb{R}$ ($\operatorname{dom} f = \mathbb{R}^n$) and a convex function $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ such that $\operatorname{ri}(\operatorname{dom} f) \neq \varnothing$, condition (a) is verified. For instance, f can be the indicator function of a nonempty convex subset U of $\mathbb{R}^n$.

A more *symmetrical-looking* formula can be obtained by introducing the function $\gamma(x) = -f(x)$ for which $\operatorname{dom} \gamma = \operatorname{dom} f$ rather than f but at the price of an additional definition:

$$\gamma_*(x^*) \stackrel{\text{def}}{=} \inf_{x\in\mathbb{R}^n} \left[x^* \cdot x - \gamma(x)\right] \tag{9.51}$$

not to be confused with γ^*. It is easy to check that

$$f^*(-x^*) = \sup_{x\in\mathbb{R}^n} \left[-x^* \cdot x - f(x)\right] = \sup_{x\in\mathbb{R}^n} \left[-x^* \cdot x + \gamma(x)\right] = -\inf_{x\in\mathbb{R}^n} \left[x^* \cdot x - \gamma(x)\right]$$
$$\Rightarrow \gamma_*(x^*) = -f^*(-x^*) \quad \text{and} \quad \operatorname{dom} \gamma_* = -\operatorname{dom} f^*.$$

[23] Cf. R. T. Rockafellar [**1**, Thm. 31,1, p. 327], the original papers of W. Fenchel [**1**, **2**, **4**], and the numerous results which followed.

[24] f and g are proper for the infimum.

With the initial assumptions on g and $f = -\gamma$, we get

$$\inf_{x\in\mathbb{R}^n} [g(x) - \gamma(x)] \geq \sup_{x^*\in\mathbb{R}^n} \left[\gamma_*(x^*) - g^*(x^*)\right]. \tag{9.52}$$

However, this neither adds anything to the problem nor helps in its resolution. It is more a matter of aesthetics and formalism.

Theorem 9.9. *Let $g : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be convex proper for the infimum and $\gamma : \mathbb{R}^n \to \mathbb{R}\cup\{-\infty\}$ be concave proper for the supremum. Then*

$$\inf_{x\in\mathbb{R}^n} [g(x) - \gamma(x)] = \sup_{x^*\in\mathbb{R}^n} \left[\gamma_*(x^*) - g^*(x^*)\right] = - \inf_{x^*\in\mathbb{R}^n} \left[g^*(x^*) - \gamma_*(x^*)\right] \tag{9.53}$$

if

(a) *either* $\operatorname{ri}(\operatorname{dom} g) \cap \operatorname{ri}(\operatorname{dom} \gamma) \neq \varnothing$,

(b) *or g is lsc, γ is usc, and* $\operatorname{ri}(\operatorname{dom} g^*) \cap \operatorname{ri}(\operatorname{dom} \gamma_*) \neq \varnothing$,

where $\operatorname{ri} X$ *is the* relative interior *of a subset X of* $\mathbb{R}^n$.

10 Exercises

Exercise 10.1.
Let $f, g : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be two functions that are lsc at a point $a \in \mathbb{R}^n$. Show that the following functions are lsc at a:

$$\text{(i)} \quad (f+g)(x) \stackrel{\text{def}}{=} f(x) + g(x) \quad \text{and} \quad \text{(ii)} \quad \forall \alpha \geq 0, \quad (\alpha f)(x) \stackrel{\text{def}}{=} \alpha f(x).$$ □

Exercise 10.2.
Given $f : \mathbb{R}^n \to \mathbb{R}$, prove that f is continuous at a if and only if f is lsc and usc at a. □

Exercise 10.3.
Let $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ and $g : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be two lsc functions at a. Prove that the function

$$(f \wedge g)(x) \stackrel{\text{def}}{=} \inf\{f(x), g(x)\}$$

is lsc at a. □

Exercise 10.4.
Let $\{f_\alpha\}_{\alpha\in A}$, $f_\alpha : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, be a family of functions where the *index set* A is possibly infinite. Its *upper envelope* is defined as

$$f(x) \stackrel{\text{def}}{=} \sup_{\alpha\in A} f_\alpha(x), \quad x \in \mathbb{R}^n.$$

If f_α is lsc at $x_0 \in \mathbb{R}^n$ for each $\alpha \in A$, prove that f is lsc at x_0. □

Exercise 10.5.
Consider the following function:

$$f(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } x \text{ is rational} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

Prove that f is lsc and not usc at each rational point; prove that f is usc and not lsc at each irrational point. □

Exercise 10.6.
Consider an lsc function $f : \mathbb{R}^n \to \mathbb{R}$. Prove or give a counterexample in each of the following cases.

(i) Is the function $f(x)^2$ lsc on $\mathbb{R}^n$?

(ii) Is the function $e^{f(x)}$ lsc on $\mathbb{R}^n$?

(iii) Is the function $e^{-f(x)}$ usc on $\mathbb{R}^n$?

(iv) Is the function (here $n = 1$)

$$F(x) \stackrel{\text{def}}{=} \begin{cases} f(x), & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

lsc on $\mathbb{R}$? □

Exercise 10.7.
Consider the function

$$f(x) = |x|^{17} - x + 3x^2, \quad x \in \mathbb{R}.$$

(i) Do we have existence of minimizers of f in $\mathbb{R}$? If so, provide the elements of the proof.

(ii) Assuming that there are minimizers of f on $\mathbb{R}$, prove that they are positive or zero ($x \geq 0$). □

Exercise 10.8.
Let A be an $n \times n$ matrix and b a vector in $\mathbb{R}^n$. Define the function

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x, \quad x \in \mathbb{R}^n.$$

(i) Give necessary and sufficient conditions on A and b under which f is convex on $\mathbb{R}^n$.

(ii) Give necessary and sufficient conditions on A and b under which f has a unique minimizer in $\mathbb{R}^n$.

(iii) Give necessary and sufficient conditions on A and b under which f has a minimizer in $\mathbb{R}^n$.

(iv) Are the functions f associated with the following matrices A and vectors b convex on $\mathbb{R}^2$:

$$A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, b = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}?$$ □

Exercise 10.9.
Let A be a symmetric $n \times n$ positive definite matrix and b a vector in $\mathbb{R}^n$. Prove that the function

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x, \quad x \in \mathbb{R}^n$$

is strongly convex on $\mathbb{R}^n$. $\square$

Exercise 10.10.
Find and justify the Fenchel–Legendre transforms of the following functions: (i) $f(x) = a \cdot x + b$,

$$f^*(x^*) = \begin{cases} -b, & \text{if } x^* = a \\ +\infty, & \text{if } x^* \neq a \end{cases} \quad \text{and} \quad f^{**}(x) = a \cdot x + b = f(x);$$

and (ii) $f(x) = \|x\|$,

$$f^*(x^*) = \begin{cases} 0, & \text{if } \|x^*\| \leq 1 \\ +\infty, & \text{if } \|x^*\| > 1 \end{cases} \quad \text{and} \quad f^{**}(x) = \|x\| = f(x). \quad \square$$

Exercise 10.11.
Consider the function

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x,$$

where A is an $n \times n$ symmetric matrix and $b \in \mathbb{R}^n$. Find and justify the Fenchel–Legendre transform when A is (i) positive definite and (ii) positive semidefinite. $\square$

Exercise 10.12.
The *infimum convolution* of two functions $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined as

$$(f \,\square\, g)(x) \stackrel{\text{def}}{=} \inf_{y \in \mathbb{R}^n} [f(x-y) + g(y)]. \tag{10.1}$$

(i) Find the Fenchel–Legendre transform of $f \,\square\, g$.

(ii) Prove that $f \,\square\, g$ is convex if f and g are convex.

(iii) Compute $f \,\square\, g$ for $f(x) = \|x\|_{\mathbb{R}^n}$ and $g = I_U$, the indicator function of the nonempty set $U \subset \mathbb{R}^n$. $\square$

Exercise 10.13.
(i) Let $B : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, and define the *marginal value* function

$$(Bf)(y) \stackrel{\text{def}}{=} \begin{cases} \inf\limits_{x \in \mathbb{R}^n} \{f(x) : Bx = y\}, & \text{if } y \in \operatorname{Im} B \\ +\infty, & \text{if } y \notin \operatorname{Im} B. \end{cases} \tag{10.2}$$

Show that $(Bf)^* = f^* \circ B^\top$ and that, if f is convex, then Bf and $(Bf)^*$ are convex.

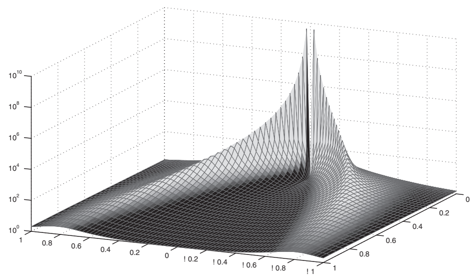
(ii) Let $A : \mathbb{R}^m \to \mathbb{R}^n$ be a linear map and $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Show that

$$(g^{**} \circ A)^* = (A^\top g^*)^{**}$$

and that, for g convex, we get the formula

$$(\mathrm{cl}\ g \circ A)^* = \mathrm{cl}\ (A^\top g^*)$$

(see R. T. ROCKAFELLAR [**1**, p. 38, Thm. 16.1, p. 142]). □



Chapter 3
Semidifferentiability, Differentiability, Continuity, and Convexities

1 Introduction

According to some historians the differential or infinitesimal calculus has been implicitly present very early. For instance, the mathematician-astronomer Aryabhata (476–550 CE) in 499 CE used a notion of infinitesimals and expressed an astronomical problem in the form of a basic differential equation.[1] For other historians the differential calculus was invented in the 17th century.

Accepting this second point of view, the first idea of the differential calculus and the rule for computing the extrema of a function go back to Pierre de Fermat[2] in 1638. He developed a method called *maximis et minimis* for determining maxima, minima, and tangents to various curves that was equivalent to differentiation.[3] Ideas leading up to the notions of function, derivative, and integral were developed throughout the 17th century. It is generally acccepted that the notion of derivative is due to Leibniz[4] and Newton.[5] Fermat's rule for the extremum of a function is then de facto generalized in the form $f'(x) = 0$. It is used in the proof of the theorem of Rolle[6] in 1691 that leads to the rule of l'Hôpital[7] in 1696.

---

[1] George G. Joseph, *The Crest of the Peacock*, Princeton University Press (2000), pp. 298–300.

[2] Pierre de Fermat (1601–1665).

[3] First written in a letter to Mersenne (who was corresponding with numerous scientists at the time and was seeing at the diffusion of new results) in 1638, the first printed version of the method can be found in the fifth volume of *Supplementum Cursus Mathematici* (1642) written by Herigone, and it is only in 1679 that it appears in *Varia opera mathematica* under the title *Methodus ad disquirendam Maximam et Minimam* followed by *De tangentibus linearum curvarum*.

[4] Gottfried Wilhelm Leibniz (1646–1716).

[5] Sir Isaac Newton (1643–1728).

[6] Michel Rolle (1652–1719).

[7] Guillaume François Antoine de l'Hôpital (1661–1704).

Publication of Newton's main treatises[8] took many years, whereas Leibniz published first (Nova methodus,[9] 1684) and the whole subject was subsequently marred by a priority dispute between the two inventors of the calculus.

> "In early calculus the use of infinitesimal quantities was thought unrigorous, and was fiercely criticized by a number of authors, most notably Michel Rolle and Bishop Berkeley. Bishop Berkeley famously described infinitesimals as the ghosts of departed quantities in his book *The Analyst* in 1734. Several mathematicians, including Maclaurin, attempted to prove the soundness of using infinitesimals, but it would be 150 years later, due to the work of Cauchy and Weierstrass, where a means was finally found to avoid mere "notions" of infinitely small quantities, that the foundations of differential and integral calculus were made firm. In Cauchy's writing, we find a versatile spectrum of foundational approaches, including a definition of continuity in terms of infinitesimals, and a (somewhat imprecise) prototype of an (ε, δ)-definition of limit in the definition of differentiation. In his work Weierstrass formalized the concept of limit and eliminated infinitesimals. Following the work of Weierstrass, it eventually became common to base calculus on limits instead of infinitesimal quantities. This approach formalized by Weierstrass came to be known as the standard calculus. Informally, the name "infinitesimal calculus" became commonly used to refer to Weierstrass' approach" (see Wikipedia, Infinitesimal Calculus).

To characterize the extremum of a smooth function of several variables, the classical notion is the one of *total differential* generalized by Fréchet to spaces of functions (hence of infinite dimension). However, important classes of functions in optimization are not *differentiable* in that sense. In fact, the first natural notion for the characterization of an extremum is rather the *directional derivative* at the point where the extremum is achieved. In this chapter, we shall go back to the older notions of *first variation* and of *differential* that will be relaxed to weaker notions of *semidifferentials*.

Nevertheless, existence of directional derivatives or semidifferentials does not guarantee existence of the total differential or of the simple gradient (linearity with respect to the direction). We shall see that the Gateaux semidifferential (or first variation) is not sufficient to preserve the basic properties of the total differential such as the continuity of the function and the chain rule for the composition of functions. It is the stronger semidifferential in the sense of Hadamard that will make it possible to enlarge the family of classically differentiable functions while preserving those two properties and adding new functional operations to the calculus that will become a *semidifferential calculus*. For instance, the lower and upper envelopes of finite (and in some cases infinite) families of Hadamard semidifferential functions are Hadamard semidifferentiable. It will also be shown that continuous convex functions that play a central role in optimization (as we have seen in Chapter 2) are Hadamard semidifferentiable.

[8]The *Method of Fluxions* completed in 1671 and published in 1736 and *Philosophiæ Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy), often called the *Principia* (Principles), 1687 and 1726 (third edition).

[9]*Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractas nec irrationales quantitates moratur, et singulare pro illis calculi genus (A new method for maxima and minima and their tangents, that are not limited to fractional or irrational expressions, and a remarkable type of calculus for these))*, in Acta Eruditorum, 1684, a journal created in Leipzig two years earlier.

This chapter focuses on the properties of differentiable and semidifferentiable real-valued and vector-valued functions and their associated semidifferential calculus. Section 2 revisits real-valued functions of a single real variable. Sections 2.2 to 2.4 review classical results.

Section 3 deals with real-valued functions of several real variables. Notions of semidifferentials and Hadamard semidifferentials are introduced going back to the preoccupations of the time through the papers of J. HADAMARD [**2**] in 1923 and of M. FRÉCHET [**3**] in 1937. They will be used to fully characterize the classical notions of differentiability with which will be associated the names of Gateaux, Hadamard, and Fréchet.

Section 3.5 concentrates on Lipschitz continuous functions for which the existence of the semidifferential implies the existence of the stronger Hadamard semidifferential. However, this is not entirely satisfactory. For instance, the distance function d_U in Example 5.2 of Chapter 2 is uniformly Lipschitzian in $\mathbb{R}^n$. It is Hadamard semidifferentiable when U is convex, but this semidifferentiability might not exist at some points for a lousy set U! Yet, since d_U is Lipschitzian, its differential quotient is bounded and the liminf and limsup are both finite numbers. Without for all that going deep to the clouds and dropping the notion of semidifferential, we are led to introducing the notions of lower and upper semidifferentials by replacing the limit by the liminf and the limsup in the definitions. This idea makes it possible to relax other notions of differentiability for Lipschitzian functions. Thus, the upper semidifferential developed by Clarke[10] in 1973 under the name *generalized directional derivative* relaxes the older notion of strict differentiability. The lower and upper semidifferentials in the sense of Clarke are added to the menagerie of notions of semidifferentials that will be summarized in section 6 and compared in Chapter 4 in the context of the necessary optimality condition. In general, the classical chain rule will not hold for the composition of two lower or upper semidifferentials, but weaker geometric forms of the chain rule will be available. The section is completed with a review of functions of class $C^{(1)}$ and $C^{(p)}$.

Section 4 first deals with the characterization of convex functions that are directionally differentiable. It is followed by a characterization via the semidifferentials. It is also shown that a convex function is Hadamard semidifferentiable and continuous at interior points of its domain. At boundary points, the Hadamard semidifferential is replaced by the lower Hadamard semidifferential. The section is completed by introducing the semiconvex and semiconcave functions that are locally Lipschitzian and Hadamard semidifferentiable at each point of the interior of their domain.

Section 5 tackles the important issue of the semidifferentiability of an infimum or a supremum that depends on a set of parameters. One such problem is the optimal design of the profile of a column against buckling which is a special case of dependence of the least or greatest eigenvalue of a symmetric matrix on parameters. Another example is the infimum or supremum of the *compliance*[11] in mechanics with respect to the design parameters defining the shape of the mechanical part.

2 Real-Valued Functions of a Real Variable

Prior to study real-valued functions of several real variables, consider functions of a single real variable (see Figure 3.1).

---

[10]Frank Herbert Clarke (1948–). See F. H. CLARKE [**1**, **2**].

[11]The compliance in mechanics corresponds to the work of the applied forces.

Definition 2.1.
Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function of a real variable.

(i) f is *differentiable from the right* or *right differentiable* in $x \in \mathbb{R}$ if the limit

$$\lim_{t \searrow 0^+} \frac{f(x+t)-f(x)}{t} \tag{2.1}$$

exists and is finite, where $\lim_{t\searrow 0^+}$ means that t goes to 0 by strictly positive values. That limit will be denoted $df(x;+1)$ or $df(x;1)$.

f is *continuous from the right* or *right continuous* at $x \in \mathbb{R}$ if

$$\lim_{t \searrow 0^+} f(x+t) = f(x) \tag{2.2}$$

(ii) f is *differentiable from the left*[12] or *left differentiable* in $x \in \mathbb{R}$ if the limit

$$\lim_{t \searrow 0^+} \frac{f(x-t)-f(x)}{t} \tag{2.3}$$

exists and is finite. That limit will be denoted $df(x;-1)$.

f is *continuous from the left* or *left continuous* at $x \in \mathbb{R}$ if

$$\lim_{t \searrow 0^+} f(x-t) = f(x) \tag{2.4}$$

(iii) f is *differentiable* at $x \in \mathbb{R}$ if the limit

$$\lim_{t \to 0} \frac{f(x+t)-f(x)}{t} \tag{2.5}$$

exists and is finite, where $\lim_{t\to 0}$ means that $t \in \mathbb{R}$ goes to 0. It will be denoted $f'(x)$, $df(x)/dx$, or $f^{(1)}(x)$.

The notions of right and left derivatives at x are what we shall later call semidifferentials at x in the respective directions $+1$ and -1. The right and left differentials are special cases of Dini's[13] differentials that we shall see later in Definition 3.9. □

Remark 2.1.
If f is right differentiable at $x \in \mathbb{R}$, we have the *positive homogeneity*:

$$\forall \alpha \geq 0, \quad df(x;\alpha) \overset{\text{def}}{=} \lim_{t \searrow 0^+} \frac{f(x+\alpha t)-f(x)}{t} = \alpha\, df(x;+1), \quad df(x;0)=0; \tag{2.6}$$

similarly, if f is left differentiable at $x \in \mathbb{R}$, we have the *positive homogeneity*:

$$\forall \alpha \geq 0, \quad df(x;-\alpha) \overset{\text{def}}{=} \lim_{t \searrow 0^+} \frac{f(x-\alpha t)-f(x)}{t} = \alpha\, df(x;-1), \quad df(x;0)=0. \tag{2.7}$$

[12] Technically speaking, it is the derivative defined in (i) in the direction -1.

[13] Ulisse Dini (1845–1918), U. Dini [**1**].

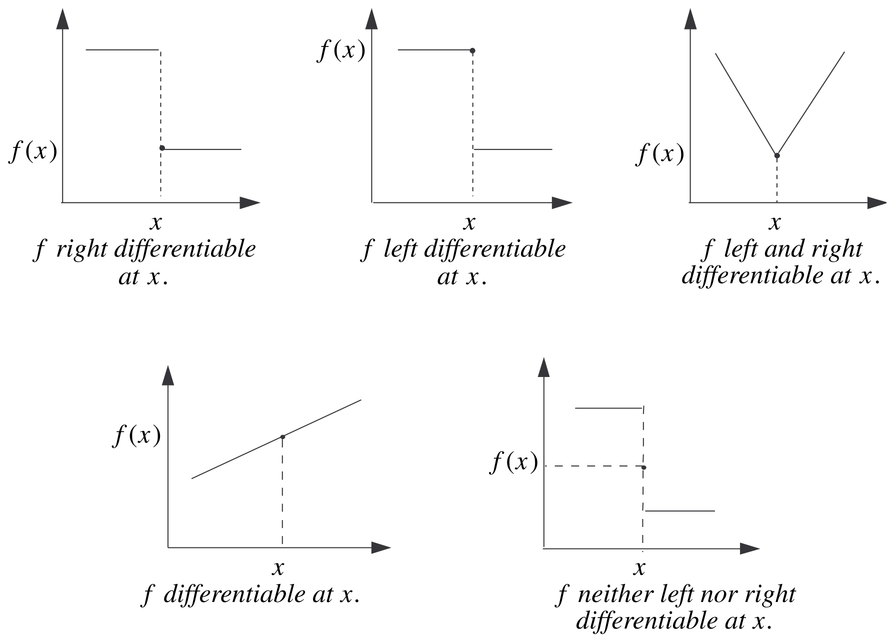


Figure 3.1. *Example of right and left differentiability.*

It is readily checked that if f is differentiable at x, then

$$f'(x) = df(x;+1) = -df(x;-1) \tag{2.8}$$
$$\Rightarrow \forall \alpha \geq 0, \quad df(x;-\alpha) \overset{\text{def}}{=} \lim_{t\to 0} \frac{f(x-\alpha t)-f(x)}{t} = \alpha\, df(x;-1) = -\alpha\, f'(x)$$

and we have the *homogeneity*

$$\forall \alpha \in \mathbb{R}, \quad df(x;\alpha) \overset{\text{def}}{=} \lim_{t\to 0} \frac{f(x+\alpha t)-f(x)}{t} = \alpha\, df(x;+1) = \alpha\, f'(x). \tag{2.9}$$

Therefore, if f is differentiable at $x \in \mathbb{R}$, the positive homogeneity, combined with the condition

$$df(x;-1) = -df(x;1), \tag{2.10}$$

implies that the map

$$v \mapsto df(x;v) \overset{\text{def}}{=} \lim_{t \searrow 0^+} \frac{f(x+tv)-f(x)}{t} : \mathbb{R} \to \mathbb{R} \tag{2.11}$$

is homogeneous and hence linear and continuous:

$$\forall \alpha, \beta \in \mathbb{R} \text{ and } \forall v, w \in \mathbb{R}, \quad df(x;\alpha v + \beta w) = \alpha\, df(x;v) + \beta\, df(x;w). \qquad \square$$

2.1 Continuity and Differentiability

Theorem 2.1. *If f is right* (resp., *left*) *differentiable at a point $x \in \mathbb{R}$, then f is right* (resp., *left*) *continuous at x. If f is differentiable at x, then f is continuous at x.*

Proof. It is sufficient to prove the right continuity from the right differentiability. By definition of the right derivative: for all $\varepsilon > 0$, there exists $\delta(x) > 0$ such that

$$\forall t,\ 0 < t < \delta(x), \qquad \left|\frac{f(x+t)-f(x)}{t} - df(x;1)\right| < \varepsilon.$$

This implies that

$$|f(x+t)-f(x)| < t\,(|df(x;+1)|+\varepsilon).$$

The number $c(x,\varepsilon) = |df(x;+1)|+\varepsilon > 0$ depends only on x and ε. Choosing

$$\delta'(x) = \min\left\{\delta(x), \frac{\varepsilon}{c(x,\varepsilon)}\right\} > 0,$$

then

$$0 < t < \delta'(x) \quad \text{yields} \quad |f(x+t)-f(x)| < t\,c(\varepsilon,x) < \varepsilon$$

and, thence, the right continuity of f, since $\varepsilon/c(x,\varepsilon) \le 1$. □

In general, the derivative of a function is not continuous as shown in the following example.

Example 2.1.
Consider the function

$$f(x) = \begin{cases} x^2 \sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases} \tag{2.12}$$

which is differentiable for $x \neq 0$:

$$f'(x) = 2x\sin\frac{1}{x} - \cos\frac{1}{x}, \quad x \neq 0.$$

For $x = 0$, go back to the definition of the differential quotient

$$\left|\frac{f(t)-f(0)}{t} = 0\right| = \left|t\sin\frac{1}{t} - 0\right| \le |t|, \quad t \neq 0;$$

by letting $t \to 0$, it is readily seen that the limit exists and that $f'(0) = 0$. Therefore, f is differentiable everywhere on $\mathbb{R}$, but f' is not a continuous function since $\cos(1/t)$ does not converge as t goes to 0. There exist no right or left limits, but

$$\liminf_{x\to 0} f'(x) = -1 \quad \text{and} \quad \limsup_{x\to 0} f'(x) = +1,$$

and, since $f'(0) = 0$, f' is neither lsc nor usc. □

This example shows that a *differentiable function* at each point can have as *derivative* a function *that is not continuous* at some points. However, that does not mean that any function can be the derivative of a *continuous function differentiable everywhere*. The derivatives of such functions in an open interval share an important property of continuous functions in an open interval: it goes through all intermediary points (section 2.3, Theorem 2.5).

2.2 Mean Value Theorem

With the notion of derivative, it is easy to recover Fermat's rule.

Theorem 2.2 (Fermat's rule)**.** *Let $f : [a,b] \to \mathbb{R}$, $a < b$. Assume that f has a* local maximum *at $x \in]a,b[$; that is,*

$$\begin{cases} \exists \text{ a neighborhood } V(x) \text{ of } x \text{ in }]a,b[\text{ such that} \\ f(x) \geq f(y), \quad \forall y \in V(x). \end{cases} \tag{2.13}$$

If f is differentiable at x, then

$$f'(x) = 0. \tag{2.14}$$

Proof. As x is an interior of $[a,b]$, choose $\delta > 0$ such that $]a,b[\supset V(x) \supset I_\delta =]x-\delta, x+\delta[$. As a result,

$$\forall y \in I_\delta, \quad f(y) \leq f(x).$$

For t, $0 < t < \delta$, we have $f(x-t) \leq f(x)$ that implies

$$\frac{f(x-t)-f(x)}{t} \leq 0 \quad \Rightarrow \quad df(x;-1) \leq 0$$

since f is differentiable at x; similarly $f(x+t) \leq f(x)$ implies

$$\frac{f(x+t)-f(x)}{t} \leq 0 \quad \Rightarrow \quad df(x;+1) \leq 0.$$

But, since f is differentiable at x, we have $0 \geq df(x;-1) = -df(x;+1) \geq 0$ and hence $f'(x) = df(x;+1) = 0$. □

The next theorem is a generalized form of the Mean Value theorem involving two functions.

Theorem 2.3. *Let f and g be two continuous functions in $[a,b]$ that are differentiable on $]a,b[$. Then*

$$\exists x \in]a,b[\text{ such that } \quad [f(b)-f(a)]\, g'(x) = [g(b)-g(a)]\, f'(x). \tag{2.15}$$

Proof. Define

$$h(t) \stackrel{\text{def}}{=} [f(b)-f(a)]g(t) - [g(b)-g(a)]f(t), \quad a \leq t \leq b.$$

It is readily seen that h is continuous on $[a,b]$, differentiable on $]a,b[$, and that

$$h(a) = f(b)g(a) - f(a)g(b) = h(b).$$

To prove the theorem, it is sufficient to show that $h'(x) = 0$ for some point $x \in]a,b[$. If h is constant, it is true for all $x \in]a,b[$. If $h(t) > h(a)$ for some $t \in]a,b[$, let x be the point of $[a,b]$ at which h achieves its maximum. It exists by the Weierstrass theorem since f is continuous on the compact interval $[a,b]$. Since $h(a) = h(b)$, the point x belongs to the open interval $]a,b[$ and $h'(x) = 0$ by Fermat's rule. If $h(t) < h(a)$ for some $t \in]a,b[$, repeat the same argument by choosing as x a point of $[a,b]$ where $-h$ achieves its maximum. □

The Mean Value theorem is now a corollary to the previous theorem.

Theorem 2.4. *Let f be continuous on $[a,b]$ and differentiable everywhere on $]a,b[$. Then there exists a point $x \in\,]a,b[$ such that*

$$f(b) - f(a) = (b-a)f'(x). \tag{2.16}$$

Proof. Set $g(x) = x$ in the previous theorem. □

2.3 Property of the Derivative of a Function Differentiable Everywhere

Go back to the issue raised in the context of Example 2.1 that shows that, in general, the derivative of a differentiable function is not continuous. It still retains the following property of continuous functions on an open interval.

Theorem 2.5 (W. RUDIN [**1**, Thm. 5.8, p. 92])**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable on $[a,b]$. Given λ such that*

$$f'(a) < \lambda < f'(b) \quad (\text{resp., } f'(a) > \lambda > f'(b)), \tag{2.17}$$

there exists a point $x \in]a,b[$ such that $f'(x) = \lambda$.

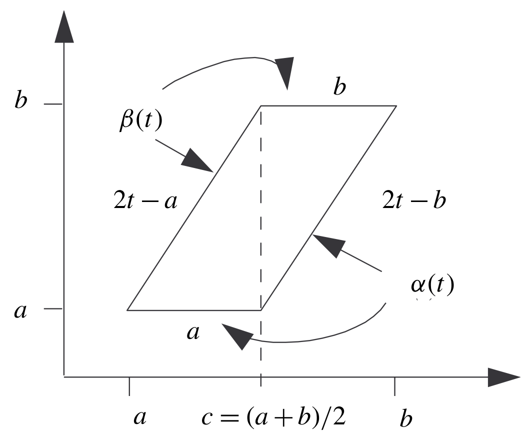


Figure 3.2. *Region determined by the functions α and β.*

Proof (see Figure 3.2 for the definitions of α and β). Let $c = (a+b)/2$. If $a \le t \le c$, define $\alpha(t) = a$, $\beta(t) = 2t - a$. If $c \le t \le b$, define $\alpha(t) = 2t - b$, $\beta(t) = b$. Then $a \le \alpha(t) < \beta(t) \le b$ on $]a,b[$. The function

$$g(t) = \frac{f(\beta(t)) - f(\alpha(t))}{\beta(t) - \alpha(t)}, \quad a < t < b,$$

is well defined and continuous on $]a,b[$. Moreover, $g(t) \to f'(a)$ as $t \to a$, $g(t) \to f'(b)$ as $t \to b$. Since the image $g([a,b])$ of the connected set $[a,b]$ by the continuous function

g is connected, we have $g(t_0) = \lambda$ for some $t_0 \in]a,b[$. Fix t_0. By Theorem 2.4 applied to f, there exists a point x such that $\alpha(t_0) < x < \beta(t_0)$ and such that $f(\beta(t_0)) - f(\alpha(t_0)) = (\beta(t_0) - \alpha(t_0))f'(x)$. Therefore, by definition of g, $f'(x) = g(t_0)$ and $f'(x) = \lambda$. $\square$

Recall that if f is discontinuous at a point x and if the limit from the right $f(x^+)$ and the limit from the left $f(x^-)$ exist and are not equal, we say that f has a discontinuity of the *first kind*. Otherwise, the discontinuity is of the *second kind*. Theorem 2.5 says that *the derivative of a differentiable function on* $[a,b]$ *cannot have discontinuities of the first kind.*

2.4 Taylor's Theorem

When f has a derivative f' on an interval and f' also has a derivative on the same interval, this second derivative will be denoted $f^{(2)}$. In a similar fashion, denote by $f^{(n)}$ the nth-order derivative of f, $n \geq 1$.

For the existence of $f^{(n)}(x)$ at a point x, it is necessary that $f^{(n-1)}$ exists in a neighborhood of x and that it is differentiable (and hence continuous) at that point. Since $f^{(n-1)}$ must exist on a neighborhood of x, $f^{(n-2)}$ must exist and must be differentiable on that neighborhood and ... we can go up to f. As a result, a function f for which $f^{(n)}$ exists on $]a,b[$ is a function such that f and its derivatives up to $n-1$ are continuous and differentiable on $]a,b[$.

Theorem 2.6 (Taylor[14]). *Let* $f :]a,b[\to \mathbb{R}$ *and assume that* $f^{(n)}$ *exists on* $]a,b[$ *for some integer* $n \geq 1$. *Given* x, $a < x < b$, *define the* $(n-1)$*th-order polynomial*

$$P_x(y) \overset{\text{def}}{=} \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!}(y-x)^k, \quad a < y < b. \tag{2.18}$$

For any $y \neq x$, $a < y < b$, *there exists* θ, $0 < \theta < 1$, *such that*

$$f(y) = P_x(y) + \frac{f^{(n)}(x + \theta\,(y-x))}{n!}(y-x)^n. \tag{2.19}$$

Proof. If $y > x$, let M be the real number implicitly defined by the identity

$$f(y) = P_x(y) + M(y-x)^n.$$

Define

$$g(t) \overset{\text{def}}{=} f(t) - P_x(t) - M(t-x)^n, \quad a < t < b.$$

It remains to prove the existence of a $z \in]x,y[$ such that $n!M = f^{(n)}(z)$ that will give $\theta = (z-x)/(y-x) \in]0,1[$. By definition of g and P_x,

$$g^{(n)}(t) = f^{(n)}(t) - n!M, \quad a < t < b$$

since $P_x^{(n)} = 0$ in $]a,b[$. So, this is equivalent to proving that $g^{(n)}(z) = 0$ for some $z \in]x,y[$. By definition of P_x,

$$P_x^{(k)}(x) = f^{(k)}(x), \quad 0 \leq k \leq n-1,$$

[14]Brook Taylor (1685–1731).

and hence

$$g(x) = g'(x) = \cdots = g^{(n-1)}(x) = 0.$$

By choice of M we also have $g(y) = 0$. By Theorem 2.4, there exists $x_1 \in]x,y[$ such that $0 = g(y) - g(x) = (y-x)\,g'(x_1)$. Since $g'(x_1) = 0$, we conclude that there exists $x_2 \in]x,x_1[$ such that $g^{(2)}(x_2) = 0$. By repeating n times, we get the existence of $x_n \in]x,x_{n-1}[$ such that $g^{(n)}(x_n) = 0$. Since $x_n \in]x,x_{n-1}[\subset]x,y[$, choose $z = x_n$ and this completes the proof for $y > x$. If $y < x$, set $\bar{f}(t) = f(-t)$, $-b < t < -a$, and apply the previous result with $-y > -x$. □

3 Real-Valued Functions of Several Real Variables

3.1 Geometrical Approach via the Differential

It is straightforward to extend the results of the previous section to vector-valued functions of a single real variable $t \mapsto h(t) : \mathbb{R} \to \mathbb{R}^n$. The derivative $h'(t)$ and the derivatives from the right $dh(t;+1)$ and from the left $dh(t;+1)$ at t are defined in the same way, but the convergence of the differential quotients takes place in the space $\mathbb{R}^n$ instead of $\mathbb{R}$ or component by component.

It is quite different for functions of several real variables $f : \mathbb{R}^n \to \mathbb{R}$. The oldest notion in the literature is the one of *differential*. For instance, for a function $f(x,y)$ of two variables, the reasoning is on the *increment* or the *variation* $\Delta f(x,y)$ of the function f resulting from the variations Δx and Δy of the variables x and y:

$$\Delta f(x,y) = \frac{\Delta f(x,y)}{\Delta x}\,\Delta x + \frac{\Delta f(x,y)}{\Delta y}\,\Delta y.$$

Assuming that, as Δx and Δy go to zero, the quotients

$$\frac{\Delta f(x,y)}{\Delta x} \to \frac{\partial f}{\partial x}(x,y) \quad \text{and} \quad \frac{\Delta f(x,y)}{\Delta y} \to \frac{\partial f}{\partial y}(x,y)$$

converge in $\mathbb{R}$, the differential is formally written as

$$df(x,y) = \frac{\partial f}{\partial x}(x,y)\,dx + \frac{\partial f}{\partial y}(x,y)\,dy \tag{3.1}$$

that underlies the notion of partial derivative in the directions of the x and y axes.

But, for J. HADAMARD[15] [**2**] in 1923, this expression is only an *operational symbol*:

> "What is the meaning of equality (3.1)? That, if x, y and hence $g = f(x,y)$ are expressed as a function of some auxiliary variable t, we have, whatever those expressions be,
>
> $$\frac{dg}{dt} = \frac{dg}{dx}\frac{dx}{dt} + \frac{dg}{dy}\frac{dy}{dt}. \tag{3.2}$$

[15] Jacques-Salomon Hadamard (1865–1963). He obtained important results on partial differential equations of mathematical physics. He was also one of the contributors to the elaboration of the modern theory of functional analysis.

> Such is the unique meaning of equality (3.1). The equality (3.2) taking place whatever be the expression of the independent variable as a function of the other two, t is deleted. The invaluable advantage of the differential notation precisely consists of the possibility not to specify what is the variable that we consider as independent."[16]

This quotation gives a precise meaning to the notion of differential. Consider the vector function

$$t \mapsto h(t) \overset{\text{def}}{=} (x(t), y(t)) : \mathbb{R} \to \mathbb{R}^2 \tag{3.3}$$

of the auxiliary real variable t and the composition g of f and h

$$t \mapsto g(t) \overset{\text{def}}{=} f(h(t)) = f(x(t), y(t)) : \mathbb{R} \to \mathbb{R}.$$

The vector function $h(t) = (x(t), y(t))$ defines a *path* or a *trajectory* in $\mathbb{R}^2$ as a function of t. Assuming, without loss of generality, that $h(0) = (x, y)$, the trajectory goes through the point $(x(0), y(0)) = (x, y)$ where the tangent vector is $(x'(0), y'(0))$. The differential at (x, y) exists when there exists a linear[17] map $L : \mathbb{R}^2 \to \mathbb{R}$ such that $g'(0) = L(h'(0))$ for all paths h through (x, y). Note that the map L depends on (x, y) but is independent of the choice of the path h. This is what we call the *geometric point of view*. It can readily be extended to functions defined on manifolds by choosing paths lying in that manifold.

If the paths are limited to lines through (x, y), we get the weaker notion of *directional derivative* in the direction (v, w) at (x, y) that is obtained by choosing paths of the form

$$t \mapsto h(t) \overset{\text{def}}{=} (x + tv, y + tw) : \mathbb{R} \to \mathbb{R}^2 \tag{3.4}$$

that yields $t \mapsto g(t) = f(h(t)) = f(x + tv, y + tw) : \mathbb{R} \to \mathbb{R}$. Equation (3.2) now becomes

$$\frac{dg}{dt}(0) = \frac{d(f \circ h)}{dt}(0) = \frac{\partial f}{\partial x}(h(0))\, v + \frac{\partial f}{\partial y}(h(0))\, w = \begin{bmatrix} \dfrac{\partial f}{\partial x}(h(0)) \\[2ex] \dfrac{\partial f}{\partial y}(h(0)) \end{bmatrix} \cdot \begin{bmatrix} v \\ w \end{bmatrix}.$$

Denote by $L(v, w)$ the right-hand side of this identity and observe that it is a scalar product. So, the function $L : \mathbb{R}^2 \to \mathbb{R}$ is linear with respect to the direction (v, w): for all $\alpha, \beta \in \mathbb{R}$ and all $(v_1, w_1), (v_2, w_2) \in \mathbb{R}^2$,

$$L(\alpha v_1 + \beta v_2, \alpha w_1 + \beta w_2) = \alpha\, L(v_1, w_1) + \beta\, L(v_2, w_2).$$

[16] From the French: "Que signifie l'égalité (3.1)? Que si x, y et dès lors $g = f(x, y)$ sont exprimés en fonction d'une variable auxiliaire quelconque t, on a, quelles que soient ces expressions,

$$\frac{dg}{dt} = \frac{dg}{dx}\frac{dx}{dt} + \frac{dg}{dy}\frac{dy}{dt}. \tag{3.2}$$

Tel est le sens unique de l'égalité (3.1). L'égalité (3.2) ayant lieu quelle que soit la variable indépendante en fonction de laquelle les deux autres variables sont exprimées, on supprime la mention de t. L'avantage précieux de la notation différentielle consiste précisément en la possibilité de ne pas préciser quelle est la variable que l'on considère comme indépendante."

[17] See Definition 4.10 in Chapter 1.

Retain two elements from the quotation of Hadamard: (a) identity (3.2) must be verified for all paths $h(t) = (x(t), y(t))$ and not only along lines; (b) the differential must be linear with respect to the tangent vector $h'(0) = (x'(0), y'(0))$.

We now give two examples to illustrate those considerations.

Example 3.1.
Let $f(x, y) = x^2 + 2\,y^2$. It is easy to check that

$$\frac{dg}{dt}(0) = 2x\; x'(0) + 4y\; y'(0) = \begin{bmatrix} 2x \\ 4y \end{bmatrix} \cdot \begin{bmatrix} x'(0) \\ y'(0) \end{bmatrix}$$

for all paths $h(t) = (x(t), y(t))$ given by (3.3) satisfying $h(0) = (x, y)$. □

Example 3.2.
Let

$$f(x, y) \overset{\text{def}}{=} \frac{x^3}{x^2 + y^2}, \text{ if } (x, y) \neq (0,0), \quad f(0,0) \overset{\text{def}}{=} 0. \tag{3.5}$$

The function f is continuous at every point of the plane (including $(0,0)$), and for all paths h given by (3.3) satisfying $h(0) = (x, y)$, we have

$$\frac{dg}{dt}(0) = \frac{x'(0)^3}{x'(0)^2 + y'(0)^2}.$$

Specializing to paths along lines in the direction (v, w) through $(x, y) = (0,0)$,

$$t \mapsto h(t) \overset{\text{def}}{=} (tv, tw) : \mathbb{R} \to \mathbb{R}^2,$$

consider the composition

$$g(t) \overset{\text{def}}{=} f(h(t)) = \frac{(tv)^3}{(tv)^2 + (tw)^2} = t\,\frac{v^3}{v^2 + w^2} \quad \Rightarrow \frac{dg}{dt}(0) = \frac{v^3}{v^2 + w^2}.$$

Observe that the map $(v, w) \mapsto L(v, w) = g'(0) : \mathbb{R}^2 \to \mathbb{R}$ is well defined and independent of the choice of the path, but it is not linear in (v, w). We obtain this result in spite of the fact that the partial derivatives

$$\frac{\partial f}{\partial x}(0,0) = 1 \quad \text{and} \quad \frac{\partial f}{\partial y}(0,0) = 0$$

exist. So, we would have expected identity (3.2)

$$\frac{dg}{dt}(0) = \begin{bmatrix} \dfrac{\partial f}{\partial x}(0,0) \\[2ex] \dfrac{\partial f}{\partial y}(0,0) \end{bmatrix} \cdot \begin{bmatrix} \dfrac{dx}{dt}(0) \\[2ex] \dfrac{dy}{dt}(0) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x}(0,0) \\[2ex] \dfrac{\partial f}{\partial y}(0,0) \end{bmatrix} \cdot \begin{bmatrix} v \\[2ex] w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} v \\ w \end{bmatrix} = v \neq \frac{v^3}{v^2 + w^2}$$

for $w \neq 0$ and $v \neq 0$. However, only the linearity is missing. □

The second example clearly illustrates that, working with the notion of *differential* in dimensions greater than one, does not always yield the equivalent of a derivative in dimension one and a *gradient* in higher dimensions. It is customary to say that the function of the second example is *not differentiable* at (x,y), even if there exists a *differential* independent of the path through (x,y). Hence, it is necessary to carefully review the notion of differential and determine how far it can be relaxed while preserving the basic elements of the differential calculus.

3.2 Semidifferentials, Differentials, Gradient, and Partial Derivatives

The privileged approach in this section is the one of the first variation of the *calculus of variations* as opposed to the geometric approach of the previous section. We begin with the notions of *semidifferentials* and *differentials*. Then, we shall relate them to the previous geometric notion of differential in the next section. Several names will come up in that context: Karl Weierstrass (1815–1897) who is most likely the first to have given a correct definition of the differential of a function of several variables, Otto Stolz (1842–1905), James Pierpont (1866–1938), William Henry Young (1863–1942), Jacques Hadamard (1865–1963), Maurice Fréchet (1873–1973), and René Gateaux (1889–1914) as well as the one of Paul Lévy (1886–1971), who carried out the publication of the work of Gateaux after his death in the early moments of the First World War (1914–1918).

3.2.1 Definitions

Let $f:\mathbb{R}^n \to \mathbb{R}$. Starting from a point x and a *direction*[18] $v \in \mathbb{R}^n$, we can consider the new real-valued function of the variable t

$$t \mapsto g(t) \overset{\text{def}}{=} f(x+tv) : \mathbb{R} \to \mathbb{R} \tag{3.6}$$

of the real variable t and we are back to the framework and conditions of section 2. Yet, some stronger definitions will be required that will involve the variation of f with respect to both t and v.

Definition 3.1.
Let $f:\mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$, and let $v \in \mathbb{R}^n$ be a direction.

(i) • f is *semidifferentiable* at x in the direction v if the following limit exists:

$$\lim_{t \searrow 0^+} \frac{f(x+tv)-f(x)}{t}. \tag{3.7}$$

When the limit (3.7) exists, it will be denoted $df(x;v)$. By definition, we get $df(x;0)=0$ and the positive homogeneity

$$\forall \alpha \geq 0, \quad df(x;\alpha v) \text{ exists and } df(x;\alpha v) = \alpha\, df(x;v).$$

[18] A direction is often interpreted as a vector $v \in \mathbb{R}^n$ of norm one. In this book, the term *direction* is used for any vector $v \in \mathbb{R}^n$ including 0.

• f is *semidifferentiable* at x if $df(x;v)$ exists for all $v \in \mathbb{R}^n$.

• f is *Gateaux differentiable*[19] at x if the semidifferential $df(x;v)$ exists in all directions $v \in \mathbb{R}^n$ and the map

$$v \mapsto df(x)(v) \stackrel{\text{def}}{=} df(x;v) : \mathbb{R}^n \to \mathbb{R} \tag{3.8}$$

is linear.[20]

(ii) • f is *Hadamard semidifferentiable*[21] at x in the direction v if the following limit exists:

$$\lim_{\substack{t \searrow 0^+ \\ w \to v}} \frac{f(x+tw)-f(x)}{t}. \tag{3.9}$$

When the limit (3.9) exists, it will be denoted $d_H f(x;v)$. By definition, we have $d_H f(x;v) = df(x;v)$.

• f is *Hadamard semidifferentiable* at x if $d_H f(x;v)$ exists for all $v \in \mathbb{R}^n$.

• f is *Hadamard differentiable* at x if the semidifferential $d_H f(x;v)$ exists in all directions $v \in \mathbb{R}^n$ and the map

$$v \mapsto d_H f(x)(v) \stackrel{\text{def}}{=} d_H f(x;v) : \mathbb{R}^n \to \mathbb{R} \tag{3.10}$$

is linear. □

By definition, if $d_H f(x;v)$ exists, then $df(x;v)$ exists and $d_H f(x;v) = df(x;v)$. However, even if $df(x;0)$ always exists and is equal to 0, $d_H f(x;0)$ does not always exist as shown in Example 3.6 (Figure 3.5).

Remark 3.1.
In Definition 3.1, checking the existence of the limit as $w \to v$ and $t \searrow 0$ is equivalent to using all sequences $\{w_n\}$ and $\{t_n > 0\}$ converging to v and 0, respectively. For instance, $d_H f(x;v)$ exists if there exists $q \in \mathbb{R}$ such that for all sequences $\{t_n > 0\}$, $t_n \to 0$, and $\{w_n\}$, $w_n \to v$,

$$\frac{f(x+t_n w_n)-f(x)}{t_n} \to q.$$

□

Example 3.3.
Consider the square $f(x) = \|x\|^2$ of the norm $\|x\| = \sqrt{x \cdot x}$ of x in $\mathbb{R}^n$. It is continuous on $\mathbb{R}^n$. For all $w \to v$ and $t > 0$ going to 0,

[19] René Eugène Gateaux (1889–1914). In his birth certificate as in his texts, his letters, and his publications before his death (R. GATEAUX [**1**, **2**, **3**, **4**, **5**]) his name is spelled without circumflex accent (see L. MAZILAK [**1**], M. BARBUT, B. LOCKER, L. MAZILAK, and P. PRIOURET [**1**], L. MAZILAK and R. TAZZIOLI [**1**]). The accent appeared in his three posthumous publications from 1919 to 1922 in the *Bulletin de la Société Mathématique de France* probably by homonymy with the French word "gâteaux" (cakes) (see L. MAZILAK [**2**]). His important unpublished work was left in Jacques Hadamard's care. It was completed by Paul Lévy who saw to their publication in the *Bulletin de la Société Mathématique de France* (see R. GATEAUX [**6**, **7**, **8**]). Gateaux was awarded (post-mortem) the Francœur Prize of the Academy of Sciences in 1916 for his work on the *functional calculus*. Paul Lévy (1886–1971) made Gateaux's work known in his *Leçons d'analyse fonctionnelle* (1922).

[20] Cf. Definition 4.10 of Chapter 1.

[21] Cf. footnote 36 on page 96 and the comments at the end of section 3.3.

$$\frac{f(x+tw)-f(x)}{t}=\frac{\|x+tw\|^2-\|x\|^2}{t}=\frac{(2x+tw)\cdot tw}{t}=(2x+tw)\cdot w\to 2x\cdot v$$

by continuity of the scalar product. Therefore, $d_H f(x;v)$ exists for all x and all v,

$$d_H f(x;v)=2x\cdot v,$$

and as $v\mapsto d_H f(x;v)$ is linear[22] in v, it is Hadamard differentiable. □

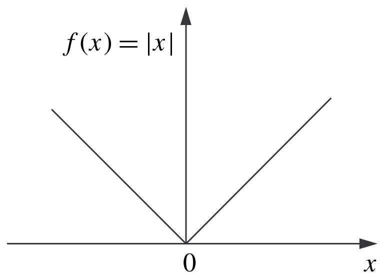


Figure 3.3. *The function $f(x)=|x|$ in a neighborhood of $x=0$ for $n=1$.*

The second example is the norm which is Hadamard semidifferentiable at 0, but not linear with respect to the direction v.

Example 3.4 (Figure 3.3)**.**
Consider the norm $f(x)=\|x\|=\sqrt{x\cdot x}$ of x in $\mathbb{R}^n$. It is continuous at $\mathbb{R}^n$. First consider the case $x\neq 0$. For all $w\to v$ and $t>0$,

$$\begin{aligned}\frac{f(x+tw)-f(x)}{t}=\frac{\|x+tw\|-\|x\|}{t}&=\frac{\|x+tw\|^2-\|x\|^2}{t}\,\frac{1}{\|x+tw\|+\|x\|}\\ &\to 2x\cdot v\,\frac{1}{2\|x\|}=\frac{x}{\|x\|}\cdot v\end{aligned}$$

by continuity of the norm. Therefore, $d_H f(x;v)$ exists,

$$d_H f(x;v)=\frac{x}{\|x\|}\cdot v,$$

and since $d_H f(x;v)$ is linear with respect to v, the norm is Hadamard differentiable at every point $x\neq 0$. We now show that f has a Hadamard semidifferential at $x=0$ that is not linear in v. Indeed, for all $w\to v$ and $t>0$,

$$\frac{f(0+tw)-f(0)}{t}=\frac{\|tw\|}{t}=\frac{t}{t}\|w\|=\|w\|\to\|v\|$$

by continuity of the norm. Therefore, $d_H f(0;v)$ exists,

$$d_H f(0;v)=\|v\|,$$

for all directions v, but the map $v\mapsto d_H f(0;v)$ is not linear. □

We now complete the definitions by introducing the notion of directional derivative and the special case of partial derivatives.

[22] Cf. Definition 4.10 of Chapter 1.

Definition 3.2.
Let $f : \mathbb{R}^n \to \mathbb{R}, x \in \mathbb{R}^n$, and let $v \in \mathbb{R}^n$ be a direction.

(i) f has a *derivative in the direction* v at the point x if the following limit exists:

$$\lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}. \tag{3.11}$$

By definition, we get $df(x;-v) = -df(x;v)$ and the homogeneity

$$\forall \alpha \in \mathbb{R}, \quad df(x;\alpha v) \text{ exists and } df(x;\alpha v) = \alpha\, df(x;v).$$

(ii) Let $\{e_i : 1 \le i \le n\}$, $(e_i)_j = \delta_{ij}$, be the canonical orthonormal basis in $\mathbb{R}^n$. f has *partial derivatives* at x if for each i, f is differentiable at x in the direction e_i; that is,

$$\lim_{t \to 0} \frac{f(x+t\,e_i) - f(x)}{t} \text{ exists.}$$

The limit is denoted $\partial_i f(x)$ or $\partial f/\partial x_i(x)$. By definition, $\partial_i f(x) = df(x;e_i)$ and the function $\alpha \mapsto df(x;\alpha\, e_i) : \mathbb{R} \to \mathbb{R}$ is homogeneous. □

The intermediary notion of directional derivative is of limited interest since it hides the more fundamental notion of semidifferential and is not yet the Gateaux differential.

3.2.2 Examples and Counterexamples

It is useful to consider a series of examples to fully appreciate the differences between the previous definitions:

(i) in general, a continous function with partial derivatives may not be semidifferentiable in all directions $v \in \mathbb{R}^n$ (see Example 3.5);

(ii) $df(x;0)$ always exists and $df(x;0) = 0$, but $d_H f(x;0)$ does not necessarily exist (see Example 3.6 and Theorem 3.7 that will establish that when $d_H f(x;0)$ exists, the function f is continuous at x);

(iii) $df(x;v)$ can exist in all directions while $d_H f(x;v)$ does not (see Example 3.6).

(iii) $d_H f(x;v)$ can exist in all directions without the linearity of the map $v \mapsto d_H f(x;v)$ (Example 3.7 is the same as Example 3.2 already introduced in section 3.1).

Moreover,

$$\begin{array}{lll}
\boxed{f \text{ Hadamard differentiable at } x} & \left\langle \begin{array}{l} \Rightarrow \\ \not\Leftarrow \text{ (Example 3.5)} \end{array} \right. & \\
\quad \Downarrow (\not\Uparrow \text{ Counterexample 3.6}) & & \\
\boxed{f \text{ Gateaux differentiable at } x} & \left\langle \begin{array}{l} \not\Rightarrow \text{ (Example 3.6)} \\ \not\Leftarrow \text{ (Example 3.5)} \end{array} \right. & \boxed{\begin{array}{c} f \text{ continuous} \\ \text{at } x \end{array}} \\
\quad \Downarrow (\not\Uparrow \text{ Counterexample 3.7}) & & \\
\boxed{f \text{ differentiable in all directions at } x} & \left\langle \begin{array}{l} \not\Rightarrow \text{ (Example 3.6)} \\ \not\Leftarrow \text{ (Example 3.5)} \end{array} \right. &
\end{array}$$

In general, the converse statements are not true as will be shown in the examples.

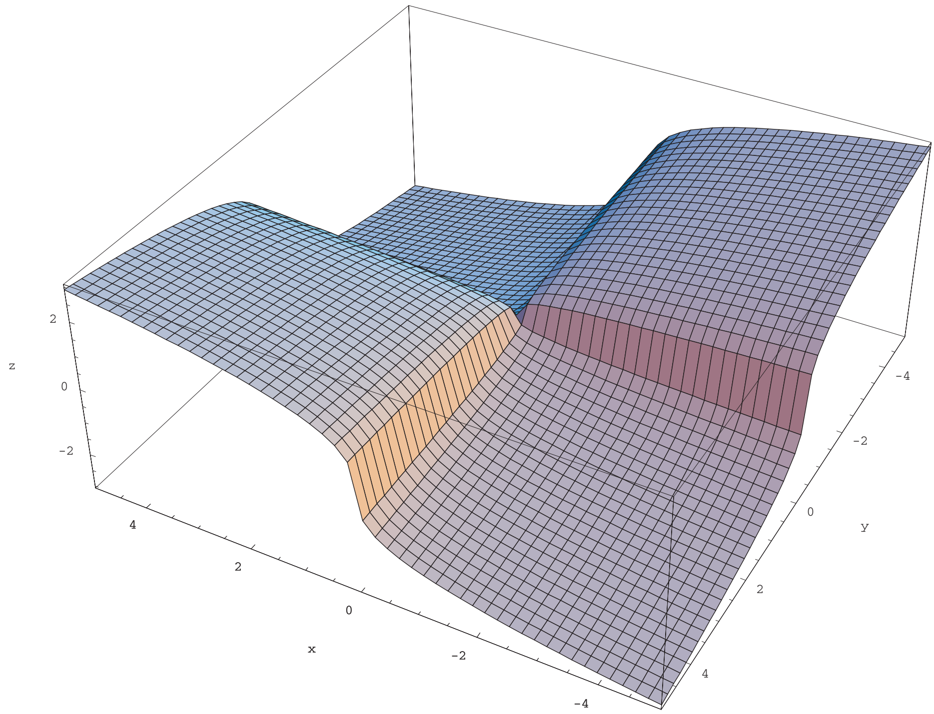


Figure 3.4. *Example* 3.5.

The first example is a function which is continuous with partial derivatives at $(0,0)$ that is not Gateaux semidifferentiable in all directions.

Example 3.5 (Figure 3.4)**.**
Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x,y) = \begin{cases} (xy)^{1/3}, & \text{if } xy \geq 0 \\ -|xy|^{1/3}, & \text{if } xy < 0. \end{cases}$$

It is continuous on $\mathbb{R}^2$. Compute its directional derivative at $(0,0)$ in the direction $v = (v_1, v_2)$. For $v_1 v_2 \geq 0$ and $t > 0$,

$$\frac{f(0+tv) - f(0)}{t} = \frac{(tv_1 tv_2)^{1/3}}{t} = \frac{1}{t^{1/3}} (v_1 v_2)^{1/3},$$

and for $v_1 v_2 < 0$,

$$\frac{f(0+tv) - f(0)}{t} = -\frac{|tv_1 tv_2|^{1/3}}{t} = -\frac{1}{t^{1/3}} |v_1 v_2|^{1/3}.$$

The differential quotients converge if and only if $v_1 v_2 = 0$. Therefore,

$$df((0,0);(v_1,v_2)) = 0 \quad \text{if} \quad v_1 v_2 = 0$$

and does not exist for $v_1 v_2 \neq 0$. In fact, under the condition $v_1 v_2 = 0$, we have

$$\lim_{t \to 0} \frac{f(0+tv) - f(0)}{t} = 0,$$

that is, a semidifferential and even a directional derivative.

Specializing to the canonical orthonormal basis $e_1 = (1,0)$ and $e_2 = (0,1)$ of $\mathbb{R}^2$, we get

$$\frac{\partial f}{\partial x_1}(0,0) = df((0,0);e_1) = 0 \text{ and } \frac{\partial f}{\partial x_2}(0,0) = df((0,0);e_2) = 0,$$

where $\partial f/\partial x_1(0,0)$ and $\partial f/\partial x_2(0,0)$ are the partial derivatives of f at $(0,0)$. This example shows that for directions (v_1, v_2), $v_1\, v_2 \neq 0$, the semidifferential $df((0,0);v_1,v_2)$ does not exist and, hence, is not equal to

$$\left(\frac{\partial f}{\partial x_1}(0,0), \frac{\partial f}{\partial x_2}(0,0) \right) \cdot (v_1\, v_2). \qquad \square$$

The second example is both rich and important. The function is discontinuous and Gateaux differentiable, but not Hadamard semidifferentiable at $(0,0)$. In particular, $d_H f(x;0)$ does not exist, even if $df(x;0)$ exists and $df(x;0) = 0$.

Example 3.6 (Figure 3.5)**.**
Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$f(x,y) \overset{\text{def}}{=} \begin{cases} \dfrac{x^6}{(y-x^2)^2 + x^8}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

It is directionally differentiable and even Gateaux differentiable at $x = (0,0)$, but not Hadamard semidifferential at $(0,0)$ in the directions $(0,0)$ and $(1,0)$, and f is not continuous at $(0,0)$.

First compute the directional derivative of f at $(0,0)$. For $v = (v_1, v_2) \neq (0,0)$, consider the following two cases: $v_2 = 0$ and $v_2 \neq 0$. If $v_2 = 0$ and $v_1 \neq 0$,

$$\begin{aligned} \frac{f(tv_1,0) - f(0,0)}{t} &= \frac{1}{t} \frac{(tv_1)^6}{(0-(tv_1)^2)^2 + (tv_1)^8} = \frac{1}{t} \frac{(tv_1)^6}{(tv_1)^4 + (tv_1)^8} \\ &= t \frac{v_1^2}{1+(tv_1)^4} \end{aligned}$$

and as t goes to 0,

$$df((0,0);(v_1,v_2)) = 0.$$

If $v_2 \neq 0$,

$$\frac{f(tv_1,tv_2) - f(0,0)}{t} = \frac{1}{t} \frac{(tv_1)^6}{(tv_2 - (tv_1)^2)^2 + (tv_1)^8} = t^3 \frac{v_1^6}{(v_2 - tv_1^2)^2 + t^6 v_1^8}$$

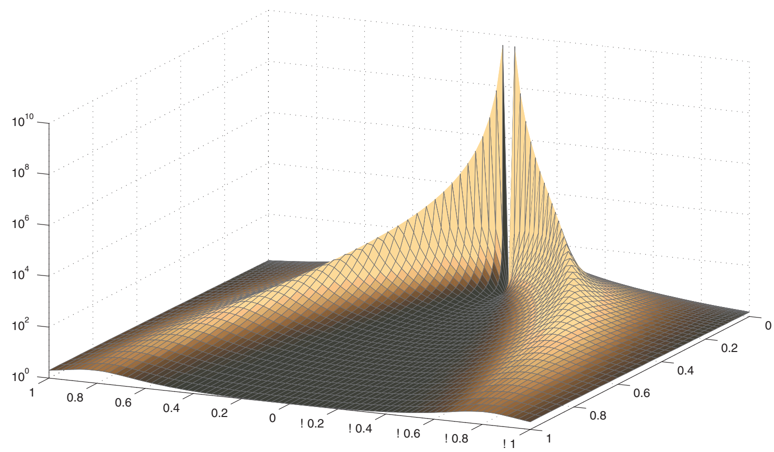


Figure 3.5. *Examples* 3.6 *and* 3.8 *in logarithmic scale.*

and as t goes to 0,

$$df((0,0);(v_1,v_2)) = 0 \text{ if } v_2 \neq 0.$$

Therefore,

$$\forall v = (v_1,v_2) \in \mathbb{R}^2, \quad df((0,0);(v_1,v_2)) = 0,$$

the map $(v_1,v_2) \mapsto df((0,0);(v_1,v_2)) : \mathbb{R}^2 \to \mathbb{R}$ is linear, and f is Gateaux differentiable.

To show that $d_H f((0,0);(0,0))$ does not exist, choose the following sequences:

$$t_n = \frac{1}{n} \searrow 0 \text{ and } w_n = \left(\frac{1}{n}, \frac{1}{n^3}\right) \to (0,0) \text{ as } n \to +\infty.$$

Consider the quotient

$$q_n \stackrel{\text{def}}{=} \frac{f((0,0)+t_n w_n) - f(0,0)}{t_n}.$$

As $n \to \infty$,

$$q_n = \frac{(\frac{1}{n^2})^6}{\frac{1}{n}(\frac{1}{n^2})^8} = n^5 \to +\infty$$

and $d_H f((0,0);(0,0))$ does not exist. It is also readily seen that f is discontinuous at $x = (0,0)$ by following the path (α,α^2) as α goes to 0:

$$|f(\alpha,\alpha^2) - f(0,0)| = \left|\frac{\alpha^6}{\alpha^8}\right| = \frac{1}{\alpha^2} \to +\infty \text{ when } \alpha \to 0.$$

To show that $d_H((0,0);(1,0))$ does not exist, compute the differential quotient for two different sequences $(w_n,t_n) \to (v,0)$ and show that the two limits are different. First choose

$$w_n = (1,0), \ \forall n \ \text{ and } \ t_n = \frac{1}{n}, \ \forall n$$

$$\frac{f(t_n w_n) - f(0,0)}{\frac{1}{n}} = \frac{1}{\frac{1}{n}} \frac{(\frac{1}{n})^6}{(\frac{1}{n})^4 + (\frac{1}{n})^8} = \frac{(\frac{1}{n})}{1 + (\frac{1}{n})^4} \to 0.$$

Then choose

$$w_n = \left(1, \frac{1}{n}\right), \ \forall n \ \text{ and } \ t_n = \frac{1}{n}, \ \forall n$$

$$\frac{f(t_n w_n) - f(0,0)}{\frac{1}{n}} = \frac{1}{\frac{1}{n}} \frac{(\frac{1}{n})^6}{(\frac{1}{n})^8} = n^3 \to +\infty. \qquad \square$$

The last example is the same as Example 3.2 already introduced in section 3.1. It is a continuous Hadamard semidifferentiable function, but not linear with respect to the direction.

Example 3.7.
Consider the function

$$f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

It is readily seen that it is continuous at $(0,0)$:

$$\left|\frac{x^3}{x^2+y^2}\right| = |x| \left|\frac{x^2}{x^2+y^2}\right| \quad \leq \quad |x| \to 0 \ \text{ as } \ (x,y) \to (0,0).$$

For $v = (v_1,v_2) \neq 0$ and $w = (w_1,w_2) \to v = (v_1,v_2)$,

$$\frac{f(tw) - f(0)}{t} = \frac{1}{t} \frac{(tw_1)^3}{(tw_1)^2 + (tw_2)^2} = \frac{w_1^3}{w_1^2 + w_2^2} = f(w_1,w_2) \to f(v_1,v_2) = \frac{v_1^3}{v_1^2+v_2^2}$$

and necessarily

$$\forall v \in \mathbb{R}^2, \quad d_H f(0;v) = \left\{\begin{matrix} v_1^3/(v_1^2+v_2^2), & \text{if } (v_1,v_2) \neq (0,0) \\ 0, & \text{if } (v_1,v_2) = (0,0) \end{matrix}\right\} = f(v_1,v_2).$$

So f is not Gateaux differentiable since $v \mapsto d_H f(0;v)$ is not linear. $\square$

3.2.3 Gradient

Let $\{e_i : 1 \leq i \leq n\}$, $(e_i)_j = \delta_{ij}$, be the canonical orthonormal basis in $\mathbb{R}^n$. Any element $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ can be written in the form $x = \sum_{i=1}^n x_i\, e_i$ and the inner product is

$$x \cdot y \overset{\text{def}}{=} \sum_{i=1}^n x_i\, y_i. \tag{3.12}$$

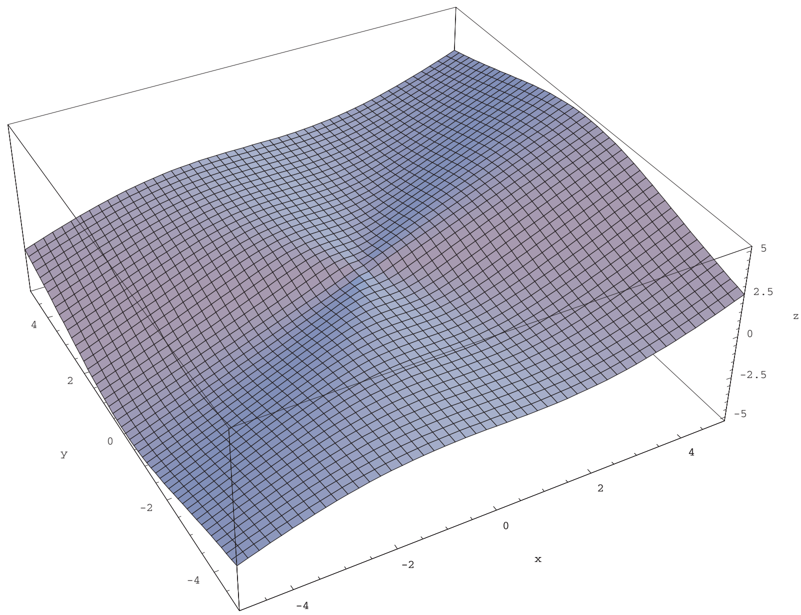


Figure 3.6. *Example* 3.7.

Similarly, any direction $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ can be written as

$$v = \sum_{i=1}^{n} v_i e_i \quad \text{and} \quad v_i = v \cdot e_i.$$

As f is Gateaux differentiable at x, the map $v \mapsto df(x;v) : \mathbb{R}^n \to \mathbb{R}$ is linear. Therefore,

$$df(x;v) = df\left(x; \sum_{i=1}^{n} v_i e_i\right) = \sum_{i=1}^{n} v_i \, df(x;e_i) = g(x) \cdot v,$$

where

$$g(x) \stackrel{\text{def}}{=} \sum_{i=1}^{n} df(x;e_i) e_i \in \mathbb{R}^n.$$

The vector $g(x)$ is unique. Indeed, if there exist g_1 and g_2 in $\mathbb{R}^n$ such that

$$\forall v \in \mathbb{R}^n, \quad g_1 \cdot v = df(x;v) = g_2 \cdot v,$$

then for all $v \in \mathbb{R}^n$ we have $(g_1 - g_2) \cdot v = 0$ and hence $g_1 = g_2$.

Definition 3.3.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be *Gateaux differentiable* in $x \in \mathbb{R}^n$. The *gradient* of f at x is the unique vector $\nabla f(x)$ of $\mathbb{R}^n$ such that

$$\forall v \in \mathbb{R}^n, \quad \nabla f(x) \cdot v = df(x;v). \tag{3.13}$$

In particular,

$$\nabla f(x) = \sum_{i=1}^{n} \partial_i f(x) e_i,$$

where $\partial_i f(x) = df(x; e_i)$ is the *partial derivative* of f at x in the direction e_i. $\square$

Example 3.5 shows that, even if the partial derivatives exist, the gradient may not exist and, a fortiori, semidifferentials $df(x;v)$ may not exist in some directions v.

3.2.4 Fréchet Differential

In this section, we turn to the notion of differential that is usually found in contemporary books in analysis. For terminological reasons, it will be referred to as the *Fréchet differentiability*. In finite dimensions, it is equivalent to Hadamard differentiability that is simpler to characterize and can be extended to infinite-dimensional topological vector spaces that do not have a metric structure.

Definition 3.4.
$f : \mathbb{R}^n \to \mathbb{R}$ is *Fréchet differentiable*[23] at $x \in \mathbb{R}^n$ if there exists a linear[24] map $L(x) : \mathbb{R}^n \to \mathbb{R}$ such that

$$\lim_{v \to 0} \frac{f(x+v) - f(x) - L(x)v}{\|v\|} = 0. \tag{3.14}$$

$\square$

Remark 3.2.
This definition was initially given by M. FRÉCHET [**1**] in 1911 in the context of functionals, that is, functions of functions. But, in finite dimension, his definition is equivalent to the earlier notion of *total differential* used by O. Stolz[25] in 1893, J. Pierpont[26] in 1905, and W. H. Young[27] in 1908–1909:

> "In fact, *an equivalent definition had been given in 1908 by M.* W. H. YOUNG [**1**, p. 157], [**2**, p. 21], who had, besides, explicitly developed the consequences." (translated from M. FRÉCHET [**2**])[28]
>
> "But, I noticed that this definition was already in Stolz, *Grundzüge der Differential und Integral-Rechnung*, t. I, p. 133, and in James Pierpont, *The theory of functions of real variables*, t. I, p. 268. But, it is W. H. Young who was the first to truly show all the advantages in his small Book: *The fundamental*

[23] Maurice René Fréchet (1873–1973) made important contributions to real analysis and built the foundations of abstract spaces. He wrote his thesis under his supervision of Hadamard in 1906. He introduced the concept of metric space and the abstract formulation of compactness.

[24] The map $v \mapsto f(x) + L(x)v$ can be seen as an affine approximation of $f(x+v)$ in $(x, f(x))$ at the infinitesimal scale.

[25] Otto Stolz (1842–1905) (see O. STOLZ [**1**, 133]).

[26] James Pierpont (1866–1938) (see J. PIERPONT [**1**, p. 268]).

[27] William Henry Young (1863–1942) (see W. H. YOUNG [**1**, p. 157], [**2**, p. 21]).

[28] From the French: "En fait, *une définition équivalente avait été donnée en 1908 par M.* W. H. YOUNG [**1**, p. 157], [**2**, p. 21], qui avait, en outre, développé explicitement les conséquences."

> *theorems of Differential Calculus* and in a few Memoirs." (translated from M. FRÉCHET [**3**])[29]

According to V. M. TIHOMIROV [**1**], "the correct definitions of derivative and differential of a function of many variables were given by K. Weierstrass[30] in his lectures in the eighties of the 19th century. These lectures were published in the thirties of our century (20th). The correct definitions of the derivative in the multidimensional case appear also at the beginning of the century in some German and English text-books (Scholz, Young) under the influence of Weierstrass." He does not provide a more specific reference but both of them had had contacts with Weierstrass over long stays In Germany. □

It is readily seen that a Fréchet differentiable function at x is Gateaux differentiable at x and that

$$df(x;v) = L(x)v, \quad \forall v \in \mathbb{R}^n.$$

Indeed, the result is true for $v = 0$ since $df(x;0) = 0 = L(x)0$. For $v \neq 0$ and $t > 0$,

$$tv \to 0 \text{ as } t \searrow 0.$$

As f is Fréchet differentiable at x,

$$\lim_{t \searrow 0} \frac{f(x+tv) - f(x) - L(x)(tv)}{\|tv\|} = 0.$$

Since $t > 0$ and $v \neq 0$, eliminate $\|v\|$ and rewrite the expression as

$$\lim_{t \searrow 0} \left| \frac{f(x+tv) - f(x)}{t} - L(x)v \right| = 0$$

that is the semidifferential $df(x;v)$ of f at x in the direction v and hence

$$df(x;v) = \lim_{t \searrow 0} \frac{f(x+tv) - f(x)}{t} = L(x)v.$$

This is precisely the definition of the semidifferential of f at x in the direction v. Since the map $L(x) : \mathbb{R}^n \to \mathbb{R}$ is linear, f is Gateaux differentiable at x and by Definition 3.3 of the gradient,

$$\forall v \in \mathbb{R}^n, \quad \nabla f(x) \cdot v = df(x;v) = L(x)v.$$

The next example shows that a Gateaux differentiable function f at x is not necessarily Fréchet differentiable at x, and not even continuous at x.

Example 3.8.
Go back to the function $f : \mathbb{R}^2 \to \mathbb{R}$ of Example 3.6 defined by

$$f(x,y) \stackrel{\text{def}}{=} \begin{cases} \dfrac{x^6}{(y-x^2)^2 + x^8}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

[29] From the French: "Mais je me suis aperçu qu'on trouve déjà cette définition dans Stolz, *Grundzüge der Differential und Integral-Rechnung*, t. I, p. 133, et James Pierpont, *The theory of functions of real variables*, t. I, p. 268. Mais c'est W. H. Young qui en a véritablement montré le premier tous les avantages dans son petit Livre: *The fundamental theorems of Differential Calculus* et dans quelques Mémoires."

[30] Karl Theodor Wilhelm Weierstrass (1815–1897).

It was shown in Example 3.6 that $df((0,0);(v_1,v_2)) = 0$ for all (v_1,v_2), that f is Gateaux differentiable in $(0,0)$, and that f is discontinuous at $(0,0)$. We now show that f is not Fréchet differentiable at $(0,0)$. Choose

$$v(\alpha) = (\alpha,\alpha^2), \quad \alpha \neq 0.$$

As α goes to 0, $v(\alpha)$ goes to $(0,0)$. Compute the Fréchet quotient

$$\begin{aligned} q(\alpha) &\stackrel{\text{def}}{=} \frac{f(\alpha,\alpha^2) - f(0,0) - df((0,0);(\alpha,\alpha^2))}{\|(\alpha,\alpha^2)\|} \\ &= \frac{1}{\alpha^3}\frac{1}{(1+\alpha^2)^{\frac{1}{2}}} \to +\infty \text{ as } \alpha \to 0. \end{aligned}$$

By definition, the function f is not Fréchet differentiable at $(0,0)$. □

The next theorem throws the bridge between the Fréchet and the Hadamard differentiabilities.

Theorem 3.1. *Let $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$. The following conditions are equivalent:*

(i) f *is* Fréchet differentiable *at x;*

(ii) f *is* Hadamard differentiable *at x.*

In dimension $n = 1$, Fréchet, Hadamard, and Gateaux differentiabilities in x coincide and correpond to the derivative in x of Definition 2.1(iii).

Proof. (i) ⇒ (ii). Consider the quotient

$$q(t,w) \stackrel{\text{def}}{=} \frac{f(x+tw) - f(x)}{t}$$

as $w \to v$ and $t \searrow 0$. We have $h(t,w) \stackrel{\text{def}}{=} tw \to 0$ since $w \to v$ and $t \searrow 0$. Then

$$q(t,w) = Q(h(t,w))\,\|w\| + L(x)\,w,$$

where

$$Q(h) \stackrel{\text{def}}{=} \frac{f(x+h) - f(x) - L(x)h}{\|h\|} \text{ if } h \neq 0 \quad \text{and} \quad Q(0) \stackrel{\text{def}}{=} 0. \tag{3.15}$$

Since f is Fréchet differentiable at x and $h(t,w) = tw \to 0$, $Q(h(t,w)) \to 0$. Moreover, by continuity of $L(x)$, $L(x)\,w \to L(x)\,v$ as $w \to v$. Therefore,

$$\lim_{\substack{t \searrow 0^+ \\ w \to v}} q(t,w) = L(x)\,v,$$

$d_H f(x;v)$ exists, and $d_H f(x;v) = L(x)\,v$ is linear (and continuous) with respect to v.

(ii) ⇒ (i) As f is Hadamard differentiable, $d_H(x;v)$ exists for all v and the map $v \mapsto L(x)v \stackrel{\text{def}}{=} d_H f(x;v)$ is linear. Letting

$$\overline{Q} \stackrel{\text{def}}{=} \limsup_{\|h\| \to 0} |Q(h)|,$$

there exists a sequence $\{h_n\}$, $0 \neq h_n \to 0$, such that $|Q(h_n)|$ converges to $\overline{Q} \in [0,\infty]$.

Since $\{h/\|h\| : \forall h \in \mathbb{R}^n,\, h \neq 0\}$ is the compact sphere S of radius 1 in $\mathbb{R}^n$, there exists a subsequence $\{h_{n_k}\}$ and a point $v \in S$ such that

$$w_{n_k} \stackrel{\text{def}}{=} \frac{h_{n_k}}{\|h_{n_k}\|} \to v \in S.$$

As $h \neq 0$,

$$Q(h) = Q\left(\|h\| \frac{h}{\|h\|}\right) = \frac{f\left(x + \|h\| \frac{h}{\|h\|}\right) - f(x)}{\|h\|} - L(x)\frac{h}{\|h\|}.$$

Since $d_H f(x;v)$ exists and $L(x)w_{n_k} \to L(x)v = d_H f(x;v)$, choosing the subsequence $t_{n_k} = \|h_{n_k}\|$ that goes to 0, we get

$$\begin{aligned} Q(h_{n_k}) = & \frac{f(x + t_{n_k} w_{n_k}) - f(x)}{t_{n_k}} - L(x)\, w_{n_k} \\ & \to d_H f(x;v) - L(x)\, v = d_H f(x;v) - d_H f(x;v) = 0 \\ & \Rightarrow |Q(h_{n_k})| \to 0 \text{ and } \overline{Q} = \limsup_{h \to 0} |Q(h)| = 0. \end{aligned}$$

Since $|Q(h)| \geq 0$ and the limsup $\overline{Q}$ is equal to zero, the limsup is equal to the limit and the limit of the quotient $Q(h)$ exists and is 0 as h goes to 0. By definition, f is Fréchet differentiable at x. □

3.3 Hadamard Differential and Semidifferential

By using the terminilogy *Hadamard differentiability* in Definition 3.1(ii), we have slightly cheated. This rather applies to the geometric notion of differential in the quotation of section 3.1 which is the original differential in the sense of Hadamard, while Definition 3.1(ii) is merely an equivalent characterization, as can be seen from the equivalence of (a) and (c) in part (ii) of the following theorem.

Theorem 3.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $x \in \mathbb{R}^n$.*

(i) *Given a direction $v \in \mathbb{R}^n$, the following conditions are equivalent:*

  (a) *$d_H f(x;v)$ exists;*

  (b) *there exists $g(x,v) \in \mathbb{R}$ such that for all paths $h : [0,\infty[\to \mathbb{R}^n$ for which $h(0) = x$ and $dh(0;+1) = v$, $d(f \circ h)(0;+1)$ exists and $d(f \circ h)(0;+1) = g(x,v)$.*

(ii) *The following conditions are equivalent:*

  (a) *f is Hadamard differentiable at x;*

  (b) *there exists a linear map $L(x) : \mathbb{R}^n \to \mathbb{R}$ such that for all paths $h : \mathbb{R} \to \mathbb{R}^n$ for which $h(0) = x$ and $dh(0;+1)$ exists, $d(f \circ h)(0;+1)$ exists and*

$$d(f \circ h)(0;+1) = L(x)\, dh(0;+1); \tag{3.16}$$

(c) *there exists a linear map* $L(x) : \mathbb{R}^n \to \mathbb{R}$ *such that for all paths* $h : \mathbb{R} \to \mathbb{R}^n$ *for which* $h(0) = x$ *and* $h'(0)$ *exists,* $(f \circ h)'(0)$ *exists and*

$$(f \circ h)'(0) = L(x)\,h'(0). \tag{3.17}$$

Proof. It is sufficient to prove the equivalence of (a) and (b) in part (i). The equivalence of (a) and (b) in part (ii) is a consequence of part (i). It will remain to prove the equivalence of (b) and (c) in part (ii) to complete the proof.

Part (i) (a) $\Rightarrow$ (b). As $dh(0;+1)$ exists, for any sequence $\{t_n > 0\}$ going to 0, we have

$$w_n \overset{\text{def}}{=} \frac{h(t_n) - h(0)}{t_n} \to dh(0;+1) = v.$$

But, since $d_H f(x;v)$ exists, we have for all sequences $\{t_n > 0\}$, $t_n \searrow 0$, $w_n \to v$,

$$\frac{f(h(t_n)) - f(h(0))}{t_n} = \frac{f(x + t_n w_n) - f(x)}{t_n} \to d_H f(x;v),$$

and $d(f \circ h)(0;+1) = d_H f(x;v)$ exists and we can choose $g(x,v) = d_H f(x;v)$.

(b) $\Rightarrow$ (a). By contradiction. Assume that there exist sequences $w_n \to v$ and $\{t_n > 0\}$, $t_n \searrow 0$, such that the sequence of differential quotients

$$q_n \overset{\text{def}}{=} \frac{f(x + t_n w_n) - f(x)}{t_n}$$

does not converge to $g(x,v)$. So, there exists $\eta > 0$ such that, for all $k \geq 1$, there exists $n_k \geq k$ such that $|q_{n_k} - g(x,v)| \geq \eta$. To simplify the notation, denote by (t_n, w_n) the subsequence (t_{n_k}, w_{n_k}).

Construct the following new subsequence (t_{n_k}, w_{n_k}). Let n_1 be the first $n \geq 1$ such that $t_n \leq 1$. Let n_2 be the first $n > n_1$ such that $t_n \leq t_{n_1}/2$. At step $k+1$, let n_{k+1} be the first $n > n_k$ such that $t_n \leq t_{n_k}/2$. By construction, $n_k > n_{k+1}$ and $t_{n_{k+1}} \leq t_{n_k}/2 < t_{n_k}$. The subsequence $\{(t_{n_k}, x_{n_k})\}$ is such that $\{t_{n_k}\}$ is monotone strictly decreasing to 0 and $w_{n_k} \to v$. As a result, we can assume, without loss of generality, that for our initial sequence $\{(t_n, w_n)\}$, the sequence $\{t_n\}$ is monotone strictly decreasing. As $\{w_n\}$ is convergent, there exists a constant c such that $|w_n| \leq c$ for all n. For all $\varepsilon > 0$, there exists N such that

$$\forall n > N, \quad \|w_n - v\| < \varepsilon \quad \text{and} \quad t_n < \varepsilon/c.$$

Define the function $h : [0,\infty[\to \mathbb{R}^n$ as follows:

$$h(t) \overset{\text{def}}{=} \begin{cases} x + tv, & \text{if } t \leq 0 \\ x + t\,w_n, & \text{if } t_n \leq t < t_{n-1}, \quad n \geq 2, \\ x + t\,w_1, & \text{if } t \geq t_1. \end{cases}$$

The vector function h is continuous at $t = 0$. Indeed, it is continuous from the left since $h(t) = x + tv \to x = h(0)$ as $t < 0$ goes to 0. On the right, for $\delta = t_{N+1} > 0$ and $0 < t < \delta$, there exists $n > N+1$ such that $t_n \leq t < t_{n-1}$ and hence

$$\|h(t) - h(0)\| = t\,\|w_n\| < t_{n-1}\,\|w_n\| < (\varepsilon/c)c = \varepsilon$$

and h is continuous from the right at 0.

For the derivative from the right, for $\delta = t_{N+1} > 0$ and $0 < t < \delta$, there exists $n > N+1$ such that $t_n \le t < t_{n-1}$ and

$$\left\| \frac{h(t)-h(0)}{t} - v \right\| = \|w_n - v\| < \varepsilon$$

and $dh(0;+1) = v$. For the derivative from the left, we have $dh(0;-1) = -v$, $-dh(0;-1) = v = dh(0;+1)$ and, in fact, the derivative exists and $h'(0) = v$. But, by hypothesis, for such a function h, $d(f \circ h)(0;+1)$ exists and is equal to $g(x;v)$. On the other hand, by construction of the function h,

$$q_n = \frac{f(x+t_n w_n) - f(x)}{t_n} = \frac{f(h(t_n)) - f(h(0))}{t_n} \to d(f \circ h)(0;+1) = g(x,v).$$

This contradicts our initial assumption that $q_n \not\to g(x,v)$.

Part (ii) (b) $\Rightarrow$ (c). Let h be such that $h(0) = x$ and $h'(0)$ exists. Then,

$$dh(0;+1) = h'(0) = -dh(0;-1)$$

and from (b),

$$d(f \circ h)(0;+1) = L(x)\,dh(0;+1) = L(x)\,h'(0). \tag{3.18}$$

The function $\bar{h}(t) \stackrel{\text{def}}{=} h(-t)$ is such that $\bar{h}(0) = x$ and $\bar{h}'(0) = -h'(0)$. By (b),

$$d(f \circ \bar{h})(0;+1) = L(x)\,d\bar{h}(0;+1) = -L(x)\,h'(0). \tag{3.19}$$

Therefore,

$$\lim_{t \searrow 0} \frac{f(h(0-t)) - f(h(0))}{t} = \lim_{t \searrow 0} \frac{f(\bar{h}(t)) - f(\bar{h}(0))}{t} = -L(x)\,h'(0)$$
$$\Rightarrow d(f \circ h)(0;-1) \text{ exists and } d(f \circ h)(0;-1) = -L(x)\,h'(0) = -d(f \circ h)(0;+1).$$

So we have existence of the derivative $(f \circ h)'(0)$ and $(f \circ h)'(0) = L(x)\,h'(0)$.

(c) $\Rightarrow$ (b). Let h be such that $h(0) = x$ and $dh(0;+1)$ exists. Construct the new function $\bar{h} : \mathbb{R} \to \mathbb{R}^n$ as follows:

$$\bar{h}(t) \stackrel{\text{def}}{=} \begin{cases} h(t), & \text{if } t \ge 0 \\ h(0) - t\,dh(0;+1), & \text{if } t < 0. \end{cases}$$

It is easy to check that $\bar{h}(0) = h(0) = x$ and that

$$d\bar{h}(0;+1) = dh(0;+1) \quad \text{and} \quad d\bar{h}(0;-1) = -dh(0;+1)$$
$$\Rightarrow \bar{h}'(0) \text{ exists and } \bar{h}'(0) = dh(0;+1).$$

For part (c),

$$d(f \circ \bar{h})(0;+1) = (f \circ \bar{h})'(0) = L(x)\,\bar{h}'(0) = L(x)\,dh(0;+1).$$

But for $t > 0$,

$$\frac{f(\bar{h}(t)) - f(\bar{h}(0))}{t} = \frac{f(h(0+t)) - f(h(0))}{t} \Rightarrow d(f \circ \bar{h})(0;+1) = d(f \circ h)(0;+1)$$
$$\Rightarrow d(f \circ h)(0;+1) = L(x)\, dh(0;+1)$$

and (c) is true. □

From the abstract of the paper entitled "*Sur la notion de différentielle*" of M. FRÉCHET [**3**] in 1937

> The author shows that the total differential of Stolz-Young[31] is equivalent to the definition due to Hadamard (Theorem 3.2 (ii) (c)). On the other hand, when the latter is extended to functionals it becomes more general that the one of the author ...[32]

since it applies to function spaces (infinite dimension) without metric. In the same paper, M. FRÉCHET [**3**, p. 239] proposes the following definition.

Definition 3.5 (Notion proposed by Fréchet)**.**
The function $f : \mathbb{R}^n \to \mathbb{R}$ is *differentiable* at $x \in \mathbb{R}^n$ if there exists a function $g(x) : \mathbb{R}^n \to \mathbb{R}$ such that for all paths $h : \mathbb{R} \to \mathbb{R}^n$ for which $h(0) = x$ and $h'(0)$ exists, we have

$$(f \circ h)'(0) = g(x)(h'(0)). \tag{3.20}$$

□

By Theorem 3.2(i), Definition 3.5 implies (b) that implies (a), and, from there, the function is Hadamard semidifferentiable; that is, the semidifferential $d_H f(x;v)$ exists for all $v \in \mathbb{R}^n$. By using $h'(0)$ and $(f \circ h)'(0)$ rather than $dh(0;+1)$ and $d(f \circ h)(0;+1)$, Fréchet was losing some Hadamard semidifferentiable functions (in the sense of Definition 3.1(ii)) such as the norm $f(x) = \|x\|$ on $\mathbb{R}^n$ at $x = 0$ since the differential quotient

$$\frac{\|0 + t_n v_n\| - \|0\|}{t_n} = \frac{|t_n|}{t_n}\|v_n\| \tag{3.21}$$

does not converge as $v_n \to v$ and $t_n \to 0$ (the sequence $|t_n|/t_n$ has subsequences converging to 1 and -1). It is really necessary to use sequences $\{t_n\}$ of positive numbers to get the convergence of the differential quotient

$$d_H f(0;v) = \lim_{\substack{t_n \searrow 0 \\ v_n \to v}} \frac{|t_n|}{t_n}\|v_n\| = \lim_{v_n \to v} \|v_n\| = \|v\|.$$

So, Definition 3.5 is slightly stronger than Definition 3.1(ii) of $d_H f(x;v)$. Yet, it is quite remarkable since, up to the use of the semiderivatives $dh(0;+1)$ and $d(f \circ h)(0;+1)$ rather than $h'(0)$ and $(f \circ h)'(0)$, he introduces a nondifferentiable infinitesimal calculus. We

[31] See Remark 3.2 on page 88.

[32] From the French: L'auteur montre que la différentielle totale de Stolz-Young est équivalente à la définition due à Hadamard (Theorem 3.2 (ii) (c)). Par contre, quand on étend cette dernière aux fonctionnelles elle devient plus générale que celle de l'auteur ...

shall see that such functions retain two important properties of differentiable functions: they are continuous at x (Theorem 3.7) and the chain rule is applicable to the composition (Theorem 3.4).

Unfortunately, yielding to criticisms, he does not push this new notion further.

> "But as was pointed out by Mr Paul Lévy, such a definition is not sufficient, since a differentiable function in that sense could loose important properties of the differential of simple functions and in particular property (3) (the linearity!). Such is, for instance, the case for the function
>
> $$f(x,y) = x\sqrt{\frac{x^2}{x^2+y^2}} \quad \text{for } (x,y) \neq (0,0) \quad \text{with } f(0,0) = 0.\text{"} \tag{3.22}$$
>
> (M. FRÉCHET [**3**, p. 239])[33]

Indeed, it is readily checked that for $(v,w) \neq (0,0)$,

$$d_H f((0,0);(v,w)) = \lim_{\substack{t_n \to 0 \\ (v_n,w_n)\to(v,w)}} \frac{f(t_n\, v_n, t_n\, w_n) - f(0,0)}{t_n} = v\sqrt{\frac{v^2}{v^2+w^2}}$$

is not linear in (v,w). Far from discrediting this new notion, his example shows that such functions exist.

Definition 3.1(i) of the semidifferential $df(x;v)$ and of the differential can be found in the *posthumous work*[34] of R. GATEAUX [**6**, **7**] published in 1919 and 1922 up to the difference that he is using directional derivatives in Definition 3.2, where the variable t goes to 0, rather than the semidifferential where the variable t goes to 0 by positive values as U. DINI [**1**] was already doing in 1878. This notion would have been inspired to him during a visit to V. Volterra[35] in the early times of the calculus of variations.

Definition 3.1(ii) of the Hadamard semidifferential $d_H f(x;v)$ and of the differential can be found in J.-P. PENOT [**1**, p. 250] in 1978 under the name M-semidifferentiability and M-differentiability. He attributes M-differentiability to A. D. MICHAL [**1**] in 1938 for infinite-dimensional spaces. A. D. MICHAL [**1**] makes the connection between his M-differentiability and what he calls the MH-differentiability that is nothing but the geometric version of the Hadamard differential in infinite dimension (see part (c) of Theorem 3.2(ii)) that Fréchet had already introduced and promoted in his 1937 paper. A. D. MICHAL [**1**]

[33] From the French: "Mais comme l'a fait observer M. Paul Lévy, une telle définition n'est pas suffisante, car une fonction différentiable à ce sens peut perdre d'importantes propriétés de la différentielle des fonctions simples et en particulier la propriété (3) (la linéarité!). Tel est, par exemple, le cas pour la fonction

$$f(x,y) = x\sqrt{\frac{x^2}{x^2+y^2}} \quad \text{pour } (x,y) \neq (0,0) \quad \text{avec } f(0,0) = 0.\text{"} \tag{3.22}$$

(M. FRÉCHET [**3**, p. 239])

[34] "... Nous allons emprunter au Calcul functionnel la notion de variation, qui nous rendra les services que rend la differential totale dans la théorie des functions d'un nombre fini de variables. $\delta F(x) = [\frac{d}{d\lambda} F(x + \lambda\,\delta x)]_{\lambda=0}$ (see R. GATEAUX [**7**, page 83]). ..."

"... Considérons $U(z+\lambda\, t_1)$. Supposons que $[\frac{d}{d\lambda} U(z+\lambda\, t_1)]_{\lambda=0}$ existe quelque soit t_1. On l'appelle la *variation première* de U au point z : $\delta U(z,t_1)$. C'est une function de z et t_1, qu'on suppose habituellement linéaire, en chaque point z par rapport à t_1." (see R. GATEAUX [**6**, page 11]).

[35] Vito Volterra (1860–1940).

also introduces the equivalent in infinite dimension of the semidifferential of Definition 3.5 that Fréchet had proposed in 1937.

However, as pointed out in (3.21), the norm is not semidifferentiable according to the definitions of the semidifferential introduced by Fréchet and Michal, but it becomes semidifferentiable for the slightly weaker definition introduced by Penot (our Definition 3.1(ii) of $d_H f(x;v)$).[36] Keeping all this in mind, we have opted for the terminology Hadamard semidifferential rather than M-semidifferential or Penot semidifferential.

3.4 Operations on Semidifferentiable Functions

3.4.1 Algebraic Operations, Lower and Upper Envelopes

Theorem 3.3. *Let $x \in \mathbb{R}^n$ be a point and $v \in \mathbb{R}^n$ be a direction.*

(i) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be such that $df(x;v)$ and $dg(x;v)$ exist. Then*

$$d(f+g)(x;v) = df(x;v) + dg(x;v) \tag{3.23}$$
$$d(fg)(x;v) = df(x;v)\, g(x) + f(x)\, dg(x;v) \tag{3.24}$$

by defining

$$(f+g)(x) \overset{\text{def}}{=} f(x) + g(x), \quad \forall x \in \mathbb{R}^n \tag{3.25}$$
$$(fg)(x) \overset{\text{def}}{=} f(x)\, g(x), \quad \forall x \in \mathbb{R}^n. \tag{3.26}$$

If $\alpha \in \mathbb{R}$, then

$$d(\alpha f)(x;v) = \alpha\, df(x;v)$$

by defining

$$(\alpha f)(x) \overset{\text{def}}{=} \alpha f(x).$$

(ii) *Let I be a* finite *set of indices, $\{f_i : i \in I\}$, $f_i : \mathbb{R}^n \to \mathbb{R}$, a family of functions such that $df_i(x;v)$ exists, and the* upper envelope

$$h(x) \overset{\text{def}}{=} \max_{i \in I} f_i(x), \quad x \in \mathbb{R}^n. \tag{3.27}$$

Then

$$dh(x;v) = \max_{i \in I(x)} df_i(x;v), \text{ where } I(x) \overset{\text{def}}{=} \{i \in I \,:\, f_i(x) = h(x)\}. \tag{3.28}$$

Similarly, for the lower envelope

$$k(x) \overset{\text{def}}{=} \min_{i \in I} f_i(x), \quad x \in \mathbb{R}^n \tag{3.29}$$

we have

$$dk(x;v) = \min_{i \in J(x)} df_i(x;v), \text{ where } J(x) \overset{\text{def}}{=} \{i \in I \,:\, f_i(x) = k(x)\}. \tag{3.30}$$

[36] To our best knowledge, J.-P. PENOT [**1**, p. 250] seems to have been the first author to explicitly introduce Definition 3.1(ii).

Corollary 1. *Under the assumptions of Theorem* 3.3, *all the operations are verified for Hadamard semidifferentiable functions.*

Proof. (i) By definition.

(ii) It is sufficient to prove the result for two functions f_1 and f_2

$$h_1(x) = \max\{f_1(x), f_2(x)\}.$$

From this, we can go to three functions by letting

$$h_2(x) = \max\{h_1(x), f_3(x)\} = \max\{f_i(x) \,|\, 1 \le i \le 3\},$$

and so on. Then, repeat the construction for a finite number of functions. We distinguish three cases (see Figures 3.7 and 3.8):

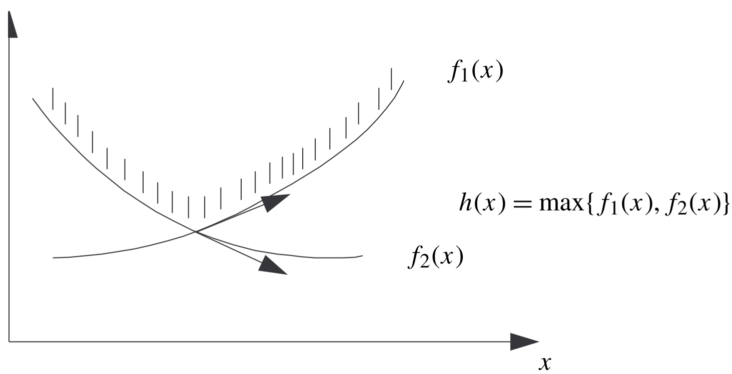


Figure 3.7. *Upper envelope of two functions.*

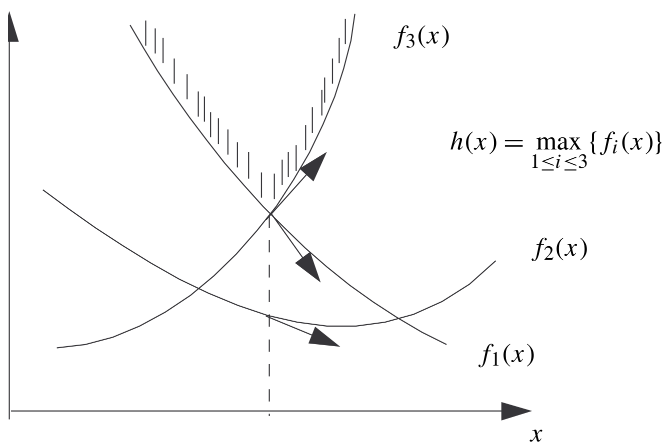


Figure 3.8. *Upper envelope of three functions.*

(a) $f_1(x) = f_2(x)$, (b) $f_1(x) > f_2(x)$, (c) $f_1(x) < f_2(x)$.

(a) $\underline{f_1(x) = f_2(x)}$. On a $h_1(x) = f_1(x) = f_2(x)$. For $t > 0$,

$$\frac{h_1(x+tv) - h_1(x)}{t} = \frac{\max\{f_1(x+tv), f_2(x+tv)\} - h_1(x)}{t}$$
$$= \max\left\{\frac{f_1(x+tv) - f_1(x)}{t}, \frac{f_2(x+tv) - f_2(x)}{t}\right\}.$$

As both limits exist,

$$dh(x;v) = \max\{df_1(x;v), df_2(x;v)\}.$$

(b) $\underline{f_1(x) > f_2(x)}$. Then $h_1(x) = f_1(x)$ and, by continuity of $f_1(x+tv)$ and $f_2(x+tv)$ with respect to $t > 0$, there exists $\bar{t} > 0$ such that

$$f_1(x+tv) > f_2(x+tv), \quad \forall 0 < t < \bar{t}.$$

Then, for $0 < t < \bar{t}$,

$$\frac{h_1(x+tv) - h_1(x)}{t} = \frac{f_1(x+tv) - f_1(x)}{t} \to df_1(x;v).$$

(c) $\underline{f_1(x) < f_2(x)}$. Repeat (b) by interchanging the indices. □

This key theorem better motivates the introduction of semidifferentials. Indeed, the classical notion of differential fails when applied to the max of two (however smooth) functions at points where both functions are "active" and it becomes necessary to weaken the notion of differential to the one of semidifferential.

3.4.2 Chain Rule for the Composition of Functions

One important operation in a good differential calculus is the *chain rule* for the *composition* h of two functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}^n$, where m and n are positive integers,

$$\begin{array}{l|l} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^n \xrightarrow{f} \mathbb{R} & \displaystyle\frac{\partial h}{\partial x_j}(x) = \sum_{k=1}^{n} \frac{\partial f}{\partial y_k}(g(x)) \frac{\partial g_k}{\partial x_j}(x),\ 1 \le j \le m. \\ x \mapsto h(x) \stackrel{\text{def}}{=} f(g(x)) : \mathbb{R}^m \to \mathbb{R}, & \end{array}$$

An analogous result holds for semidifferentials: given $v = \sum_{j=1}^{m} v_j e_j$,

$$\sum_{j=1}^{m} \frac{\partial h}{\partial x_j}(x) v_j = \sum_{j=1}^{m}\sum_{k=1}^{n} \frac{\partial f}{\partial y_k}(g(x)) \frac{\partial g_k}{\partial x_j}(x) v_j = \sum_{k=1}^{n} \frac{\partial f}{\partial y_k}(g(x)) \sum_{j=1}^{m} \frac{\partial g_k}{\partial x_j}(x) v_j$$
$$\Rightarrow \quad dh(x;v) = \sum_{k=1}^{n} \frac{\partial f}{\partial y_k}(g(x)) dg_k(x;v) = df\left(g(x); \sum_{k=1}^{n} dg_k(x;v) e_k\right)$$
$$\Rightarrow \quad dh(x;v) = df(g(x); dg(x;v)),$$

where $g = (g_1, \ldots, g_n)$ and $dg(x;v) = (dg_1(x;v), \ldots, dg_n(x;v))$ is the extension of the notion of semidifferential to vector-valued functions $g : \mathbb{R}^m \to \mathbb{R}^n$.

Definition 3.6.
Let $m \geq 1$, $n \geq 1$ be two integers, $g : \mathbb{R}^m \to \mathbb{R}^n$ be a vector-valued function, $x \in \mathbb{R}^m$ be a point, and $v \in \mathbb{R}^m$ be a direction.

(i) g is (Gateaux) *semidifferentiable* at x in the direction v if

$$\lim_{t \searrow 0} \frac{g(x+tv) - g(x)}{t} \text{ exists in } \mathbb{R}^n.$$

(ii) g is *Hadamard semidifferentiable* at x in the direction v if

$$\lim_{\substack{t \searrow 0 \\ w \to v}} \frac{g(x+tw) - g(x)}{t} \text{ exists in } \mathbb{R}^n.$$

(iii) g is *Gateaux differentiable* at x if $dg(x;v)$ exists for all $v \in \mathbb{R}^m$ and the map

$$v \mapsto dg(x;v) : \mathbb{R}^m \to \mathbb{R}^n$$

is linear. Define the linear map

$$v \mapsto Dg(x)v \stackrel{\text{def}}{=} dg(x;v) : \mathbb{R}^m \to \mathbb{R}^n$$

that will be referred to as the *Jacobian mapping*[37] of g at x. □

The above definitions are equivalent to the definitions for real-valued functions applied to each component $g_i(x) = g(x) \cdot e_i^n$, $1 \leq i \leq n$, of the vector function $g = (g_1, \ldots, g_n) : \mathbb{R}^m \to \mathbb{R}^n$.

Definition 3.7.
Let $m \geq 1$, $n \geq 1$ be two integers, $g : \mathbb{R}^m \to \mathbb{R}^n$ be a vector-valued function, $x \in \mathbb{R}^m$ be a point, and $v \in \mathbb{R}^m$ be a direction.

(i) g is (Gateaux) *semidifferentiable* at x in the direction v if for each component g_i, $dg_i(x;v)$ exists.

(ii) g is *Hadamard semidifferentiable* at x in the direction v if for each component g_i, $d_H g_i(x;v)$ exists.

(iii) g is *Gateaux differentiable* at x if each component g_i of g is Gateaux differentiable at x. The mapping

$$v \mapsto Dg(x)v \stackrel{\text{def}}{=} \sum_{i=1}^{n} dg_i(x;v)\, e_i^n : \mathbb{R}^m \to \mathbb{R}^n$$

is called the *Jacobian mapping* of g at x and the associated $n \times m$ matrix

$$Dg(x)_{ij} \stackrel{\text{def}}{=} dg_i(x;e_j^m), \quad 1 \leq i \leq n, 1 \leq j \leq m,$$

is called the *Jacobian matrix*, where $\{e_j^m : j = 1, \ldots, m\}$ and $\{e_i^n : i = 1, \ldots, n\}$ are the respective canonical orthonormal bases in $\mathbb{R}^m$ and $\mathbb{R}^n$. □

[37] Carl Gustav Jacob Jacobi (1804–1851), brother of the physicist Moritz Hermann von Jacobi. He established the theory of functional determinants, that are now called *Jacobians*.

Remark 3.3.
When $m = n$, the absolute value of the determinant of the matrix $Dg(x)$ is called the *Jacobian* of g at x. □

Theorem 3.4 (Semidifferential of the composition of two functions)**.** *Let $n \geq 1$ and $m \geq 1$ be two integers, $g : \mathbb{R}^m \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be two functions, x be a point in $\mathbb{R}^m$, and v be a direction in $\mathbb{R}^m$. Consider the composition $(f \circ g)(x) = f(g(x))$. Assume that*

(a) *$dg(x;v)$ exists in $\mathbb{R}^n$,*

(b) *$d_H f(g(x);dg(x;v))$ exists in $\mathbb{R}$.*

Then,

(i) *$d(f \circ g)(x;v)$ exists and*

$$d(f \circ g)(x;v) = d_H f(g(x);dg(x;v)); \tag{3.31}$$

(ii) *if, in addition, $d_H g(x;v)$ exists, then $d_H(f \circ g)(x;v)$ exists and*

$$d_H(f \circ g)(x;v) = d_H f(g(x);d_H g(x;v)). \tag{3.32}$$

Corollary 1. *If, in addition, f is Gateaux differentiable at $g(x)$ and g is Gateaux differentiable at x, then*[38]

$$\forall v \in \mathbb{R}^m, \quad \nabla(f \circ g)(x) \cdot v = \nabla f(g(x)) \cdot Dg(x)v. \tag{3.33}$$

The result can also be written in matrix form

$$\underbrace{\nabla(f \circ g)(x)}_{m\times 1} = \underbrace{[Dg(x)]^\top}_{m\times n} \underbrace{\nabla f(g(x))}_{n\times 1}, \tag{3.34}$$

where $Dg(x)$ is the Jacobian $m \times n$ matrix *of $g : \mathbb{R}^m \to \mathbb{R}^n$,*

$$[Dg(x)]_{ij} = \partial_j g_i(x), \;\; 1 \leq i \leq n, \;\; 1 \leq j \leq m, \tag{3.35}$$

and $\nabla f(g(x))$ is considered as a column vector *or an $n \times 1$ matrix. If $\nabla f(g(x))$ is considered as a* line vector *or a $1 \times n$ matrix, the formula takes the form*

$$\underbrace{\nabla(f \circ g)(x)}_{1\times m} = \underbrace{\nabla f(g(x))}_{1\times n} \underbrace{Dg(x)}_{n\times m}.$$

Remark 3.4.
Theorem 3.4 extends to the composition h of two vector-valued functions $g : \mathbb{R}^m \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^p$, where m, n, and p are positive integers,

$$\mathbb{R}^m \xrightarrow{g} \mathbb{R}^n \xrightarrow{f} \mathbb{R}^p, \quad x \mapsto h(x) \overset{\text{def}}{=} f(g(x)) : \mathbb{R}^m \to \mathbb{R}^p$$
$$dh(x;v) = df(g(x);dg(x;v))$$

[38]We shall see later that f being Hadamard semidifferentiable plus Gateaux differentiable is equivalent to f being Fréchet differentiable (see Definition 3.4 and Theorem 3.1 of section 3.2.4).

under the assumption that $d_H f(g(x);dg(x;v))$ exists in $\mathbb{R}^p$. When the functions are Hadamard differentiable,

$$\frac{\partial h_i}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k}(g(x)) \frac{\partial g_k}{\partial x_j}(x), \quad 1 \le i \le p, 1 \le j \le m,$$

$$Dh(x) = Df(g(x))\,Dg(x).$$

Similarly, it is possible to consider the composition of a finite number of functions. For instance, the Hadamard semidifferential of the composition of three functions $g_1 \circ g_2 \circ g_3$ that are Hadamard semidifferentiable is given by

$$d_H(g_1 \circ g_2 \circ g_3)(x;v) = d_H g_1(g_2(g_3(x)); d_H g_2(g_3(x); d_H g_3(x;v))).$$

All this extends to *tensors*. □

Proof of Theorem 3.4. (i) For $t > 0$, consider the limit of the differential quotient

$$q(t) \overset{\text{def}}{=} \frac{1}{t}[f(g(x+tv)) - f(g(x))].$$

The term $f(g(x+tv))$ can be written as

$$g(x+tv) = g(x) + t\,\frac{g(x+tv)-g(x)}{t} = g(x) + t\,v(t)$$

by introducing the function

$$v(t) \overset{\text{def}}{=} \frac{g(x+tv)-g(x)}{t}.$$

By assumption of the existence of $dg(x;v)$,

$$v(t) \to dg(x;v) \text{ as } t \searrow 0^+.$$

The differential quotient is now of the form

$$q(t) = \frac{1}{t}[f(g(x) + t\,v(t)) - f(g(x))].$$

By definition and existence of $d_H f(g(x);dg(x;v))$, the limit exists and

$$\lim_{t \searrow 0^+} q(t) = d_H f(g(x);dg(x;v)).$$

(ii) For $t > 0$ and $w \in \mathbb{R}^m$, consider the differential quotient

$$q(t,w) \overset{\text{def}}{=} \frac{1}{t}[f(g(x+tw)) - f(g(x))].$$

The term $f(g(x+tw))$ can be rewritten in the form

$$g(x+tw) = g(x) + t\,\frac{g(x+tw)-g(x)}{t} = g(x) + t\,v(t,w)$$

by introducing the function

$$v(t,w) \stackrel{\text{def}}{=} \frac{g(x+tw)-g(x)}{t}.$$

But, by assumption, $d_H g(x;v)$ exists and

$$v(t,w) \to d_H g(x;v) \text{ as } t \searrow 0^+ \text{ and } w \to v.$$

The differential quotient is now of the form

$$q(t,w) = \frac{1}{t}[f(g(x)+t\,v(t,w)) - f(g(x))].$$

By definition and existence of $d_H f(g(x);d_H g(x;v))$, the limit exists and

$$\lim_{\substack{t\searrow 0^+\\ w\to v}} q(t,w) = d_H f(g(x);d_H g(x;v)). \qquad \square$$

Remark 3.5.
It is important to recall that the semidifferential $d_H f(x;v)$ need not be linear in v as illustrated by the function $f(x) = \|x\|$ for which $d_H f(0;v) = \|v\|$. $\square$

The weaker assumption that $dg(x;v)$ and $df(g(x);dg(x;v))$ exist is not sufficient to prove the theorem. The proof critically uses the stronger assumption that $d_H f(g(x); dg(x;v))$ also exists. This is now illustrated by an example of the composition $f \circ g$ of a Gateaux differentiable function f and an infinitely differentiable function[39] g. The composition $f \circ g$ is not Gateaux differentiable and not even Gateaux semidifferentiable at 0 in any direction $v \neq 0$.

Example 3.9.
Consider the functions

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \frac{x^6}{(y-x^2)^2 + x^8} \text{ if } (x,y) \neq (0,0) \text{ and } f(0,0) = 0$$

$$g : \mathbb{R} \to \mathbb{R}^2, \quad g(x) = \begin{bmatrix} x \\ x^2 \end{bmatrix}.$$

We have seen in Example 3.6 that f is Gateaux differentiable in $(0,0)$ and that

$$\nabla f(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

It is readily seen that g is infinitely continuously differentiable or of class $C^{(\infty)}$ in $\mathbb{R}$ and that the associated Jacobian matrix is

$$Dg(x) = \begin{bmatrix} 1 \\ 2x \end{bmatrix}.$$

The composition of f and g is given by

$$h(x) = f(g(x)) = \begin{cases} 1/x^2, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

[39] A function that is differentiable as well as all its partial derivatives of all orders.

By applying the chain rule, we get

$$h'(0) = [Dg(0)]^\top \nabla f(g(0)) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0.$$

However, this result is false since the composition $h(x) = f(g(x))$ as a real-valued function of the real variable x is neither continuous at 0 nor right or left continuous at 0. So, it is not (Gateaux) semidifferentiable at $x = 0$ in any direction $v \neq 0$. It illustrates the fact that the Gateaux differentiability of f is not sufficient. It would require the Hadamard semidifferentiability or differentiability of f at $(0,0)$. □

3.5 Lipschitzian Functions

3.5.1 Definitions and Their Hadamard Semidifferential

Definition 3.8.
Let $f : \mathbb{R}^n \to \mathbb{R}$.

(i) f is *Lipschitzian* at x if there exist $c(x) > 0$ and a neighborhood $V(x)$ of x such that

$$\forall y, z \in V(x), \quad |f(z) - f(y)| \leq c(x)\|z - y\|.$$

(ii) f is *locally Lipschitzian* on a subset U of $\mathbb{R}^n$ if, for each $x \in U$, there exist $c(x) > 0$ and a neighborhood $V(x)$ of x such that

$$\forall y, z \in V(x) \cap U, \quad |f(z) - f(y)| \leq c(x)\|z - y\|.$$

(iii) f is *Lipschitzian* on a subset U of $\mathbb{R}^n$ if there exists $c(U) > 0$ such that

$$\forall y, z \in U, \quad |f(z) - f(y)| \leq c(U)\,\|z - y\|.$$ □

Example 3.10.
The norm $f(x) = \|x\|$ is Lipschitzian on $\mathbb{R}^n$ since

$$\forall y, z \in \mathbb{R}^n, \quad |f(y) - f(z)| = |\|y\| - \|z\|| \leq \|y - z\|,$$

with Lipschitz constant $c(\mathbb{R}^n) = 1$. The function $f(x) = \|x\|^2$ is locally Lipschitzian on $\mathbb{R}^n$ since for all $x \in \mathbb{R}^n$, $r > 0$,

$$\forall y, z \in B_r(x), \quad |f(y) - f(z)| = \left| \|y\|^2 - \|z\|^2 \right| \leq \|y + z\|\,\|y - z\| \leq 2(r + \|x\|)\|y - z\|,$$

with the local constant $c(x) = 2(r + \|x\|)$. □

Theorem 3.5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian at $x \in \mathbb{R}^n$.*

(i) *If $df(x;v)$ exists, then $d_H f(x;v)$ exists. In particular, $d_H f(x;0)$ exists.*

(ii) *If $df(x;v)$ exists for all $v \in \mathbb{R}^n$, then $d_H f(x;v)$ exists for all $v \in \mathbb{R}^n$ and*

$$\forall v, w \in \mathbb{R}^n, \quad |d_H f(x;v) - d_H f(x;w)| \leq c(x)\,\|v - w\|, \tag{3.36}$$

where $c(x)$ is the Lipschitz constant associated with the neighborhood $V(x)$ of x in Definition 3.8(i).

Proof. (i) Consider an arbitrary sequence $\{w_n\}$ converging to v, $t > 0$, and the differential quotient

$$\frac{f(x+tw_n)-f(x)}{t} = \frac{f(x+tw_n)-f(x+tv)}{t} + \frac{f(x+tv)-f(x)}{t}.$$

By assumption, the second term goes to $df(x;v)$ as $t \searrow 0$. As f is Lipschitzian at x,

$$\exists c(x), \forall y,z \in V(x), \quad |f(y)-f(z)| \le c(x)\|y-z\|$$

and the absolute value of the first term can be estimated:

$$\left|\frac{f(x+tw_n)-f(x+tv)}{t}\right| \le c(x)\|w_n - v\| \to 0 \text{ as } n \to \infty.$$

As the limit 0 is independent of the choice of the sequence $w_n \to v$, by definition of the Hadamard semiderivative, $d_H f(x;v)$ exists and

$$d_H f(x;v) = df(x;v), \;\; \forall v \in \mathbb{R}^n.$$

For $v = 0$, $df(x;0) = 0$ always exists. Hence, $d_H f(x;0)$ exists.

(ii) For all $v, w \in \mathbb{R}^n$ and $t > 0$ sufficiently small, $x + tv$ and $x + tw$ belong to $V(x)$ and

$$\left|\frac{f(x+tv)-f(x)}{t} - \frac{f(x+tw)-f(x)}{t}\right| \le c(x)\,\|v-w\|.$$

As t goes to zero, $|d_H f(x;v) - d_H f(x;w)| = |df(x;v) - df(x;w)| \le c(x)\,\|v-w\|$. □

Example 3.11.
Since the norm $f(x) = \|x\|$ is Lipschitzian on $\mathbb{R}^n$ and $df(x;v)$ exists, so does $d_H f(x;v)$. □

3.5.2 ▶ Dini and Hadamard Upper and Lower Semidifferentials

In order to get the existence of $d_H f(x;v)$, the existence of $df(x;v)$ is necessary. In general, this is not true for an arbitrary Lipschitzian function. For instance, $d_H d_U(x;v)$ exists for a convex set U since d_U is convex and uniformly Lipschitzian on $\mathbb{R}^n$, but the semidifferential might not exist at some points for a lousy set U. However, as f is Lipschitzian at x, the differential quotient

$$\left|\frac{f(x+tv)-f(x)}{t}\right| \le c(x)\,\|v\|$$

is bounded as t goes to zero and the liminf and limsup exist and are finite. In fact such limits exist for a much larger class of functions.

Definition 3.9. (i) Given a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $x \in \text{dom}\, f$,

$$\underline{d}f(x;v) \overset{\text{def}}{=} \liminf_{t \searrow 0^+} \frac{f(x+tv)-f(x)}{t}$$

will be referred to as the *Dini lower semidifferential*.[40]

[40] Ulisse Dini (1845–1918), U. Dini [**1**].

(ii) Given a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ and $x \in \text{dom } f$,

$$\overline{d} f(x;v) \overset{\text{def}}{=} \limsup_{t \searrow 0^+} \frac{f(x+tv) - f(x)}{t}$$

will be referred to as the *Dini upper semidifferential*.[41] □

By definition, $\overline{d} f(x;v) = -\underline{d}(-f)(x;v)$, so that it is sufficient to study the lower notion to get the properties of the upper notion.

In fact, when f is Lipschitzian at x, we even get something stronger since the differential quotient

$$\left| \frac{f(x+tw) - f(x)}{t} \right| \le c(x) \left(\|v\| + \|w - v\| \right)$$

is also bounded as $t \to 0$ and $w \to v$.

Definition 3.10. (i) Given a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $x \in \text{dom } f$,

$$\underline{d}_H f(x;v) \overset{\text{def}}{=} \liminf_{\substack{t \searrow 0^+ \\ w \to v}} \frac{f(x+tw) - f(x)}{t}$$

will be referred to as the *Hadamard lower semidifferential*.

(ii) Given a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ and $x \in \text{dom } f$,

$$\overline{d}_H f(x;v) \overset{\text{def}}{=} \limsup_{\substack{t \searrow 0^+ \\ w \to v}} \frac{f(x+tw) - f(x)}{t}$$

will be referred to as the *Hadamard upper semidifferential*. □

Again, the upper notion can be obtained from the lower one by observing that

$$\overline{d}_H f(x;v) = -\underline{d}_H(-f)(x;v). \tag{3.37}$$

When f is Lipschitzian at x, the Dini and Hadamard notions coincide and the semidifferentials are finite. This is in line with previous notions such as the lower and upper semicontinuities.

3.5.3 ► Clarke Upper and Lower Semidifferentials

Lipschitzian functions were one of the motivations behind the introduction of another semidifferential by F. H. CLARKE [**1**] in 1973. He considered the differential quotient

$$\frac{f(y+tv) - f(y)}{t}.$$

[41] *Upper* and *lower Dini derivatives* in P. CANNARSA and C. SINESTRARI [**1**, Dfn. 3.1.3, p. 50].

For a Lipschitzian function at x, this quotient is bounded as $t \searrow 0$ and $y \to x$,

$$\left|\frac{f(y+tv)-f(y)}{t}\right| \le c(x)\,\|v\|,$$

and both liminf and limsup exist and are finite. Here, it is not sufficient that $x \in \text{dom } f$ to make sense of the difference $f(y+tv)-f(y)$ entering the differential quotient in a neighborhood of x as can be seen from the convex function of Example 9.1 of Chapter 2 at the point $(0,0)$ of its effective domain.

Definition 3.11.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschizian at x and $v \in \mathbb{R}^n$ be a direction.[42]

(i) The *Clarke upper semidifferential* of f at x in the direction v is[43]

$$\overline{d}_C f(x;v) \overset{\text{def}}{=} \limsup_{\substack{t\searrow 0^+\\ y\to x}} \frac{f(y+tv)-f(y)}{t}. \tag{3.38}$$

(ii) The *Clarke lower semidifferential* of f at x in the direction v is[44]

$$\underline{d}_C f(x;v) \overset{\text{def}}{=} \liminf_{\substack{t\searrow 0^+\\ y\to x}} \frac{f(y+tv)-f(y)}{t}. \tag{3.39}$$

Again, the lower notion can be obtained from the upper one by observing that

$$\underline{d}_C f(x;v) = -\overline{d}_C(-f)(x;v). \tag{3.40}$$

□

For Lipschitzian functions at x, the upper and lower Clarke semiderivatives are relaxations of the notion of *strict differentiability*[45] at x that will not be used in this book.

[42] For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the definition of Lipschitzian at x would require that f be finite in a neighborhood of x, that is, $x \in \text{int}(\text{dom } f)$.

[43] *Generalized directional derivative* in F. H. Clarke [**2**, p. 10].

[44] *Generalized lower derivative* in P. Cannarsa and C. Sinestrari [**1**, Dfn. 3.1.11, p. 54].

[45] A function f Lipschitzian at x is *strictly differentiable* at x (see, for instance, J. M. Borwein and A. S. Lewis1[**1**, p. 132]) if

$$\exists L(x) : \mathbb{R}^n \to \mathbb{R}^n \text{ linear such that } \forall v \in \mathbb{R}^n,\ \lim_{\substack{t\searrow 0^+\\ y\to x}} \frac{f(y+tv)-f(y)}{t} = L(x)v. \tag{3.41}$$

It is stronger than Fréchet differentiability as can be seen from the following example of the function:

$$f(x) \overset{\text{def}}{=} x^2 \sin(1/x),\ x \neq 0, \text{ and } f(0)=0.$$

It is clearly differentiable at $x \neq 0$. At $x = 0$, consider the differential quotient for $v \neq 0$

$$\frac{f(tv)-f(0)}{t} = \frac{(tv)^2 \sin(1/(tv)) - 0}{t} = t\,v^2 \sin(1/(tv)) \to 0.$$

It is not strictly differentiable in $x = 0$. Choose $v = 1$ and the sequence of points $y_n = 1/[(n+1/2)\pi] \to 0$ and $t_n = 1/[(n+1/2)\pi] - 1/[(n+3/2)\pi] \searrow 0$. Then $(f(y_n+t_n v)-f(y_n))/t_n$ oscillates towards $\pm 1/\pi$ as $n \to \infty$.

Remark 3.6.
We shall see in section 4.4 that five of the six notions of Definitions 3.9, 3.10, and 3.11 can coincide when the Hadamard semidifferential exists: for semiconvex functions,

$$\forall x \in \text{dom } f, \quad df(x;v) = d_H f(x;v) = \overline{d}_C f(x;v) = \limsup_{\substack{t \searrow 0^+ \\ y \to x}} \frac{f(y+tv) - f(y)}{t} \tag{3.42}$$

and for semiconcave functions,

$$\forall x \in \text{dom } f, \quad df(x;v) = d_H f(x;v) = \underline{d}_C f(x;v) = \liminf_{\substack{t \searrow 0^+ \\ y \to x}} \frac{f(y+tv) - f(y)}{t}. \tag{3.43}$$

In general, the existence of $d_H f(x;v)$ does not imply that $\underline{d}_C f(x;v) = d_H f(x;v) = \overline{d}_C f(x;v)$ as can be seen from the following example. Consider the function $f(x) = |x|$ in $\mathbb{R}$. We know that $d_H f(0;v) = |v|$. Choose the direction $v = 1$ and the sequences $y_n = (-1)^n/n$ and $t_n = 1/n$. The differential quotient

$$\frac{|y_n + t_n v| - |y_n|}{t_n} = (-1)^n$$

oscillates between $+1$ and -1 as $n \to \infty$ and

$$\underline{d}_C f(0;1) = -1 < d_H f(0;1) = |1| = 1 = \overline{d}_C f(0;1). \qquad \square$$

3.5.4 ► Properties of Upper and Lower Subdifferentials

The next theorem gives the properties and relations between the above semidifferentials for Lipschitzian functions.

Theorem 3.6. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian at $x \in \mathbb{R}^n$.*

(i) *For all $v \in \mathbb{R}^n$,*

$$\underline{d}_C f(x;v) \le \underline{d}_H f(x;v) = \underline{d} f(x;v) \le \overline{d} f(x;v) = \overline{d}_H f(x;v) \le \overline{d}_C f(x;v). \tag{3.44}$$

(ii) *$d_H f(x;0)$ exists and $\underline{d}_C f(x;0) = d_H f(x;0) = \overline{d}_C f(x;0) = 0$.*

(iii) *The mappings $v \mapsto \underline{d}_C f(x;v)$, $v \mapsto \underline{d}_H f(x;v)$, $v \mapsto \overline{d}_H f(x;v)$, and $v \mapsto \overline{d}_C f(x;v)$ are positively homogeneous and for all $v, w \in \mathbb{R}^n$,*

$$|\underline{d}_C f(x;w) - \underline{d}_C f(x;v)| \le c(x)\,\|w - v\| \tag{3.45}$$

$$|\underline{d}_H f(x;w) - \underline{d}_H f(x;v)| \le c(x)\,\|w - v\| \tag{3.46}$$

$$|\overline{d}_H f(x;w) - \overline{d}_H f(x;v)| \le c(x)\,\|w - v\| \tag{3.47}$$

$$|\overline{d}_C f(x;w) - \overline{d}_C f(x;v)| \le c(x)\,\|w - v\|. \tag{3.48}$$

Moreover, $v \mapsto \underline{d}_C f(x;v)$ is sup additive,

$$\forall v, w \in \mathbb{R}^n, \quad \underline{d}_C f(x;v+w) \ge \underline{d}_C f(x;v) + \underline{d}_C f(x;w), \tag{3.49}$$

and $v \mapsto \overline{d}_C f(x;v)$ is sub additive,

$$\forall v, w \in \mathbb{R}^n, \quad \overline{d}_C f(x;v+w) \le \overline{d}_C f(x;v) + \overline{d}_C f(x;w). \tag{3.50}$$

Proof. (i) First we prove the last inequality in (3.44). Since f is Lipschitzian at x, $\overline{d}_C f(x;v)$ and $\overline{d}_H f(x;v)$ exist, and $d_H(x;0)$ exists and is equal to 0 by Theorem 3.5. So, there exist sequences $\{t_n\}$, $t_n > 0$, and $\{w_n\}$, $w_n \neq v$, such that $t_n \searrow 0$, $w_n \to v$, and

$$\overline{d}_H f(x;v) = \limsup_{\substack{t \searrow 0^+ \\ y \to x}} \frac{f(y+tv)-f(y)}{t} = \lim_{n\to\infty} \frac{f(x+t_n w_n)-f(x)}{t_n}.$$

Consider the differential quotient

$$\begin{aligned}\frac{f(x+t_n w_n)-f(x)}{t_n} &= \frac{f(x+t_n(w_n-v)+t_n v)-f(x+t_n(w_n-v))}{t_n} \\ &\quad + \frac{f(x+t_n(w_n-v))-f(x)}{t_n}.\end{aligned}$$

Since $w_n - v \to 0$ and $y_n = x + t_n(w_n - v) \to x$, we have

$$\begin{aligned}\lim_{n\to\infty} \frac{f(x+t_n(w_n-v))-f(x)}{t_n} &= d_H f(x;0) = 0 \\ \limsup_{n\to\infty} \frac{f(x+t_n(w_n-v)+t_n v)-f(x+t_n(w_n-v))}{t_n} &\leq \overline{d}_C f(x;v) \\ \Rightarrow \overline{d}_H f(x;v) \leq \overline{d}_C f(x;v).\end{aligned}$$

As for the first inequality in (3.44), apply the above result to $-f$ since it is also Lipschitzian:

$$\begin{aligned}&\overline{d}_H(-f)(x;v) \leq \overline{d}_C(-f)(x;v) \\ \Rightarrow\ &\underline{d}_C f(x;v) = -\overline{d}_C(-f)(x;v) \leq -\overline{d}_H(-f)(x;v) = \underline{d}_H f(x;v).\end{aligned}$$

(ii) From part (i).

(iii) It is sufficient to give the proof for the upper notions. Indeed, for $\alpha > 0$ and $v, w \in \mathbb{R}^n$,

$$\begin{aligned}\underline{d}_H f(x;\alpha v) &= -\overline{d}_H(-f)(x;\alpha v) = -\alpha\,\overline{d}_H(-f)(x;v) = \alpha\,\underline{d}_H f(x;v) \\ \left|\underline{d}_H f(x;w) - \underline{d}_H f(x;v)\right| &= \left|-\overline{d}_H(-f)(x;w) + \overline{d}_H(-f)(x;v)\right| \leq c(x)\,\|w-v\|.\end{aligned}$$

For $\overline{d}_H f(x;\alpha v)$, consider the differential quotient

$$q(t,w) = \frac{f(x+tw)-f(x)}{t}, \quad t > 0,\ w \in \mathbb{R}^n.$$

For $\alpha > 0$, $q(t,\alpha w) = \alpha\, q(\alpha t, w)$ and

$$\limsup_{\substack{t\searrow 0 \\ w\to v}} q(t,\alpha w) = \alpha \limsup_{\substack{t\searrow 0 \\ w\to v}} q(\alpha t, w) = \alpha \limsup_{\substack{t\searrow 0 \\ w\to v}} q(t,w).$$

Given $w_1 \to v_1$, $w_1 \neq v_1$, and $w_2 \to v_2$, $w_2 \neq v_2$,

$$q(t,w_2) = q(t,w_1) + \frac{f(x+tw_2)-f(x+tw_1)}{t} \leq q(t,w_1) + c(x)\,\|w_2 - w_1\|.$$

For $\varepsilon > 0$, $0 < t < \varepsilon$, $\|w_1 - v_1\| < \varepsilon$, and $\|w_2 - v_2\| < \varepsilon$, we have

$$q(t, w_2) \le q(t, w_1) + c(x)\,(\|v_2 - v_1\| + 2\varepsilon)$$

$$\sup_{\substack{0<t<\varepsilon \\ 0<\|w_2-v_2\|<\varepsilon}} q(t, w_2) \le \sup_{\substack{0<t<\varepsilon \\ 0<\|w_1-v_1\|<\varepsilon}} q(t, w_1) + c(x)\,(\|v_2 - v_1\| + 2\varepsilon)$$

$$\lim_{\varepsilon \searrow 0} \sup_{\substack{0<t<\varepsilon \\ 0<\|w_2-v_2\|<\varepsilon}} q(t, w_2) \le \lim_{\varepsilon \searrow 0} \sup_{\substack{0<t<\varepsilon \\ 0<\|w_1-v_1\|<\varepsilon}} q(t, w_1) + c(x)\,\|v_2 - v_1\|$$

$$\overline{d}_H f(x; v_2) \le \overline{d}_H f(x; v_1) + c(x)\,\|v_2 - v_1\|.$$

Since the roles of v_1 and v_2 can be interchanged,

$$\left|\overline{d}_H f(x; v_2) - \overline{d}_H f(x; v_1)\right| \le c(x)\,\|v_2 - v_1\|.$$

For $\overline{d}_C f(x; \alpha v)$, consider the differential quotient

$$q(t, v, y) = \frac{f(y + tv) - f(y)}{t}, \quad t > 0,\ y, v \in \mathbb{R}^n.$$

For $\alpha > 0$, $q(t, \alpha v, y) = \alpha\, q(\alpha t, v, y)$ and

$$\limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, \alpha v, y) = \alpha \limsup_{\substack{t \searrow 0 \\ y \to x}} q(\alpha t, v, y) = \alpha \limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, v, y).$$

Given $v_1, v_2 \in \mathbb{R}^n$, for $t > 0$ sufficiently small and $y \neq x$ close to x so that $y + tv_1$ and $y + tv_2$ belong to $V(x)$,

$$q(t, v_2, y) = q(t, v_1, y) + \frac{f(y + tv_2) - f(y + tv_1)}{t} \le q(t, v_1, y) + c(x)\,\|v_2 - v_1\|.$$

For $\varepsilon > 0$, $0 < t < \varepsilon$, and $\|y - x\| < \varepsilon$, we have

$$q(t, v_2, y) \le q(t, v_1, y) + c(x)\,\|v_2 - v_1\|$$

$$\sup_{\substack{0<t<\varepsilon \\ 0<\|y-x\|<\varepsilon}} q(t, v_2, y) \le \sup_{\substack{0<t<\varepsilon \\ 0<\|y-x\|<\varepsilon}} q(t, v_1, y) + c(x)\,\|v_2 - v_1\|$$

$$\lim_{\varepsilon \searrow 0} \sup_{\substack{0<t<\varepsilon \\ 0<\|y-x\|<\varepsilon}} q(t, v_2, y) \le \lim_{\varepsilon \searrow 0} \sup_{\substack{0<t<\varepsilon \\ 0<\|y-x\|<\varepsilon}} q(t, v_1, y) + c(x)\,\|v_2 - v_1\|$$

$$\overline{d}_C f(x; v_2) \le \overline{d}_C f(x; v_1) + c(x)\,\|v_2 - v_1\|.$$

Since the roles of v_1 and v_2 can be interchanged,

$$\left|\overline{d}_C f(x; v_2) - \overline{d}_C f(x; v_1)\right| \le c(x)\,\|v_2 - v_1\|.$$

Finally, given $v_1, v_2 \in \mathbb{R}^n$, for $t > 0$ sufficiently small and $y \neq x$ close to x so that $y + tv_1$, $y + tv_2$, and $y + t(v_1 + v_2)$ belong to $V(x)$,

$$q(t, v_2 + v_1, y) = q(t, v_2, y + tv_1) + q(t, v_1, y)$$

$$\limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, v_2 + v_1, y) \le \limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, v_2, y + tv_1) + \limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, v_1, y)$$

$$= \limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, v_2, y) + \limsup_{\substack{t \searrow 0 \\ y \to x}} q(t, v_1, y)$$

$$\Rightarrow \overline{d}_C f(x; v_2 + v_1) \le \overline{d}_C f(x; v_2) + \overline{d}_C f(x; v_1). \qquad \square$$

For the norm $f(x) = \|x\|$ which is convex and Lipschitzian and, more generally, for semiconvex functions,[46] the identity $d_H f(x;v) = \overline{d}_C f(x;v)$ holds. For functions enjoying that property, the nice semidifferential calculus and the chain rule are available and the mapping $v \mapsto d_H f(x;v)$ is convex and uniformly Lipschizian. But this is not necessarily true for an arbitrary Lipschitzian function. The use of upper or lower semidifferentials involving a limsup or a liminf considerably weakens the nice semidifferential calculus associated with Hadamard semidifferentiable functions. For instance, the basic functional operations take the following form:

$$\begin{aligned}
\overline{d}_C(f+g)(x;v) &\leq \overline{d}_C f(x;v) + \overline{d}_C g(x;v)\\
\overline{d}_C(f \vee g)(x;v) &\leq \overline{d}_C f(x;v) \vee \overline{d}_C g(x;v),
\end{aligned}$$

where $(f \vee g)(x) = \max\{f(x), g(x)\}$. In general, the classical chain rule will not hold for the composition of upper semidifferentiable functions[47] and only weaker geometric forms will be available.[48]

In practice, the choice of a semidifferential and its associated calculus can be important. It should be sufficiently general to effectively deal with the problem at hand but not too general to retain as many features of the classical differential calculus as possible. On this subject, it is amusing to read what Maurice Fréchet was writing about a notion of differential suggested by Paul Lévy

> ...Lastly the definition due to M. Paul Lévy, not necessarily verifying the theorem of composite functions, is still more general, but for this very reason, perhaps too general. ...[49]

3.6 Continuity, Hadamard Semidifferential, and Fréchet Differential

In dimension $n = 1$, Fréchet and Gateaux differentials coincide and correspond to the usual derivative of Definition 2.1(iii) and such functions are continuous by Theorem 2.1.

Example 3.8 shows that, in dimension $n \geq 2$, a Gateaux differentiable function is not necessarily continuous. So, it is not the linearity of the semidifferential that causes the continuity of the function. Example 3.7 considers a continuous function that is *Hadamard semidifferentiable*, but not Gateaux differentiable (that is, $d_H(x;v)$ is not linear in v). It is to preserve the continuity of the function that the stronger notion of Fréchet differential has been introduced.

From Theorem 3.1, it is necessary that the function be Hadamard semidifferentiable to make it Fréchet differentiable and thence continuous. It turns out that only the existence of $d_H f(x;0)$ is sufficient for the continuity of f at x.

---

[46] For semiconcave functions, we have $\underline{d}_C f(x;v) = d_H f(x;v)$ (see section 4.4).

[47] Consider the real functions $f(x) = |x|$ and $g(x) = -x$, and their composition $(g \circ f)(x) = g(f(x)) = -|x|$. It is readily seen that $\overline{d}_C(g \circ f)(0;1) = +1$, $\overline{d}_C g(0;1) = -1$, $\overline{d}_C f(0;1) = 1$, and that $\overline{d}_C(g \circ f)(0;1) = +1 > -1 = \overline{d}_C g(0;\overline{d}_C f(0;1))$.

[48] Cf. F. H. Clarke [**2**, pp. 42–47].

[49] "...Enfin la définition due à M. Paul Lévy, ne vérifiant pas nécessairement le théorème des fonctions composées, est plus générale encore, mais, pour cette même raison, peut-être trop générale. ..." (M. Fréchet [**3**, p. 233] in 1937).

Theorem 3.7. *Let $f:\mathbb{R}^n\to\mathbb{R}$. If $d_H f(x;0)$ exists at x, f is continuous at x. Moreover, for any $\alpha\in(0,1)$ and any $\varepsilon>0$,*

$$\exists\delta>0 \text{ such that } \forall y\in B_\delta(x),\quad \frac{|f(y)-f(x)|}{\|y-x\|^\alpha}<\varepsilon.$$

Corollary 1. *If $f:\mathbb{R}^n\to\mathbb{R}$ is Fréchet differentiable at x, then f is continuous at x.*

Proof. If $d_H f(x;0)$ exists, then $d_H f(x;0)=df(x;0)=0$. For all y such that $y\neq x$,

$$\frac{|f(y)-f(x)|}{\|y-x\|^\alpha}=\left|\frac{f\left(x+\|y-x\|^\alpha\frac{y-x}{\|y-x\|^\alpha}\right)-f(x)}{\|y-x\|^\alpha}-0\right|.$$

Since $d_H f(x;0)=0$, as $y\to x$, $t=\|y-x\|^\alpha\to 0$ and

$$w=\frac{y-x}{\|y-x\|^\alpha}=\|y-x\|^{1-\alpha}\frac{y-x}{\|y-x\|}\to 0 \text{ as } y\to x$$

$$\Rightarrow \lim_{y\to x}\left|\frac{f\left(x+\|y-x\|^\alpha\frac{y-x}{\|y-x\|^\alpha}\right)-f(x)}{\|y-x\|^\alpha}-d_H f(x;0)\right|=0.$$

Therefore, for all $\varepsilon>0$, there exists $\delta>0$ such that

$$\forall y\in B_\delta(x),\quad \frac{|f(y)-f(x)|}{\|y-x\|^\alpha}=\left|\frac{f\left(x+\|y-x\|^\alpha\frac{y-x}{\|y-x\|^\alpha}\right)-f(x)}{\|y-x\|^\alpha}-0\right|<\varepsilon.$$

In particular, this yields the continuity of f at x. □

It is interesting to compare Theorem 3.7 to Theorem 3.5, which says that if f is Lipschitzian[50] at x, then $d_H f(x;0)$ exists.

3.7 Mean Value Theorem for Functions of Several Variables

Theorem 3.8. *Let $f:\mathbb{R}^n\to\mathbb{R}$, $x\in\mathbb{R}^n$, and $v\in\mathbb{R}^n$. If the function $t\mapsto g(t)\overset{\text{def}}{=}f(x+tv)$ is continuous on $[0,1]$ and differentiable on $]0,1[$, then*

$$\exists\alpha\in\,]0,1[\text{ such that } f(x+v)=f(x)+df(x+\alpha v;v). \tag{3.51}$$

Proof. It is sufficient to observe that for $t\in]0,1[$,

$$g'(t)=df(x+tv;v).$$

Indeed, by definition of $g'(t)$ at a point $0<t<1$: for $|s|$ sufficiently small,

$$\begin{aligned}\frac{g(t+s)-g(t)}{s}&=\frac{f(x+(t+s)v)-f(x+tv)}{s}\\&=\frac{f((x+tv)+s\,v)-f(x+tv)}{s}\to df(x+tv;v) \text{ as } s\to 0.\end{aligned}$$

To complete the proof, it is sufficient to apply the mean value theorem (Theorem 2.4) to $g(t)=f(x+tv)$: there exists $\alpha\in]0,1[$ such that $g(1)=g(0)+g'(\alpha)$. □

[50] That corresponds to $\alpha=1$.

The following corollaries (Corollaries 1 and 2) and Remark 3.7 are from W. RUDIN [**1**].

Corollary 1. *If f is Gateaux differentiable on a convex open subset $U \subset \mathbb{R}^n$ and if*

$$\exists c \geq 0 \text{ such that } \forall x \in U, \quad \|\nabla f(x)\| \leq c, \tag{3.52}$$

then f is Lipschitzian on U,

$$|f(y) - f(x)| \leq c\|y - x\|, \quad \forall x, y \in U, \tag{3.53}$$

and f is Fréchet differentiable on U.

Proof. By the mean value theorem (Theorem 2.4) with $v = y - x$, there exists $\alpha = \alpha(y) \in]0,1[$ such that

$$f(y) - f(x) = df(x + \alpha(y - x); y - x) = \nabla f(x + \alpha(y - x)) \cdot (y - x).$$

By Cauchy inequality,

$$|f(y) - f(x)| \leq \|\nabla f(x + \alpha(y - x))\| \, \|y - x\| \leq c\|y - x\|.$$

Since f is Gateaux differentiable and Lipschitzian, then by Theorem 3.5 it is Hadamard semidifferentiable and it is Fréchet differentiable by Theorem 3.1. □

Corollary 2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Gateaux differentiable on a connected open nonempty subset U of $\mathbb{R}^n$. If*

$$\forall x \in U, \quad \nabla f(x) = 0, \tag{3.54}$$

then $f(x)$ is constant on U.

Proof. Since $U \neq \varnothing$, pick a point $x_0 \in U$ and define $U_1 = \{x \in U : f(x) = f(x_0)\}$. By definition, $U_1 \neq \varnothing$. Let $x \in U_1$. Since U is open, there exists $r > 0$ such that $B_r(x) \subset U$. Apply Corollary 1 with $c = 0$ and the open convex set $B_r(x)$ to get

$$\forall y \in B_r(x), \quad f(y) = f(x) = f(x_0).$$

Hence $B_r(x) \subset U_1$. This shows that U_1 is an open subset of $\mathbb{R}^n$. Consider the complement with respect to U

$$U \backslash U_1 \stackrel{\text{def}}{=} \{x \in U : f(x) \neq f(x_0)\}.$$

For each point $x_1 \in U \backslash U_1$, there exists $r > 0$ such that $B_r(x_1) \subset U$ and by the same argument $f(x) = f(x_1) \neq f(x_0)$ for all $x \in B_r(x_1)$. Thence, $B_r(x_1) \subset U \backslash U_1$ and $U \backslash U_1$ is open.

If $U \backslash U_1 \neq \varnothing$, then U is the union of two disjoint open sets U_1 and $U \backslash U_1$. But this is impossible since U is connected. Therefore, $U \backslash U_1 = \varnothing$ and $U = U_1$. □

Remark 3.7.
H. WHITNEY [**1**] has given an example of a convex set $U \subset \mathbb{R}^2$and a differentiable function f such that $\nabla f(x, y) = 0$ for all $(x, y) \in U$, but $f(x, y)$ is not constant on U. This set has no interior point. □

Corollary 1 seems to contradict Example 3.8 of a Gateaux differentiable function $f(x,y)$ at the point $(0,0)$, but discontinuous at that point. Indeed, the function $f(x,y)$ is Gateaux differentiable not only at $(0,0)$, but at every point of $\mathbb{R}^2$. However, the function and its gradient are not bounded in any open ball $B_\delta(0,0)$, $\delta > 0$, around $(0,0)$. It is easy to check that for $(x,y) \neq (0,0)$,

$$\frac{\partial f}{\partial y} = -\frac{2x^6(y-x^2)}{\left[(y-x^2)^2+x^8\right]^2}, \quad \frac{\partial f}{\partial x} = 2x^5\,\frac{(y-x^2)(3y-x^2)-x^8}{\left[(y-x^2)^2+x^8\right]^2}.$$

Choosing $y = x^2$, $x \neq 0$,

$$\frac{\partial f}{\partial y}(x,x^2) = 0, \quad \frac{\partial f}{\partial x}(x,x^2) = -\frac{2}{x^3}$$

and $\nabla f(x,x^2)$ is not bounded as x goes to 0.

3.8 Functions of Classes $C^{(0)}$ and $C^{(1)}$

When f is Gateaux (resp., Fréchet) differentiable in x, we have seen that the gradient can be expressed in terms of the partial derivatives of f at x. In general, the converse is not true since the existence of directional derivatives in all directions is not sufficient to get the Gateaux (and Fréchet) differentiability. However, when appropriate continuity conditions are imposed on the partial derivatives, the function becomes Fréchet differentiable and the gradient is completely specified by the partial derivatives. The main structure of the proof is borrowed from W. Rudin [**1**].

Definition 3.12.
Let $f : \mathbb{R}^n \to \mathbb{R}$, $U \subset \mathbb{R}^n$ be open, and $\{e_i\}_{i=1}^n$ be the canonical orthonormal basis of $\mathbb{R}^n$.

(i) f is of *class* $C^{(0)}$ on U if f is continuous on U.

(ii) f is of *class* $C^{(1)}$ on U if the partial derivatives $\partial_i f(x)$, $1 \le i \le n$, exist and are continuous on U.

The above definitions extend to vector-valued functions. □

Theorem 3.9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $\{e_i\}_{i=1}^n$ be the canonical orthonormal basis of $\mathbb{R}^n$.*

(i) *If f has partial derivatives $\partial_i f$, $i = 1,\dots,n$, on a neighborhood of x that are continuous on x, then f is Fréchet differentiable at x (and hence continuous at x) and*

$$\forall v \in \mathbb{R}^n, \quad L(x)v = \sum_{i=1}^n df(x;e_i)\,e_i \cdot v, \quad \nabla f(x) = \sum_{i=1}^n df(x;e_i)\,e_i.$$

(ii) *If f has partial derivatives $\partial_i f$, $i = 1,\dots,n$, on an open subset U of $\mathbb{R}^n$ that are continuous on U, then f is Fréchet differentiable on U (and hence continuous on U) and for all $y \in U$,*

$$\forall v \in \mathbb{R}^n, \quad L(y)v = \sum_{i=1}^n \partial_i f(y)\,e_i \cdot v, \quad \nabla f(y) = \sum_{i=1}^n \partial_i f(y)\,e_i,$$

and the maps

$$y \mapsto \nabla f(y) : U \to \mathbb{R}^n \quad \textit{and} \quad (y,w) \mapsto df(y;w) : U \times \mathbb{R}^n \to \mathbb{R} \tag{3.55}$$

are continuous.

Corollary 1. *If f is of class $C^{(1)}$ on an open set U, then f is Fréchet differentiable on U and hence of class $C^{(0)}$ on U. Moreover, ∇f is of class $C^{(0)}$ on U.*

Proof of Theorem 3.9. (i) It is sufficient to prove that f is Fréchet differentiable at x. The continuity follows from Theorem 3.7 and the other properties from the comments following the definitions of Gateaux and Fréchet differentials. To show that f is Fréchet differentiable, we show that f is Hadamard semidifferentiable and that $d_H f(x;v)$ is linear with respect to v, and we apply Theorem 3.1.

Any element $v = (v_1, \dots, v_n)$ of $\mathbb{R}^n$ can be written as

$$v = \sum_{i=1}^{n} v_i e_i.$$

For each y in the neighborhood $V(x)$ of x, define the linear map

$$v \mapsto L(y)v \stackrel{\text{def}}{=} \sum_{i=1}^{n} \partial_i f(y)\, v_i : \mathbb{R}^n \to \mathbb{R}.$$

Fix $v \in \mathbb{R}^n$ and consider sequences $w_k \to v$ and $t_k \searrow 0$. I There exists N such that

$$\forall k > N, \quad x + t_k w_k \in V(x).$$

Define the following points:

$$x_k^0 \stackrel{\text{def}}{=} x, \quad x_k^i \stackrel{\text{def}}{=} x_k^{i-1} + t_k (w_k)_i e_i, \quad 1 \le i \le n.$$

We want to show that the differential quotient minus $L(x)v$,

$$q_k \stackrel{\text{def}}{=} \frac{f(x + t_k w_k) - f(x)}{t_k} - L(x)v,$$

goes to 0 as k goes to infinity. That difference can be rewritten in the form

$$q_k = \sum_{i=1}^{n} \frac{f(x_k^i) - f(x_k^{i-1})}{t_k} - \partial_i f(x)\, v_i.$$

Since f is differentiable in the direction e_i at every point of $V(x)$, the function $g_i(\alpha) = f(x^{i-1} + \alpha\, t_k\, (w_k)_i e_i)$ is continuous in $[0,1]$ and differentiable in $]0,1[$. By the mean value theorem (Theorem 3.8),

$$\begin{aligned}\exists \alpha_k^i \in\,]0,1[, \quad f(x_k^i) - f(x_k^{i-1}) &= df(x_k^{i-1} + \alpha_k^i t_k (w_k)_i e_i; t_k (w_k)_i e_i) \\ &= t_k (w_k)_i\, \partial_i f(x_k^{i-1} + \alpha_k^i t_k (w_k)_i e_i)\end{aligned}$$

by homogeneity, and finally

$$\frac{f(x_k^i)-f(x_k^{i-1})}{t_k} = (w_k-v)_i\, \partial_i f(x_k^{i-1}+\alpha_k^i t_k (w_k)_i e_i) + v_i\, \partial_i f(x_k^{i-1}+\alpha_k^i t_k (w_k)_i e_i)$$

$$\begin{aligned}\frac{f(x_k^i)-f(x_k^{i-1})}{t_k} - \partial_i f(x)\, v_i &= (w_k-v)_i\, \partial_i f(x_k^{i-1}+\alpha_k^i t_k (w_k)_i e_i)\\ &\quad + v_i \left[\partial_i f(x_k^{i-1}+\alpha_k^i t_k (w_k)_i e_i) - \partial_i f(x)\right].\end{aligned}$$

But, by construction, for all i,

$$\begin{aligned}|x_k^{i-1}+\alpha_k^i t_k (w_k)_i e_i - x| &= \left\| \sum_{j=1}^{i-1} (x_k^j - x_k^{j-1}) + \alpha_k^i t_k (w_k)_i e_i \right\|\\ &= \left\| \sum_{j=1}^{i-1} t_k (w_k)_j e_j + \alpha_k^i t_k (w_k)_i e_i \right\| = t_k \left\{ \sum_{j=1}^{i-1} |(w_k)_j|^2 + |\alpha_k^i (w_k)_i|^2 \right\}^{1/2}\\ &\le t_k \left\{ \sum_{j=1}^{i} |(w_k)_j|^2 \right\}^{1/2} \le t_k \left\{ \sum_{j=1}^{n} |(w_k)_j|^2 \right\}^{1/2} = t_k\, \|w_k\|\end{aligned}$$

which goes to zero as $w_k \to v$ and $t_k \searrow 0$ and since $\partial_i f(y)$ is continuous at x, $\partial_i f(x_k^{i-1} + \alpha_k^i t_k (w_k)_i e_i) \to \partial_i f(x)$, and q_k goes to zero as k goes to infinity. We have shown that, for all $v \in \mathbb{R}^n$, $d_H f(x;v) = L(x)v$ that is linear in v by definition of $L(x)$. This proves that f is Fréchet differentiable at x.

(ii) Results for U follow from the results in part (i) at x. It remains to prove the continuity of the maps (3.55). As the expression of the gradient $\nabla f(y)$ is

$$\nabla f(y) = \sum_{i=1}^{n} df(y;e_i)\, e_i,$$

it is continuous on U as the sum of n functions continuous on U by assumption on the partial derivatives $\partial_i f(y) = df(y;e_i)$. Since f is Fréchet differentiable on U, for all y in U,

$$\forall w \in \mathbb{R}^n, \quad df(y;w) = \nabla f(y) \cdot w = L(y)w.$$

Given an arbitrary pair (x,v) in $U \times \mathbb{R}^n$ and another pair (y,w) in $U \times \mathbb{R}^n$, estimate the difference

$$\begin{aligned} df(y;w) - df(x;v) &= df(y;w) - df(x;w) + df(x;w) - df(x;v)\\ &= [\nabla f(y) - \nabla f(x)] \cdot w + \nabla f(x) \cdot (w-v)\\ \Rightarrow |df(y;w) - df(x;v)| &\le \|\nabla f(y) - \nabla f(x)\|\, \|w\| + \|\nabla f(x)\|\, \|w-v\|.\end{aligned}$$

Since x is fixed, there exists a constant $c > 0$ such that $|\nabla f(x)| \le c$ and for all $\varepsilon > 0$ and all w such that $\|w-v\| \le \varepsilon/(2c)$, we have

$$\|\nabla f(x)\|\, \|w-v\| \le \varepsilon/2 \text{ and } \|w\| \le \|v\| + \varepsilon/(2c).$$

Now, by continuity of the n partial derivatives on U, there exists $\delta(x) > 0$ such that

$$\begin{aligned}\|y-x\| < \delta(x) \quad &\Rightarrow \left\{\sum_{j=1}^{n} |\partial_j f(y) - \partial_j f(x)|^2\right\}^{1/2} \le \frac{\varepsilon}{2(\|v\| + \varepsilon/(2c))}\\ &\Rightarrow \|\nabla f(y) - \nabla f(x)\| \, \|w\| \le \varepsilon/2.\end{aligned}$$

Finally, for all $\varepsilon > 0$, there exists $\delta' = \min\{\delta(x), \varepsilon/(2c)\}$ such that

$$\|y-x\| < \delta' \text{ and } \|w-v\| \le \delta' \quad \Rightarrow |df(y;w) - df(x;v)| \le \varepsilon$$

and hence the continuity as (y,w) goes to (x,v). □

If f and g are functions of class $C^{(1)}$ on the same open set U, then the sum $(f+g)(x) = f(x)+g(x)$ is of class $C^{(1)}$ on U. In a similar fashion, the product $(fg)(x) = f(x)g(x)$ is of class $C^{(1)}$ using the property

$$\partial_i(fg) = (\partial_i f)g + f(\partial_i g), \quad 1 \le i \le n.$$

The composition of two functions of class $C^{(1)}$ is also of class $C^{(1)}$. Indeed, if $f = \psi \circ g$, $g : U \to \mathbb{R}$ ($U \subset \mathbb{R}^n$) and $\psi : \mathbb{R} \to \mathbb{R}$, then

$$\partial_i f = (\psi' \circ g)\partial_i g, \quad \text{où } \psi' = \frac{d\psi}{dx}.$$

3.9 Functions of Class $C^{(p)}$ and Hessian Matrix

Second- and higher-order *semidifferentials* are defined in the same manner as the ones of first order.

Definition 3.13.
Let $f : \mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$, and v and w be two directions in $\mathbb{R}^n$. Assume that $df(y;v)$ exists for all y on a neighborhood $V(x)$ of x. The function f has a *second-order semidifferential* in the directions (v,w) in x if the limit

$$\lim_{t \searrow 0} \frac{df(x+tw;v) - df(x;v)}{t}$$

exists. The limit is denoted $d^2 f(x;v;w)$. □

In general, the order of the directions (v,w) is important.

Theorem 3.10. *If $f : \mathbb{R}^n \to \mathbb{R}$ is Gateaux differentiable on a neighborhood $V(x)$ of x, and $\nabla f(y)$ is Gateaux differentiable at x, then the map*

$$(v,w) \mapsto d^2 f(x;v;w) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{3.56}$$

is bilinear; that is,

$$\forall v \in \mathbb{R}^n, \quad w \mapsto d^2 f(x;v;w) : \mathbb{R}^n \to \mathbb{R} \textit{ is linear,} \tag{3.57}$$

$$\forall w \in \mathbb{R}^n, \quad v \mapsto d^2 f(x;v;w) : \mathbb{R}^n \to \mathbb{R} \textit{ is linear.} \tag{3.58}$$

Moreover, there exists a unique linear map

$$Hf(x) : \mathbb{R}^n \to \mathbb{R}^n \textit{ such that } \forall v,w \in \mathbb{R}^n, \quad d^2 f(x;v;w) = Hf(x)w \cdot v. \tag{3.59}$$

Any pair of directions $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$ can be written as

$$v = \sum_{i=1}^{n} v_i e_i, \quad w = \sum_{i=1}^{n} w_i e_i,$$

in terms of the canonical orthonormal basis $\{e_i\}_{i=1}^n$. Under the assumptions of Theorem 3.10, the map $(v, w) \mapsto d^2 f(x; v; w) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is bilinear. Therefore,

$$d^2 f(x; v; w) = d^2 f\left(x; \sum_{i=1}^{n} v_i e_i; \sum_{j=1}^{n} w_j e_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} d^2 f(x; e_i; e_j) v_i w_j.$$

The elements $d^2 f(x; e_i; e_j)$ are the entries of the matrix associated with the linear map $Hf(x) : \mathbb{R}^n \to \mathbb{R}^n$. In the light of the theorem and the above comments, we now introduce the following definitions.

Definition 3.14.
Let the assumptions of Theorem 3.10 hold.

(i) The *Hessian*[51] of f at x is the linear map $Hf(x) : \mathbb{R}^n \to \mathbb{R}^n$ defined by (3.59).

(ii) The *Hessian matrix* of f at x is the $n \times n$ matrix with entries

$$Hf(x)_{ij} \stackrel{\text{def}}{=} d^2 f(x; e_i; e_j)$$

with respect to the canonical orthonormal basis $\{e_i\}_{i=1}^n$ of $\mathbb{R}^n$.

The same notation $Hf(x)$ will be used for the map and its associated matrix. □

Proof of Theorem 3.10. By assumption, for all $y \in V(x)$, f is Gateaux differentiable at y and

$$df(y; v) = \nabla f(y) \cdot v$$

and, since ∇f is Gateaux differentiable at x,

$$d\nabla f(x; v) = D(\nabla f)(x)\, v,$$

where $D(\nabla f)(x) : \mathbb{R}^n \to \mathbb{R}^n$ is the Jacobian (linear) map associated with the vector function $y \mapsto \nabla f(y)$.

For v and w in $\mathbb{R}^n$ and $t > 0$, consider the quotient

$$\frac{df(x + tw; v) - df(x; v)}{t} = \frac{\nabla f(x + tw) - \nabla f(x)}{t} \cdot v.$$

As $t \to 0$,

$$\frac{\nabla f(x + tw) - \nabla f(x)}{t} \to D(\nabla f)(x)\, w$$
$$\frac{df(x + tw; v) - df(x; v)}{t} = \frac{\nabla f(x + tw) - \nabla f(x)}{t} \cdot v \to D(\nabla f)(x)\, w \cdot v$$
$$\Rightarrow d^2 f(x; v; w) = D(\nabla f)(x)\, w \cdot v.$$

[51] The Hessian matrix was developed in the 19th century by Ludwig Otto Hesse (1811–1874) and later named after him. Hesse himself had used the term "functional determinants."

So, we get the bilinearity and the linear map $Hf(x)$ coincides with the linear map $D(\nabla f)(x)$. $\square$

We stress the fact that the definition of the Hessian matrix is compatible with the definition of the *Jacobian matrix* DF of a vector function $F : \mathbb{R}^n \to \mathbb{R}^n$. Indeed,

$$DF(x)_{ij} \stackrel{\text{def}}{=} \frac{\partial F_i}{\partial x_j}(x),$$

where F_i is the ith component of a function F. Taking $F = \nabla f$, we get

$$D(\nabla f(x))_{ij} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}(x)\right) = d^2 f(x; e_i; e_j) = Hf(x)_{ij}.$$

We shall use the notation $\partial^2_{ji} f(x)$.

The higher-order partial derivatives are defined in the same way as second-order derivatives.

$$\frac{\partial^m f(x)}{\partial x_{i_m} \ldots \partial x_{i_1}} = d^m f(x; e_{i_1}; \ldots; e_{i_m}). \tag{3.60}$$

We shall use the notation $\partial^m_{i_m \cdots i_1} f(x)$.

Definition 3.15.
Let $f : \mathbb{R}^n \to \mathbb{R}$ and U an open subset of $\mathbb{R}^n$. The function f is of class $C^{(p)}$ on U if all partial derivatives of f of order p exist and are continuous on U. $\square$

From the previous results (see Corollary 1 to Theorem 3.9), it is readily seen that functions of class $C^{(p)}$ on U are of class $C^{(p-1)}$ on U and so on.

The second-order differentials will be important to characterize the convexity of a differentiable function and hence to characterize its minimizers. In general, the Hessian matrix is not symmetrical, but it is for functions of class $C^{(2)}$.

Theorem 3.11. *Let $f : \mathbb{R}^n \to \mathbb{R}$. Let $V(x)$ be a neighborhood of x such that*

$$\forall y \in V(x), \forall v \in \mathbb{R}^n, \quad df(y; v) \text{ exists} \tag{3.61}$$

$$\forall y \in V(x), \forall v, w \in \mathbb{R}^n, \quad d^2 f(y; v; w) \text{ exists.} \tag{3.62}$$

If for all v and w in $\mathbb{R}^n$ the map

$$y \mapsto d^2 f(y; v; w) : V(x) \to \mathbb{R} \text{ is continuous at } x, \tag{3.63}$$

then

$$\forall v, w \in \mathbb{R}^n, \quad d^2 f(x; v; w) = d^2 f(x; w; v). \tag{3.64}$$

Corollary 1. *Let $\{e_i\}_{i=1}^n$ be the canonical orthonormal basis of $\mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, and U be an open subset of $\mathbb{R}^n$. Assume that f has partial derivatives $\partial_i f$ on U and that each $\partial_i f$ also has partial derivatives $\partial_j(\partial_i f)$ on U. If for each pair i and j, the map*

$$y \mapsto \partial_j(\partial_i f(y)) : U \to \mathbb{R} \text{ is continuous,} \tag{3.65}$$

then f is of class $C^{(2)}$ on U,

$$\forall y \in U, \forall v, w \in \mathbb{R}^n,\ d^2 f(y;v;w) = d^2 f(y;w;v), \quad \partial_j(\partial_i f(y)) = \partial_i(\partial_j f(y)), \tag{3.66}$$

and the Hessian matrix $Hf(y)$ is symmetrical.

Proof of Theorem 3.11. Define

$$C_{s,t} \stackrel{\text{def}}{=} f(x+sv+tw) - f(x+sv) - f(x+tw) + f(x).$$

We give two different expressions of $C_{s,t}$. For $s > 0$ and $t > 0$ sufficiently small, x, $x+sv$, $x+sv+tw$ are in U. We have

$$C_{s,t} = g(x+tw) - g(x), \quad \text{where} \quad g(z) \stackrel{\text{def}}{=} f(z+sv) - f(z).$$

By Taylor's theorem (Theorem 2.6), using a first-order expansion, there exists $\alpha_1 \in]0,1[$ such that

$$C_{s,t} = dg(x + \alpha_1 tw; tw).$$

By definition of g, this identity can be rewritten as a function of df:

$$C_{s,t} = df(x + \alpha_1 tw + sv; tw) - df(x + \alpha_1 tw; tw).$$

A second application of Taylor's theorem (Theorem 2.6) yields an $\alpha_2 \in]0,1[$:

$$C_{s,t} = d^2 f(x + \alpha_1 tw + \alpha_2 sv; tw; sv) = st\, d^2 f(x + \alpha_1 tw + \alpha_2 sv; w; v)$$

using the positive homogeneity. In an analogous fashion, by interchanging the roles of s and t, we get

$$\exists \alpha_3, \alpha_4 \in]0,1[, \quad C_{s,t} = st\, d^2 f(x + \alpha_3 tw + \alpha_4 sv; v; w).$$

Hence,

$$d^2 f(x + \alpha_1 tw + \alpha_2 sv; w; v) = d^2 f(x + \alpha_3 tw + \alpha_4 sv; v; w).$$

As $s \to 0$ and $t \to 0$, we get, by continuity of $d^2 f$, $d^2 f(x;w;v) = d^2 f(x;v;w)$. □

4 Convex and Semiconvex Functions

Recall Definitions 7.1 and 7.4 of Chapter 2 of a convex subset U of $\mathbb{R}^n$ and of convex and strictly convex functions $f : \mathbb{R}^n \to \mathbb{R}$ on a convex U.

4.1 Directionally Differentiable Convex Functions

Theorem 4.1. *Let $U \subset \mathbb{R}^n$ be a convex open subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be directionally differentiable*[52] *in U.*

[52] In the sense of Definition 3.1(iii), that is, for all $v \in \mathbb{R}^n$, the limit of the differential quotient

$$\lim_{t \to 0} \frac{f(x+tv) - f(v)}{t}$$

exists as t goes to 0. This condition is equivalent to saying that for all $v \in \mathbb{R}^n$, $df(x;v)$ exists and $df(x;-v) = -df(x;v)$.

(i) *f is convex on U if and only if*

$$\forall x, y \in U, \quad f(y) \geq f(x) + df(x; y-x). \tag{4.1}$$

(ii) *If, in addition, f is Gateaux differentiable in U, then f is convex on U if and only if*

$$\forall x, y \in U, \quad f(y) \geq f(x) + \nabla f(x) \cdot (y-x). \tag{4.2}$$

Proof. (i) ($\Rightarrow$) For all $\lambda \in]0,1[$,

$$\begin{aligned} &f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda) f(x) \\ \Rightarrow\ &f(x + \lambda(y-x)) - f(x) \leq \lambda(f(y) - f(x)). \end{aligned}$$

By dividing by $\lambda > 0$ and going to the limit as $\lambda \to 0$, we get

$$df(x; y-x) \leq f(y) - f(x).$$

($\Leftarrow$) Conversely, apply condition (4.1) twice for $\lambda \in [0,1]$ and $x, y \in U$:

$$\begin{aligned} f(x) &\geq f(x + \lambda(y-x)) + df(x + \lambda(y-x); -\lambda(y-x)) \\ f(y) &\geq f(x + \lambda(y-x)) + df(x + \lambda(y-x); (1-\lambda)(y-x)). \end{aligned}$$

Multiply the first inequality by $1-\lambda$ and the second by λ. Add them up. As U is convex, $x + \lambda(y-x) \in U$, and, by homogeneity, $df(x + \lambda(y-x); -(y-x)) = -df(x + \lambda(y-x); y-x)$. This yields

$$(1-\lambda) f(x) + \lambda f(y) \geq f(x + \lambda(y-x))$$

and the convexity of f on U.

(ii) As f is Gateaux differentiable, its gradient exists and $df(x; v) = \nabla f(x) \cdot v$. Substitute it in the expression of part (i). □

Theorem 4.2. *Let $U \subset \mathbb{R}^n$ be a convex open subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be directionally differentiable on U.*

(i) *f is strictly convex on U if and only if*

$$\forall x, y \in U, \quad x \neq y, \quad f(y) > f(x) + df(x; y-x). \tag{4.3}$$

(ii) *If, in addition, f is Gateaux differentiable in U, then f is strictly convex on U if and only if*

$$\forall x, y \in U, \quad x \neq y, \quad f(y) > f(x) + \nabla f(x) \cdot (y-x). \tag{4.4}$$

Proof. (i) If f is strictly convex on U, we get (4.1) from Theorem 4.1. Therefore, for all x and y in U such that $x \neq y$ and $t \in]0,1[$,

$$df(x; t(y-x)) \leq f(x + t(y-x)) - f(x).$$

By positive homogeneity, $t df(x;y-x) = df(x;t(y-x))$. As f is strictly convex,

$$\begin{aligned} f(x+t(y-x)) - f(x) =& f((1-t)x+ty) - f(x) \\ &< (1-t)f(x) + tf(y) - f(x) = t\,[f(y) - f(x)] \\ \Rightarrow&\ t df(x;(y-x)) < t\,[f(y) - f(x)]. \end{aligned}$$

By dividing both sides by t, we get (4.3). The proof of the converse is the same as the one of Theorem 4.1 but with $\lambda \in]0,1[$ and $x \neq y$.

(ii) Since f is Gateaux differentiable, its gradient exists and $df(x;v) = \nabla f(x) \cdot v$. Substitute in part (i). □

For the next theorem, recall Definition 5.3 of Chapter 2.

Definition 4.1.
A symmetrical matrix A is *positive definite* (resp., *positive semidefinite*) if

$$\forall x \in \mathbb{R}^n, x \neq 0, \quad (Ax)\cdot x > 0 \quad (\text{resp., } \forall x \in \mathbb{R}^n, \quad (Ax)\cdot x \geq 0).$$

The property will be denoted $A > 0$ (resp., $A \geq 0$). □

Theorem 4.3. *Let $U \subset \mathbb{R}^n$ be open convex and $f : \mathbb{R}^n \to \mathbb{R}$ be of class $C^{(2)}$ on U.*

(i) *f is convex on U if and only if $Hf(y) \geq 0$ for all $y \in U$.*

(ii) *If there exists $x \in U$ such that $Hf(x) > 0$, then there exists a convex neighborhood $V(x)$ of x such that f is strictly convex on $V(x)$.*

Remark 4.1.
The converse of part (ii) of Theorem 4.3 is not true. Indeed, consider the function $f(x) = x^4$ defined on $\mathbb{R}$. Its second-order derivative is given by $f^{(2)}(x) = 12x^2$. At $x = 0$, f is zero even if f is strictly convex on any neighborhood of $x = 0$. □

Proof of Theorem 4.3. We again make use of Taylor's theorem (Theorem 2.6) applied to the function $g(t) = f(x+t(y-x))$. For all x, y on U, there exists $\alpha \in]0,1[$ such that

$$f(y) = f(x) + \nabla f(x)\cdot(y-x) + \frac{1}{2} Hf(x+\alpha(y-x))(y-x)\cdot(y-x).$$

(i) If f is convex on U, by Theorem 4.1,

$$0 \leq f(y) - f(x) - \nabla f(x)\cdot(y-x) = \frac{1}{2} Hf(x+\alpha(y-x))(y-x)\cdot(y-x).$$

But, since U is open, there exists $r > 0$ such that $B_r(x) \subset U$. Therefore,

$$Hf(x+\alpha r b)b\cdot b \geq 0, \quad \forall b \in B_1(0).$$

Since f is of class $C^{(2)}$ and $|\alpha r b| < r$, as r goes to 0,

$$Hf(x)b\cdot b \geq 0, \quad \forall b \in B_1(0) \quad \Rightarrow \forall v \in \mathbb{R}^n, Hf(x)v\cdot v \geq 0,$$

and $Hf(x) \geq 0$ for all $x \in U$. Conversely, for all $x, y \in U$, there exists $\alpha \in]0,1[$ such that

$$f(y) - f(x) - \nabla f(x) \cdot (y - x) = \frac{1}{2} Hf(x + \alpha(y - x))(y - x) \cdot (y - x) \geq 0$$

since $x + \alpha(y - x) \in U$ and, by assumption, $Hf(x + \alpha(y - x)) \geq 0$. The function f is then convex by Theorem 4.1.

(ii) If $Hf(x) > 0$ for some $x \in U$, then by continuity, there exists $r > 0$ such that

$$\overline{B_r(x)} \subset U \text{ and } \forall y \in B_r(x), \quad Hf(y) > 0.$$

Now make use of Taylor's theorem (Theorem 2.6) in $\overline{B_r(x)}$. Then, for all $y, z \in B_r(x)$, $y \neq z$, there exists $\alpha \in]0,1[$ such that

$$f(y) - f(z) - \nabla f(z) \cdot (y - z) = \frac{1}{2} Hf(z + \alpha(y - z))(y - z) \cdot (y - z) > 0$$

since $z + \alpha(y - z) \in B_r(x)$ and, by assumption, $Hf(z + \alpha(y - z)) > 0$. Therefore, for all $y, z \in B_r(x)$, $y \neq z$,

$$f(y) - f(z) - \nabla f(z) \cdot (y - z) > 0$$

and, by Theorem 4.2, f is strictly convex on $B_r(x)$. □

4.2 ► Semidifferentiability and Continuity of Convex Functions

One of the most studied family of functions is the one of convex functions. In general, a convex function on a compact convex set U is not continuous and does not necessarily have a semidifferential in all points of U as illustrated by the graphs of Figure 3.9 corresponding to the two functions of Example 4.1.

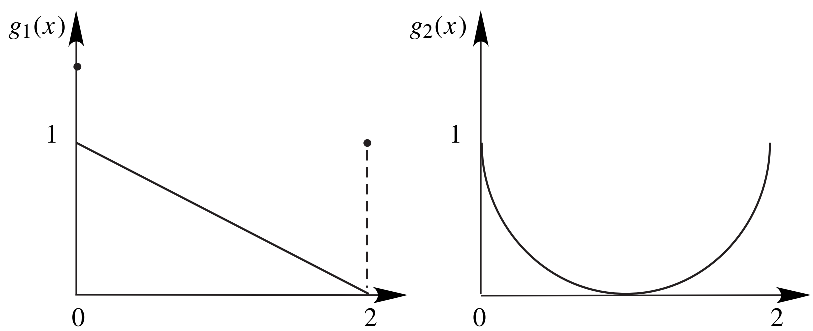


Figure 3.9. *The two convex functions* g_1 *and* g_2 *on* $[0,1]$ *of Example* 4.1.

Example 4.1.
Consider the following convex discontinuous function $g_1 : [0,2] \to \mathbb{R}$:

$$g_1(x) \stackrel{\text{def}}{=} \begin{cases} 3/2, & \text{if } x = 0 \\ 1 - x/2, & \text{if } 0 < x < 2 \\ 1, & \text{if } x = 2. \end{cases}$$

It is readily checked that $dg_1(0;1) = -\infty$ and $dg_1(2;-1) = -\infty$.

Now consider the continuous convex function $g_2 : [0,2] \to \mathbb{R}$:

$$g_2(x) \overset{\text{def}}{=} 1 - \sqrt{1-(x-1)^2}.$$

Again, it is easy to check that $dg_2(0;1) = -\infty$ and $dg_2(2;-1) = -\infty$. □

However, problems occur at the boundary of the set U and Theorem 4.4 will show that a convex function f defined on a convex neighborhood of a point x is semidifferentiable in x. If, in addition, f is continuous at x, f will be Hadamard semidifferentiable at x (see Theorem 4.7). Finally, Theorem 4.8 will show that, in finite dimensions, a convex function defined in a convex set U is continuous on its interior $\operatorname{int} U$. A complete characterization of a convex function from the existence of semidifferentials and associated conditions on them will be given in Theorem 4.5.

4.2.1 Convexity and Semidifferentiability

Theorem 4.4. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on a convex neighborhood $V(x)$ of a point $x \in \mathbb{R}^n$. Then $df(x;v)$ exists in all directions $v \in \mathbb{R}^n$ and*

$$\forall v \in \mathbb{R}^n, \quad df(x;v) + df(x;-v) \geq 0. \tag{4.5}$$

Proof. (i) *Existence.* Given $v \in \mathbb{R}^n$, there exists α_0, $0 < \alpha_0 < 1$, such that $x - \alpha v \in V(x)$, $0 < \alpha \leq \alpha_0$, and there exists θ_0, $0 < \theta_0 < 1$, such that $x + \theta v \in V(x)$, $0 < \theta \leq \theta_0$. Fix α, $0 < \alpha \leq \alpha_0$. We first show that

$$\forall \theta,\ 0 < \theta < \theta_0, \quad \frac{f(x) - f(x-\alpha v)}{\alpha} \leq \frac{f(x+\theta v) - f(x)}{\theta}. \tag{4.6}$$

Indeed, x can be written as

$$x = \frac{\alpha}{\alpha+\theta}(x+\theta v) + \frac{\theta}{\alpha+\theta}(x-\alpha v)$$

and, by convexity,

$$f(x) \leq \frac{\alpha}{\alpha+\theta} f(x+\theta v) + \frac{\theta}{\alpha+\theta} f(x-\alpha v)$$

or, by rearranging,

$$\frac{\theta}{\theta+\alpha}[f(x) - f(x-\alpha v)] \leq \frac{\alpha}{\theta+\alpha}[f(x+\theta v) - f(x)]$$

and hence we get (4.6). Define

$$\varphi(\theta) \overset{\text{def}}{=} \frac{f(x+\theta v) - f(x)}{\theta}, \quad 0 < \theta < \theta_0,$$

and show that φ is monotone increasing. For all θ_1 and θ_2, $0 < \theta_1 < \theta_2 < \theta_0$,

$$\begin{aligned}
&f(x+\theta_1 v) - f(x) = f\left(\frac{\theta_1}{\theta_2}(x+\theta_2 v) + \left(1 - \frac{\theta_1}{\theta_2}\right)x\right) - f(x) \\
&\leq \frac{\theta_1}{\theta_2} f(x+\theta_2 v) + \left(1 - \frac{\theta_1}{\theta_2}\right) f(x) - f(x) \leq \frac{\theta_1}{\theta_2}[f(x+\theta_2 v) - f(x)] \\
&\Rightarrow \varphi(\theta_1) \leq \varphi(\theta_2).
\end{aligned}$$

Since the function $\varphi(\theta)$ is increasing and bounded below for $\theta \in]0,\theta_0[$, the limit as θ goes to 0 exists. By definition, it coincides with the semidifferential $df(x;v)$.

(ii) Given $v \in \mathbb{R}^n$, there exists α_0, $0 < \alpha_0 < 1$, such that

$$\frac{f(x) - f(x - \alpha v)}{\alpha} \leq df(x;v), \; 0 < \alpha \leq \alpha_0$$

From part (i) for all $v \in \mathbb{R}^n$, $df(x;v)$ and $df(x;-v)$ exist. Letting α go to 0, we get

$$-df(x;-v) = -\lim_{\alpha \searrow 0} \frac{f(x - \alpha v) - f(x)}{\alpha} \leq df(x;v)$$

and the inequality $df(x;-v) + df(x;v) \geq 0$. □

Corollary 1. *If $f : \mathbb{R}^n \to \mathbb{R}$ is convex in $\mathbb{R}^n$, then for all $v \in \mathbb{R}^n$ and all $x \in \mathbb{R}^n$,*

$$f(x) - f(x - v) \leq -df(x;-v) \leq df(x;v) \leq f(x+v) - f(x). \tag{4.7}$$

This theorem is the first step towards the complete characterization of a convex function and the relaxation of the conditions of Theorem 4.2.

Theorem 4.5. *Let $U \subset \mathbb{R}^n$ be open convex and $f : \mathbb{R}^n \to \mathbb{R}$. Then f is convex* (resp., *strictly convex*) *on U if and only if the following conditions are satisfied:*

(a) $\forall x \in U, \forall v \in \mathbb{R}^n$, $df(x;v)$ *exists;*

(b) $\forall x \in U, \forall v \in \mathbb{R}^n$, $df(x;v) + df(x;-v) \geq 0$*;*

(c) $\forall x, y \in U,\ f(y) \geq f(x) + df(x;y-x)$

(resp. $\forall x, y \in U,\ x \neq y,\ f(y) > f(x) + df(x;y-x)$).

Proof. ($\Rightarrow$) If f is convex, conditions (a) and (b) are verified by Theorem 4.4 at all points of the open convex U. As for condition (c), let x and y be two points of U and θ, $0 < \theta \leq 1$. As f is convex on U,

$$\begin{aligned} f(\theta y + (1-\theta)x) &\leq \theta f(y) + (1-\theta) f(x) \\ \Rightarrow\ f(x + \theta(y-x)) - f(x) &\leq \theta\,[f(y) - f(x)]. \end{aligned}$$

By dividing by θ and going to the limit as θ goes to 0, we get

$$df(x;y-x) \leq f(y) - f(x).$$

When f is strictly convex, choose $x \neq y$ and θ, $0 < \theta < 1$. Then

$$df(x;y-x) = \frac{1}{\theta} df(x;\theta(y-x)) \leq \frac{1}{\theta}[f(x + \theta(y-x)) - f(x)].$$

But

$$\begin{aligned} f(x + \theta(y-x)) &= f(\theta y + (1-\theta)x) < \theta f(y) + (1-\theta) f(x) \\ \Rightarrow\ df(x;y-x) &\leq \frac{1}{\theta}[f(x+\theta(y-x)) - f(x)] < \frac{1}{\theta}\theta(f(y) - f(x)) = f(y) - f(x). \end{aligned}$$

($\Leftarrow$) Apply the inequality in (c) to x and $x+\theta(y-x)$ and to y and $x+\theta(y-x)$ with $\theta \in [0,1]$:

$$\begin{aligned} f(x) &\geq f(x+\theta(y-x)) + df(x+\theta(y-x); -\theta(y-x)) \\ f(y) &\geq f(x+\theta(y-x)) + df(x+\theta(y-x); (1-\theta)(y-x)). \end{aligned}$$

Multiply the first inequality by $1-\theta$ and the second by θ and add them up:

$$\begin{aligned} (1-\theta)f(x)+\theta f(y) \geq f(x+\theta(y-x)) &+ (1-\theta)\theta df(x+\theta(y-x); -(y-x)) \\ &+ \theta(1-\theta) df(x+\theta(y-x); y-x). \end{aligned}$$

By using (b), we get

$$f(\theta y+(1-\theta)x) \leq \theta f(y)+(1-\theta)f(x).$$

In the strictly convex case, $y \neq x$ and $\theta \in]0,1[$ yield

$$x \neq x+\theta(y-x) \quad \text{and} \quad y \neq x+\theta(y-x).$$

The previous steps can be repeated with a strict inequality. $\square$

Theorem 4.6. *Let $U \subset \mathbb{R}^n$ be open convex and $f : \mathbb{R}^n \to \mathbb{R}$ be convex on U. For each $x \in U$, the function*

$$v \mapsto df(x;v) : \mathbb{R}^n \to \mathbb{R} \tag{4.8}$$

is positively homogeneous, convex, and subadditive; that is,

$$\forall v,w \in \mathbb{R}^n, \quad df(x;v+w) \leq df(x;v)+df(x;w). \tag{4.9}$$

Proof. We want to show that for all α, $0 \leq \alpha \leq 1$, and $v,w \in \mathbb{R}^n$,

$$df(x;\alpha v+(1-\alpha)w) \leq \alpha df(x;v)+(1-\alpha)df(x;w).$$

Since $x \in U$ and U is open and convex,

$$\begin{aligned} &\exists \theta_0,\, 0<\theta_0<1, \text{ such that } \forall \theta,\, 0<\theta\leq\theta_0, \quad x+\theta v \in U \text{ and } x+\theta w \in U \\ &\Rightarrow \forall 0 \leq \alpha \leq 1, \quad x+\theta(\alpha v+(1-\alpha)w) = \alpha(x+\theta v)+(1-\alpha)(x+\theta w) \in U \end{aligned}$$

and by convexity of f,

$$\begin{aligned} f(x+\theta(\alpha v+(1-\alpha)w)) &= f(\alpha[x+\theta v]+(1-\alpha)[x+\theta w]) \\ &\leq \alpha f(x+\theta v)+(1-\alpha)f(x+\theta w) \\ \Rightarrow [f(x+\theta(\alpha v&+(1-\alpha)w)) - f(x)] \\ \leq \alpha\,[f(x+\theta v)&-f(x)]+(1-\alpha)\,[f(x+\theta w)-f(x)]. \end{aligned}$$

Dividing by θ and going to the limit as θ goes to 0, we get

$$df(x;\alpha v+(1-\alpha)w) \leq \alpha df(x;v)+(1-\alpha)df(x;w).$$

Combining the positive homogeneity and the convexity,

$$\begin{aligned} df(x;v+w) &= df\left(x; \frac{1}{2}2v+\frac{1}{2}2w\right) \\ &\leq \frac{1}{2}df(x;2v)+\frac{1}{2}df(x;2w) = df(x;v)+df(x;w), \end{aligned}$$

we get the subadditivity. $\square$

4.2.2 Convexity and Continuity

The notion of convex function in a convex set U naturally extends to sets U that are not necessarily convex.

Definition 4.2.
$f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is *locally convex* on $U \subset \mathbb{R}^n$ if f is convex on all convex subsets V of U. It is *locally concave* on $U \subset \mathbb{R}^n$ if $-f$ is locally convex in U. □

Theorem 4.7. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be locally convex on a neighborhood of a point x of $\mathbb{R}^n$. The following conditions are equivalent.*

(i) *f is bounded above on a neighborhood of x.*

(ii) *There exist a neighborhood $W(x)$ of x and a constant $c(x) > 0$ such that*

$$\forall y \in W(x), \quad \forall v, w \in \mathbb{R}^n, \quad |df(y;w) - df(y;v)| \leq c(x)\|w - v\|. \tag{4.10}$$

(iii) *f is Lipschitzian at x; that is, there exist $c(x) > 0$ and a neighborhood $W(x)$ of x such that*

$$\forall y, z \in W(x), \quad |f(y) - f(z)| \leq c(x)\,\|y - z\|.$$

(iv) *f is Hadamard semidifferentiable on a neighborhood of x.*

(v) *$d_H f(x;0)$ exists.*

(vi) *f is continuous at x.*

Proof. (i) $\Rightarrow$ (ii) By assumption, there exist a neighborhood $V(x)$ of x and a constant $d(x) \in \mathbb{R}$ such that

$$\forall y \in U, \quad f(y) \leq d(x)$$

and since $V(x)$ is a neighborhood of x, there exists $\eta > 0$ such that $B_{2\eta}(x) \subset V(x)$. Choose $c(x) = d(x) - f(x)$. From the previous discussion, $c(x) \geq 0$. We get

$$\boxed{\forall z \in B_{2\eta}(x), \quad f(z) - f(x) \leq c(x).} \tag{4.11}$$

By convexity of f on $B_{2\eta}(x)$ and property (c) of Theorem 4.5,

$$\boxed{df(x;z-x) \leq f(z) - f(x) \leq c(x),} \quad \Rightarrow \forall y \in B_{2\eta}(0), \quad df(x;y) \leq c(x). \tag{4.12}$$

For all $w \in \mathbb{R}^n$, $w \neq 0$, $\eta\, w/\|w\| \in B_{2\eta}(0)$ and

$$df\left(x; \eta \frac{w}{\|w\|}\right) \leq c(x).$$

Since $v \mapsto df(x;v)$ is positively homogeneous,

$$\forall w \in \mathbb{R}^n, \quad df(x;w) \leq \frac{c(x)}{\eta}\|w\|. \tag{4.13}$$

By convexity of f and Theorem 4.4,

$$-df(x;w) \le df(x;-w) \le \frac{c(x)}{\eta}\|-w\| = \frac{c(x)}{\eta}\|w\|$$

and, combining this last inequality with inequality (4.13), we get

$$\boxed{\forall w \in \mathbb{R}^n, \quad |df(x;w)| \le \frac{c(x)}{\eta}\|w\|.} \tag{4.14}$$

But we want more. Always by convexity of f, we get from (4.9) in Theorem 4.6 that $v \mapsto df(x;v)$ is subadditive. Therefore, for all v and w in $\mathbb{R}^n$,

$$\begin{aligned} df(x;w) - df(x;v) &= df(x;v+w-v) - df(x;v) \\ &\le df(x;v) + df(x;w-v) - df(x;v) = df(x;w-v). \end{aligned}$$

Then, using inequality (4.14), we get

$$df(x;w) - df(x;v) \le df(x;w-v) \le |df(x;w-v)| \le \frac{c(x)}{\eta}\|w-v\|.$$

By repeating this argument and interchanging the roles of v and w, we get

$$\forall v,w \in \mathbb{R}^n, \quad |df(x;w) - df(x;v)| \le \frac{c(x)}{\eta}\|w-v\|.$$

We have proved the result at x. We now extend it to a neighborhood of x. Always by convexity of f and property (b) of Theorem 4.5, we have from (4.12) and (4.13), for all $z \in B_\eta(x)$,

$$f(x) - f(z) \le -df(x;z-x) \le df(x;x-z) \le \frac{c(x)}{\eta}\|x-z\| \le c(x).$$

Therefore, by combining this inequality with inequality (4.11),

$$\forall y \in B_\eta(x), \quad |f(y) - f(x)| \le c(x).$$

Repeat the proof by replacing x by some $y \in B_\eta(x)$. Indeed, for $z \in B_\eta(y)$, we get $z \in B_{2\eta}(x)$ and

$$f(z) - f(y) \le \underbrace{f(z) - f(x)}_{\le c(x)} + \underbrace{|f(y) - f(x)|}_{\le c(x)} \le 2c(x)$$

and f is bounded above on a neighborhood $B_\eta(y)$ of y by the constant $d(y) = f(y) + 2c(x)$. So we are back to the previous conditions with y in place of x, $2c(x)$ in place of $c(x)$, and $B_\eta(y)$ in place of $B_{2\eta}(x)$. By the same technical steps as above, we finally get for all $y \in B_\eta(x)$,

$$\forall v,w \in \mathbb{R}^n, \quad |df(y;w) - df(y;v)| \le \frac{2c(x)}{\eta/2}\|w-v\| = \frac{4c(x)}{\eta}\|w-v\|.$$

Choose the neighborhood $W(x) = B_\eta(x)$.

(ii) $\Rightarrow$ (iii). By definition, $df(y;0) = 0$ and from (4.10),

$$\forall y \in W(x),\ \forall w \in \mathbb{R}^n, \quad |df(y;w)| = |df(y;w) - 0| \le c(x)\|w\|.$$

Since $W(x)$ is a neighborhood of x, there exists $\eta > 0$ such that $B_\eta(x) \subset W(x)$. From this we have for y_1 and $y_2 \in B_\eta(x)$,

$$f(y_2) - f(y_1) \le -df(y_2; y_1 - y_2) \le df(y_2; y_2 - y_1) \le c(x)\|y_2 - y_1\|.$$

By interchanging the roles of y_1 and y_2, we get

$$\forall y_1, y_2 \in B_\eta(x), \ |f(y_2) - f(y_1)| \le c(x)\|y_2 - y_1\|.$$

So, the function f is Lipschitzian at x.

(iii) $\Rightarrow$ (iv) By definition, there exist $\eta > 0$ and $c(x) > 0$ such that

$$B_\eta(x) \subset W(x) \text{ and } \forall y, z \in B_\eta(x), \quad |f(y) - f(z)| \le c(x)\,\|y - z\|.$$

For all $y \in B_\eta(x)$, there exists $\varepsilon > 0$ such that $B_\varepsilon(y) \subset B_\eta(x)$. Then

$$\forall z_1, z_2 \in B_\varepsilon(y), \quad |f(z_2) - f(z_1)| \le c(x)\,\|z_2 - z_1\|$$

and f is Lipschitzian and convex on $B_\varepsilon(y)$. Therefore, as f is convex on $B_\eta(x)$, for all $v \in \mathbb{R}^n$, the semidifferential $df(y;v)$ exists and, since f is Lipschitzian on $B_\eta(x)$, the semidifferential $d_H f(y;v)$ exists by Theorem 3.5. Therefore, we have the result with $W(x) = B_\eta(x)$.

(iv) $\Rightarrow$ (v) is obvious by choosing $v = 0$ at x.

(v) $\Rightarrow$ (vi) By Theorem 3.7.

(vi) $\Rightarrow$ (i) If f is continuous at x, then

$$\forall \varepsilon > 0, \exists \eta > 0, \forall y, \|y - x\| < \eta, \quad |f(y) - f(x)| < \varepsilon$$

and, for y on a neighborhood $W(x) = B_\eta(x)$ of x, $f(y) \le |f(y)| \le |f(x)| + \varepsilon$. Therefore, the function f is bounded on a neighborhood $W(x)$ of x. □

Since we have established that a continuous convex function is Hadamard semidifferentiable, we complete this section by proving that convex functions on convex sets U are continuous in the interior of U. We need the following lemma.

Lemma 4.1. *Given an open ball $B_\eta(x)$ at x of radius $\eta > 0$, there exist $x_0, x_1, \dots, x_n$ in $B_\eta(x)$ and $\varepsilon > 0$ such that*

$$B_\varepsilon(x) \subset M \stackrel{\text{def}}{=} \left\{ \sum_{i=0}^{n} \lambda_i x_i : \lambda_i \ge 0,\ 0 \le i \le n,\ \sum_{i=0}^{n} \lambda_i = 1 \right\}. \tag{4.15}$$

Proof. Since $B_\eta(x)$ is a neighborhood of x in $\mathbb{R}^n$, choose $x_1, \dots, x_n$ in $B_\eta(x)$ such that the directions $x_1 - x$, $x_2 - x, \dots, x_n - x$ are linearly independent and such that

$$\|x_i - x\| = \eta/2n, \ 1 \le i \le n.$$

Also choose

$$x_0 \stackrel{\text{def}}{=} x - \sum_{i=1}^{n} (x_i - x). \tag{4.16}$$

It is readily verified that

$$\|x_0 - x\| \leq \eta/2 \ \Rightarrow \ x_0 \in B_\eta(x).$$

Moreover, by definition of M,

$$x = \sum_{i=0}^{n} \frac{1}{n+1} x_i \in M.$$

We want to show that

$$\exists \varepsilon > 0 \quad \text{such that} \quad B_\varepsilon(x) \subset M. \tag{4.17}$$

We proceed by contradiction. If (4.17) is not verified, then

$$\forall m \geq 1, \exists x_m \text{ such that } \|x_m - x\| = 1/m \text{ and } x_m \notin M.$$

Since, by construction, M is a closed convex set, we can apply the separation theorem (Theorem 7.8(ii)) of Chapter 2 to separate x_m and M:

$$\exists p_m, \|p_m\| = 1, \text{ such that } \forall y \in M, \quad p_m \cdot y \leq p_m \cdot x_m. \tag{4.18}$$

As m goes to infinity, x_m goes to x. But the sequence $\{p_m\}$ belongs to the compact sphere of radius one. So, there exist a subsequence $\{p_{m_k}\}$ of $\{p_m\}$ and a point p, $\|p\| = 1$, such that $p_{m_k} \to p, \ k \to \infty$. By going to the limit in inequality (4.18) for the subsequence as k goes to infinity, we get

$$\exists p, \|p\| = 1, \quad p \cdot y \leq p \cdot x, \quad \forall y \in M. \tag{4.19}$$

In particular, for i, $1 \leq i \leq n$, $x_i \in M$ and

$$p \cdot (x_i - x) \leq 0, \ \ 1 \leq i \leq n. \tag{4.20}$$

By definition of M and inequality (4.19),

$$\begin{aligned} & p \cdot \left(\sum_{i=0}^{n} \lambda_i x_i \right) \leq p \cdot x, \ \ \forall \lambda_i \geq 0, \ \ 0 \leq i \leq n, \ \ \sum_{i=0}^{n} \lambda_i = 1 \\ \Rightarrow \ & p \cdot \left(\sum_{i=0}^{n} \lambda_i (x_i - x) \right) \leq 0, \ \ \forall \lambda_i \geq 0, \ \ 0 \leq i \leq n, \ \ \sum_{i=0}^{n} \lambda_i = 1. \end{aligned} \tag{4.21}$$

By isolating x_0,

$$p \cdot \left(\sum_{i=0}^{n} \lambda_i (x_i - x) \right) = p \cdot (\lambda_0 (x_0 - x)) + p \cdot \left(\sum_{i=1}^{n} \lambda_i (x_i - x) \right)$$

and, by definition of x_0,

$$\begin{aligned} p \cdot \left(\sum_{i=0}^{n} \lambda_i (x_i - x) \right) &= p \cdot \left(-\lambda_0 \sum_{i=1}^{n} (x_i - x) \right) + p \cdot \left(\sum_{i=1}^{n} \lambda_i (x_i - x) \right) \\ &= p \cdot \left(\sum_{i=1}^{n} (\lambda_i - \lambda_0)(x_i - x) \right) = p \cdot \left(\sum_{i=0}^{n} (\lambda_i - \lambda_0)(x_i - x) \right). \end{aligned}$$

Then (4.21) becomes

$$p \cdot \left[\sum_{i=0}^{n} (\lambda_i - \lambda_0)(x_i - x) \right] \leq 0, \quad \forall \lambda_i \geq 0, \quad 0 \leq i \leq n, \quad \sum_{i=0}^{n} \lambda_i = 1. \tag{4.22}$$

Given i, $1 \leq i \leq n$, choose

$$\lambda_0 = \frac{1 + 1/2n}{1+n} > 0, \quad \lambda_i = \frac{1}{2(n+1)} > 0 \quad \text{and} \quad \lambda_j = \lambda_0, \, j \neq i, 1 \leq j \leq n.$$

It is readily checked that

$$\sum_{j=0}^{n} \lambda_j = n\lambda_0 + \lambda_i = \frac{1+2n}{2(1+n)} + \frac{1}{2(1+n)} = 1.$$

By substituting in (4.22), we get

$$0 \geq p \cdot (\lambda_i - \lambda_0)(x_i - x) = -\frac{1}{2n} p \cdot (x_i - x) \quad \Rightarrow p \cdot (x_i - x) \geq 0.$$

Hence, from (4.20),

$$p \cdot (x_i - x) \leq 0 \quad \text{and} \quad p \cdot (x_i - x) \geq 0 \quad \Rightarrow p \cdot (x_i - x) = 0.$$

Since the n vectors $x_1 - x, \ldots, x_n - x$ are linearly independent, any $y \in \mathbb{R}^n$ can be represented as a linear combination of those n vectors. Then

$$p \cdot y = 0, \; \forall y \in \mathbb{R}^n \quad \Rightarrow p = 0$$

that contradicts the fact that $\|p\| = 1$. Hence, the assertion (4.17) is true. □

Theorem 4.8. *A function $f : \mathbb{R}^n \to \mathbb{R}$ locally convex on $U \subset \mathbb{R}^n$ is locally Lipschitzian and Hadamard semidifferentiable in the interior* $\operatorname{int} U$ *of* U.

Proof. For a locally convex function on U and an interior point $x \in U$, there exists an open ball $B_r(x) \subset U$ on which f is convex. So it is sufficient to prove the theorem for a convex function in a convex U. If $\operatorname{int} U \neq \varnothing$, then for all $x \in \operatorname{int} U$, there exists $\eta > 0$ such that $B_\eta(x) \subset \operatorname{int} U$. By Lemme 4.1, there exist $x_0, x_1, \ldots, x_n$ in $B_\eta(x)$ and $\varepsilon > 0$ such that $B_\varepsilon(x) \subset M$. So the set M is a neighborhood of x. To prove the continuity of f at x, we show that f is bounded above on M and apply Theorem 4.7 (equivalence of (i) and (v)). By Jensen's inequality (7.5) from Theorem 7.4(i) of Chapter 2, for a convex function f,

$$f\left(\sum_{i=0}^{n} \lambda_i x_i\right) \leq \sum_{i=0}^{n} \lambda_i f(x_i), \; \forall \lambda_i \geq 0, \; 0 \leq i \leq n, \; \sum_{i=0}^{n} \lambda_i = 1.$$

Hence,

$$\forall y \in M, \quad f(y) \leq \max\{f(x_i) : 0 \leq i \leq n\} < +\infty$$

and f is bounded above on M. The continuity and the Hadamard semidifferentiability now follow from Theorem 4.7. □

4.3 ▶ Lower Hadamard Semidifferential at a Boundary Point of the Domain

Section 4.2 provides a fairly complete account of the properties of a convex function f at interior points of its domain dom f. However, for the infimum problem, the minimizer may very well occur on a part of the boundary $\partial(\text{dom } f)$ contained in dom f where the semidifferential $d_H f(x;v)$ does not exist (see Example 9.1 of Chapter 2 where the minimizer $(0,0) \in \text{dom } f \cap \partial(\text{dom } f)$).

When $\text{int}\,(\text{dom } f) = \varnothing$, the properties of convex functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ obtained in section 4.2 naturally extend from interior points of dom f in $\mathbb{R}^n$ to points of the relative interior ri(dom f) in aff (dom f) by considering the function f as a function on aff (dom) rather than on $\mathbb{R}^n$. Recall, from section 8.2 in Chapter 2 (Definition 8.2 and Lemma 8.2), that we can associate with f the function

$$y \mapsto f_x(y) \stackrel{\text{def}}{=} f(y+x) - f(x) : S(\text{dom } f) \to \mathbb{R} \tag{4.23}$$

which is convex with domain dom $f_x = \text{dom } f - x$, where $S(\text{dom } f) = \text{aff}\,(\text{dom } f) - x$ is the linear subspace associated with aff (dom f). If dom $f \neq \varnothing$ is not a singleton, ri (dom f) is a convex open subset in the affine subspace aff (dom f) and the previous results readily extend to that case. We single out a few properties.

Theorem 4.9. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a convex function with* dom $f \neq \varnothing$ *that is not a singleton.*

(i) *For each $x \in \text{ri}\,(\text{dom } f)$,*

$$\liminf_{y \to x} f(y) = \lim_{y \in \text{dom } f \to x} f(y) = f(x). \tag{4.24}$$

(ii) *For each $x \in \text{ri}\,(\text{dom } f)$, $df(x,v)$ exists in all directions $v \in S(\text{dom } f)$ and*

$$\forall v \in S(\text{dom } f), \quad df(x;v) + df(x;-v) \geq 0 \tag{4.25}$$

$$\forall x, y \in \text{ri}\,(\text{dom } f), \quad f(y) \geq f(x) + df(x; y-x). \tag{4.26}$$

The function

$$v \mapsto df(x;v) : S(\text{dom } f) \to \mathbb{R} \tag{4.27}$$

is positively homogeneous, convex, subadditive, and Lipschitzian.

Even if this result is interesting, the minimizer can still occur on the boundary $\partial(\text{dom } f)$ and not on the relative interior ri (dom f). In addition, it would be nice to avoid working with relative interiors and directions in the subspace $S(\text{dom } f)$.

To get around the fact that the nice semidifferential $d_H f(x;v)$ is slightly too restrictive since it does not exist at boundary points of dom f, we turn to the lower notion $\underline{d}_H f(x;v)$ which is much better suited for our purposes since, as a liminf, it always exists at every point $x \in \text{dom } f$ where it can be finite or infinite. This can be easily seen by computing $\underline{d}_H f((0,0);(x,y))$ for the function of Example 9.1 in Chapter 2 at the point

$(0,0) \in \mathrm{dom}\, f \cap \partial(\mathrm{dom}\, f)$ where $d_H f((0,0);(x,y))$ does not exist, but

$$\underline{d}_H f((0,0);(x,y)) = 0 \text{ for all } (x,y) \in \mathrm{dom}\, f.$$

The next theorem generalizes the important property (4.1) of Theorem 4.1.

Theorem 4.10. *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a convex function,* $\mathrm{dom}\, f \neq \varnothing$.

(i) *For each* $x \in \mathrm{dom}\, f$,

$$\forall y \in \mathbb{R}^n, \quad \underline{d}_H f(x; y-x) \leq f(y) - f(x) \tag{4.28}$$

except when $\mathrm{dom}\, f = \{x\}$ *and* $y = x$ *for which* $\underline{d}_H f(x;0) = +\infty > 0$.

(ii) *For each* $x \in \mathrm{ri}\,(\mathrm{dom}\, f)$ *and all* $v \in S(\mathrm{dom}\, f)$,

$$\underline{d}_H f(x;v) = \lim_{\substack{t \searrow 0 \\ w \in S(\mathrm{dom}\, f) \to v}} \frac{f(x+tw) - f(x)}{t} \tag{4.29}$$

is finite.

Proof. (i) For $y \in \mathbb{R}^n$, $0 < t < 1$, and $z \neq y$, by convexity of f,

$$\frac{f(x+t(z-x)) - f(x)}{t} \leq f(z) - f(x).$$

If $y = x$, $\|(x+t(z-x)) - x\| = t\,\|z-y\| > 0$ and if $y \neq x$, $\|(x+t(z-x)) - x\| \geq t\,\|y - x\| - t\,\|z-y\|$ that becomes strictly positive as $z \to y$. Therefore, as $t \to 0$ and $z \to y$, $x + t(z-x) \neq x$ and

$$\underline{d}_H f(x; y-x) = \liminf_{\substack{t \searrow 0 \\ z-x \to y-x}} \frac{f(x+t(z-x)) - f(x)}{t} \leq \liminf_{z \to y} f(z) - f(x).$$

By Lemma 7.2 of Chapter 2, $\liminf_{z \to y} f(z) \leq f(y)$ except in the case where $\mathrm{dom}\, f = \{x\}$ and $y = x$.

(ii) From the previous section when f is restricted to $\mathrm{aff}\,(\mathrm{dom}\, f)$. □

4.4 ► Semiconvex Functions and Hadamard Semidifferentiability

In order to make this book in tune with the ongoing modern research, we complete this section on convexity by introducing semiconcave functions that play a key role in the study of the Hamilton–Jacobi equations and in optimal control problems (see, for instance, the book of P. CANNARSA and C. SINESTRARI [**1**]). As in the case of the notion of concavity, a function is semiconvex if $-f$ is semiconcave and vice versa. It turns out that semiconvex functions are Hadamard semidifferentiable. Moreover, the Hadamard semidifferential of a semiconvex function coincides with the generalized directional derivative (3.38) briefly introduced in section 3.5.

In its simplest form a function $f : \mathbb{R}^n \to \mathbb{R}$ is *semiconvex* on a convex subset U of $\mathbb{R}^n$ if there exists a constant $c > 0$ such that $f(x) + c\,\|x\|^2$ is convex in U. This yields, for all $x, y \in U$ and all $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) - \lambda\, f(x) - (1-\lambda)\, f(y) \leq \lambda(1-\lambda)c\,\|x-y\|^2. \tag{4.30}$$

If f is continuous and semiconvex in $\mathbb{R}^n$, then the function

$$f_c(x) \overset{\text{def}}{=} f(x) + c\,\|x\|^2 \tag{4.31}$$

is convex and continuous in $\mathbb{R}^n$. As a result $f(x) = f_c(x) - c\,\|x\|^2$ is the difference of two convex continuous functions. In particular, f is Hadamard semidifferentiable in all points x and in all directions v:

$$d_H f(x;v) = d_H f_c(x;v) - 2c\,x \cdot v.$$

When U is not necessarily convex, then we say that $f : \mathbb{R}^n \to \mathbb{R}$ is *locally semiconvex* on U if there exists a constant $c > 0$ such that $f(x) + c\,\|x\|^2$ is convex in all convex subsets V of U as in Definition 4.2. The *semiconcavity* and the *local semiconcavity* are obtained by replacing f by $-f$ in the definitions of *semiconvexity* and *local semiconvexity*. As in the semiconvex case, a semiconcave function is also the difference of two convex functions.

The above definitions can be weakened in various ways to suit specific purposes. For instance, the square of the norm could be replaced by the norm: f is semiconvex on the convex U if there exists a constant $c > 0$ such that $f(x) + c\,\|x\|$ is convex on U. We obtain an inequality similar to (4.30):

$$\begin{aligned}
&f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda) f(y) \\
&\le c\,[\lambda\|x\| + (1-\lambda)\|y\| - \|\lambda x + (1-\lambda)y\|] \\
&\le c\,[\lambda(\|x\| - \|\lambda x + (1-\lambda)y\|) + (1-\lambda)(\|y\| - \|\lambda x + (1-\lambda)y\|)] \\
&\le c\,[\lambda\,|\|x\| - \|\lambda x + (1-\lambda)y\|| + (1-\lambda)\,|\|y\| - \|\lambda x + (1-\lambda)y\||] \\
&\le 2\lambda(1-\lambda)\,c\,\|x-y\|.
\end{aligned}$$

If f is continuous and semiconvex in $\mathbb{R}^n$, then the function

$$f_c(x) \overset{\text{def}}{=} f(x) + c\,\|x\| \tag{4.32}$$

is convex and continuous in $\mathbb{R}^n$. As a result $f(x) = f_c(x) - c\,\|x\|$ is again the difference of two convex continuous functions. In particular, f is Hadamard semidifferentiable at all points x and in all directions v,

$$d_H f(x;v) = d_H f_c(x;v) - d_H n(x;v), \quad \text{where } n(x) \overset{\text{def}}{=} \|x\|,$$

since the norm is Hadamard semidifferentiable.

The definition of a semiconcave function as the one of a semiconvex function has been generalized in the literature as follows (see, for instance, P. CANNARSA and C. SINESTRARI [**1**]).

Definition 4.3.
Let $f : \mathbb{R}^n \to \mathbb{R}$ and $U \subset \mathbb{R}^n$.

(i) f is *semiconvex* on U if there exists a nondecreasing function[53] $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\omega(\rho) \to 0$ as $\rho \to 0$ and

[53] The additional condition that $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ be upper semicontinuous in the definition of P. CANNARSA and C. SINESTRARI [**1**] has been dropped since it is redundant. If, for some reason, this property becomes necessary, simply replace ω by its usc regularization

$$\text{cl}_{\text{usc}}\,\omega(x) \overset{\text{def}}{=} \inf_{\substack{g \text{ usc and} \\ \omega \le g \text{ on } \mathbb{R}_+}} g(x) \tag{4.33}$$

(see Definition 4.3(ii) of Chapter 2). It retains the nondecreasing property, the continuity in 0, and the semiconvexity property (4.34) of ω since, by definition, $\omega \le \text{cl}_{\text{usc}}\,\omega$.

$$f(\lambda x+(1-\lambda)y) \le \lambda f(x)+(1-\lambda) f(y)+\lambda(1-\lambda)\omega(\|x-y\|)\|x-y\| \tag{4.34}$$

for all pairs $x, y \in U$ such that the segment $[x, y]$ is contained in U.

(ii) f is *semiconcave* on U if $-f$ is semiconvex on U. $\square$

Locally convex (resp., concave) functions are a special case of semiconvex (resp., semiconcave) functions for which $\omega = 0$, and the next theorem is a generalization of Theorem 4.8.

Theorem 4.11. *Let $U \subset \mathbb{R}^n$, $\operatorname{int} U \neq \varnothing$, and $f : \mathbb{R}^n \to \mathbb{R}$ be semiconvex on U.*

(i) *For each $x \in \operatorname{int} U$ and all $v \in \mathbb{R}^n$, $df(x;v)$ exists and is finite and $df(x;v) + df(x;-v) \ge 0$.*

(ii) *For each $x \in \operatorname{int} U$, there exists $r > 0$ such that f is Lipschitzian in $B_r(x)$, for all $v \in \mathbb{R}^n$, $d_H f(x;v)$ exists, and*

$$\forall v, w \in \mathbb{R}^n, \quad |d_H f(x;v) - d_H f(x;w)| \le c(x,r)\,\|v-w\|, \tag{4.35}$$

where $c(x,r)$ is the Lipschitz constant of f on $B_r(x)$.

(iii) *For each $x \in \operatorname{int} U$ and all $v \in \mathbb{R}^n$, $\overline{d}_C f(x;v) = d_H f(x;v)$.*

(iv) *Let $x \in U \cap \partial U$. If there exists $y \in U$, $y \neq x$, such that $[x,y] \subset U$, then*

$$\liminf_{z \to x} f(z) \le f(x). \tag{4.36}$$

For all $y \in U$, $y \neq x$, such that $[x,y] \subset U$, we have

$$\underline{d}_H f(x; y-x) \le f(y) - f(x) + \omega(\|y-x\|)\,\|y-x\|. \tag{4.37}$$

(v) *(Chain rules) If $g : \mathbb{R}^m \to \mathbb{R}^n$, where $m \ge 1$ is an integer, is such that $d_H g(y;w)$ exists and $g(y) \in \operatorname{int} U$, then*

$$d_H(f \circ g)(y;w) = d_H f(g(y); d_H g(y;w)); \tag{4.38}$$

if $x \in \operatorname{int} U$ and $h : \mathbb{R} \to \mathbb{R}^m$ is such that $d_H h(f(x); d_H f(x;v))$ exists, then

$$d_H(h \circ f)(x;v) = d_H h(f(x); d_H f(x;v)). \tag{4.39}$$

Corollary 1. *If the assumption of semiconvexity in Theorem* 4.11 *is replaced by the semiconcavity of f in U, the conclusions remain the same except in* (iii) *where*

$$d_H f(x;v) = \underline{d}_C f(x;v) = \liminf_{\substack{t \searrow 0 \\ y \to x}} \frac{f(y+tv) - f(y)}{t} \tag{4.40}$$

and in (iv) *where* (4.36) *and* (4.37) *are replaced by*

$$\limsup_{z \to x} f(z) \ge f(x) \tag{4.41}$$

$$\overline{d}_H f(x; y-x) \ge f(y) - f(x) - \omega(\|y-x\|)\,\|y-x\|. \tag{4.42}$$

Proof. Since $x \in \operatorname{int} U$, there exist $R > 0$ and a closed cube Q of diameter L centered at x such that $\overline{B_{2R}(x)} \subset Q \subset U$.

(i) Given $v \in \mathbb{R}^n$, there exists θ_0, $0 < \theta_0 < 1$, such that $x + \theta v \in B_{2R}(x)$, $0 < \theta \le \theta_0$, and there exists α_0, $0 < \alpha_0 < 1$, such that $x - \alpha v \in B_{2R}(x)$, $0 < \alpha \le \alpha_0$. Hence, x can be written as a convex combination:

$$\begin{aligned}
x &= \frac{\theta}{\alpha+\theta}(x-\alpha v) + \frac{\alpha}{\alpha+\theta}(x+\theta v) \\
\Rightarrow f(x) &- \frac{\theta}{\alpha+\theta} f(x-\alpha v) - \frac{\alpha}{\alpha+\theta} f(x+\theta v) \\
&\le \frac{\theta}{\alpha+\theta}\,\frac{\alpha}{\alpha+\theta}\,\omega((\alpha+\theta)\|v\|)\,(\alpha+\theta)\|v\| = \theta\,\alpha\,\frac{\omega((\alpha+\theta)\|v\|)}{(\alpha+\theta)}\,\|v\|.
\end{aligned}$$

This last inequality can be rewritten as

$$\begin{aligned}
&\frac{f(x)-f(x-\alpha v)}{\alpha} \le \frac{f(x+\theta v)-f(x)}{\theta} + \omega((\alpha+\theta)\|v\|)\,\|v\| \\
\Rightarrow\ &\frac{f(x)-f(x-\alpha v)}{\alpha} \le \liminf_{\theta \searrow 0} \frac{f(x+\theta v)-f(x)}{\theta} + \omega(2\alpha\,\|v\|)\,\|v\|,
\end{aligned} \tag{4.43}$$

since ω is monotone increasing and, as $\theta \to 0$, we have $0 < \theta < \alpha$ for θ small. So, the liminf is bounded below. For all θ_1 and θ_2, $0 < \theta_1 < \theta_2 < \theta_0$, using the definition of a semiconvex function,

$$\begin{aligned}
&f(x+\theta_1 v) - f(x) = f\left(\frac{\theta_1}{\theta_2}(x+\theta_2 v) + \left(1-\frac{\theta_1}{\theta_2}\right)x\right) - f(x) \\
&\le \frac{\theta_1}{\theta_2} f(x+\theta_2 v) + \left(1-\frac{\theta_1}{\theta_2}\right) f(x) - f(x) + \frac{\theta_1}{\theta_2}\left(1-\frac{\theta_1}{\theta_2}\right)\omega(\theta_2\|v\|)\,\theta_2\|v\| \\
&\le \theta_1\,\frac{f(x+\theta_2 v)-f(x)}{\theta_2} + \theta_1\left(1-\frac{\theta_1}{\theta_2}\right)\omega(\theta_2\|v\|)\,\|v\|
\end{aligned}$$

and dividing both sides by θ_1,

$$\begin{aligned}
&\frac{f(x+\theta_1 v)-f(x)}{\theta_1} \le \frac{f(x+\theta_2 v)-f(x)}{\theta_2} + \left(1-\frac{\theta_1}{\theta_2}\right)\omega(\theta_2\|v\|)\,\|v\| \\
\Rightarrow\ &\limsup_{\theta_1 \searrow 0} \frac{f(x+\theta_1 v)-f(x)}{\theta_1} \le \frac{f(x+\theta_2 v)-f(x)}{\theta_2} + \omega(\theta_2\|v\|)\,\|v\|
\end{aligned} \tag{4.44}$$

and the liminf and limsup are bounded above. As liminf was bounded below, both are bounded. Now, take the liminf in (4.44) as $\theta_2 \to 0$,

$$\limsup_{\theta_1 \searrow 0} \frac{f(x+\theta_1 v)-f(x)}{\theta_1} \le \liminf_{\theta_2 \searrow 0} \frac{f(x+\theta_2 v)-f(x)}{\theta_2} + 0,$$

since, by assumption, $\omega(\theta_2\|v\|) \to 0$ as $\theta_2 \to 0$. Therefore, $df(x;v)$ exists and is finite for all $v \in \mathbb{R}^n$.

Finally, going back to (4.43), take the limit as $\alpha \to 0$,

$$\lim_{\alpha\searrow 0} \frac{f(x)-f(x-\alpha v)}{\alpha} \le \lim_{\theta\searrow 0} \frac{f(x+\theta v)-f(x)}{\theta} + \lim_{\alpha\searrow 0} \omega(2\alpha\,\|v\|)\,\|v\|,$$
$$\Rightarrow -df(x;-v) \le df(x;v)+0 \quad \Rightarrow df(x;-v)+df(x;v) \ge 0,$$

since, by assumption, $\omega(2\alpha\|v\|) \to 0$ as $\alpha \to 0$.

(ii) We first prove that f is bounded above and below in $B_R(x)$. For the upper bound, denote by $x_1, x_2, \dots, x_{2^n}$ the 2^n vertices of Q. Let

$$M_0 \stackrel{\text{def}}{=} \max\{u(x_i) : 1 \le i \le 2^n\}.$$

For two consecutive vertices x_i and x_j of Q, using the semiconvexity identity (4.34),

$$\begin{aligned} &f(\lambda x_i + (1-\lambda)x_j) - \lambda f(x_i) - (1-\lambda) f(x_j) \\ &\le \lambda(1-\lambda)\,\omega(\|x_i - x_j\|)\,\|x_i - x_j\| \le M_0 + \omega(L)\frac{L}{4}. \end{aligned} \tag{4.45}$$

This shows that f is bounded above on the one-dimensional faces of Q. Repeat the procedure by taking any convex combination of two points lying on different one-dimensional faces to obtain $u(z) \le M_0 + \omega(L)L/4 + \omega(L)L/4 = M_0 + \omega(L)L/2$. Iterating this procedure n times, we get the existence of a constant M such that $u(z) \le M$ for all $z \in Q$.

For the lower bound, since $\overline{B_{2R}(x)} \subset Q$, $u(x) \le M$ in $\overline{B_{2R}(x)}$. For any $z \in \overline{B_{2R}(x)}$,

$$x = \frac{\|z-x\|}{2R+\|z-x\|}\left(x - 2R\frac{z-x}{\|z-x\|}\right) + \frac{2R}{2R+\|z-x\|}z.$$

Using the semiconvexity identity (4.34),

$$\begin{aligned} &f(x) - \frac{\|z-x\|}{2R+\|z-x\|} f\left(x - 2R\frac{z-x}{\|z-x\|}\right) - \frac{2R}{2R+\|z-x\|} f(z) \\ &\le \frac{\|z-x\|}{2R+\|z-x\|}\,\frac{2R}{2R+\|z-x\|}\,\omega(2R+\|z-x\|)(2R+\|z-x\|). \end{aligned}$$

After rearranging the terms, for any $z \in \overline{B_{2R}(x)}$,

$$\begin{aligned} f(z) \ge &\frac{2R+\|z-x\|}{2R} f(x) - \frac{\|z-x\|}{2R} f\left(x - 2R\frac{z-x}{\|z-x\|}\right) \\ &- \|z-x\|\,\omega(2R+\|z-x\|) \\ \ge &-2|f(x)| - |M| - 2R\,\omega(4R). \end{aligned}$$

So f is bounded below by $m = -2|f(x)| - |M| - 2R\,\omega(4R)$ and bounded above by M in $\overline{B_{2R}(x)}$.

Now pick any two distinct points y and z in $B_R(x)$. There exists y' and z' on the line going through y and z such that $\|y'-x\| = 2R$, $\|z'-x\| = 2R$, and the four points appear in the order y', y, z, and z' on that line:

$$\|y'-z\| = \|y'-y\| + \|y-z\| \text{ and } \|y-z'\| = \|y-z\| + \|z-z'\|$$
$$\Rightarrow \|y'-y\| \ge R \text{ and } \|z-z'\| \ge R.$$

This means that y is a convex combination of y' and z, and z is a convex combination of y and z':

$$y = \frac{\|y-z\|}{\|y'-y\|+\|y-z\|}y' + \frac{\|y'-y\|}{\|y'-y\|+\|y-z\|}z$$

$$z = \frac{\|z-z'\|}{\|y-z\|+\|z-z'\|}y + \frac{\|y-z\|}{\|y-z\|+\|z-z'\|}z'.$$

Using the semiconvexity inequality (4.34) associated with the first convex combination,

$$\begin{aligned}
& f(y) - \frac{\|y-z\|}{\|y'-y\|+\|y-z\|} f(y') - \frac{\|y'-y\|}{\|y'-y\|+\|y-z\|} f(z) \\
& \le \frac{\|y-z\|}{\|y'-y\|+\|y-z\|} \frac{\|y'-y\|}{\|y'-y\|+\|y-z\|} \omega(\|y'-z\|)\,\|y'-z\| \\
& \le \frac{\|y-z\|\,\|y'-y\|}{\|y'-y\|+\|y-z\|} \omega(\|y'-z\|) \\
\Rightarrow \quad & \frac{\|y-z\|}{\|y'-y\|+\|y-z\|}(f(y)-f(y')) + \frac{\|y'-y\|}{\|y'-y\|+\|y-z\|}(f(y)-f(z)) \\
& \le \frac{\|y-z\|\,\|y'-y\|}{\|y'-y\|+\|y-z\|} \omega(\|y'-z\|) \\
\Rightarrow \quad & \frac{f(y)-f(y')}{\|y'-y\|} + \frac{f(y)-f(z)}{\|y-z\|} \le \omega(\|y'-z\|) \\
\Rightarrow \quad & \frac{f(y)-f(z)}{\|y-z\|} \le \omega(\|y'-z\|) - \frac{f(y)-f(y')}{\|y'-y\|} \le \omega(3R) + \frac{M-m}{R}.
\end{aligned}$$

Similarly, for the second convex combination,

$$\begin{aligned}
& \frac{f(z)-f(y)}{\|y-z\|} + \frac{f(z)-f(z')}{\|z-z'\|} \le \omega(\|y-z'\|) \\
\Rightarrow \; & \frac{f(y)-f(z)}{\|y-z\|} \ge -\omega(\|y-z'\|) + \frac{f(z)-f(z')}{\|z-z'\|} \ge -\omega(3R) - \frac{M-m}{R}.
\end{aligned}$$

Finally, for all $y,z \in B_R(x)$,

$$|f(y)-f(z)| \le \left(\omega(3R) + \frac{M-m}{R}\right) \|y-z\|.$$

The existence of $d_H f(x;v)$ and the identity $d_H f(x;v) = df(x;v)$ are now a consequence of part (i) and Theorem 3.5 for a Lipschitzian function.

(iii) From part (ii), $d_H f(x;v)$ exists and, since f is locally Lipschitzian in $B_R(x)$, $\overline{d}_C f(x;v)$ exists. By definition,

$$\overline{d}_C f(x;v) = \limsup_{\substack{t\searrow 0\\ y\to x}} \frac{f(y+tv)-f(y)}{t} \ge \limsup_{t\searrow 0} \frac{f(x+tv)-f(x)}{t} = df(x;v).$$

It remains to prove the inequality in the other direction. Given $v \neq 0$, choose $\rho > 0$ such that $\rho < \min\{R/(1+\|v\|),1\}$ so that for all θ, $0 < \theta \le \rho$, and $y \in B_\rho(x)$,

$$\|y+\theta v - x\| \le \|y-x\| + \theta\|v\| < \rho + \rho\,\|v\| < R \quad \Rightarrow y+\theta v \in B_R(x).$$

By the same technique as in the proof of part (i), for all θ, $0 < \theta \le \rho < 1$,

$$\begin{aligned}\frac{f(y+\theta v)-f(y)}{\theta} &\le \frac{f(y+\rho v)-f(y)}{\rho} + \left(1-\frac{\theta}{\rho}\right)\omega(\rho\|v\|)\,\|v\| \\ &\le \frac{f(y+\rho v)-f(y)}{\rho} + \omega(\rho\|v\|)\,\|v\|.\end{aligned}$$

By Lipschitz continuity in $B_R(x)$,

$$\frac{f(x+\rho v)-f(x)}{\rho} \ge \frac{f(y+\rho v)-f(y)}{\rho} - \frac{2c}{\rho}\|y-x\|,$$

where c is the Lipschitz constant of f in $B_R(x)$. Combining the above two inequalities: for all $y \in B_\rho(x)$ and $0 < \theta \le \rho$,

$$\frac{f(x+\rho v)-f(x)}{\rho} \ge \frac{f(y+\theta v)-f(y)}{\theta} - \omega(\rho\|v\|)\,\|v\| - \frac{2c}{\rho}\|y-x\|.$$

Choose an arbitrary ε such that $0 < \varepsilon < 2c$. Then for any δ, $0 < \delta < \varepsilon\rho/(2c)$, we have $\delta < \rho$ and

$$\begin{aligned}\frac{f(x+\rho v)-f(x)}{\rho} &\ge \sup_{\substack{\|y-x\|<\delta \\ 0<\theta<\delta}} \frac{f(y+\theta v)-f(y)}{\theta} - \omega(\rho\|v\|)\,\|v\| - \varepsilon \\ \Rightarrow \lim_{\delta\searrow 0} \frac{f(x+\rho v)-f(x)}{\rho} &\ge \lim_{\delta\searrow 0} \sup_{\substack{\|y-x\|<\delta \\ 0<\theta<\delta}} \frac{f(y+\theta v)-f(y)}{\theta} - \varepsilon.\end{aligned}$$

Since both limits exist, by letting ε go to zero,

$$df(x;v) = \lim_{\delta\searrow 0} \frac{f(x+\rho v)-f(x)}{\rho} \ge \limsup_{\substack{\delta\searrow 0 \\ y\to x}} \frac{f(y+\delta v)-f(y)}{\delta} = \overline{d}_C f(x;v).$$

(iv) From the semiconvexity property, for $0 < t < 1$

$$\begin{aligned} f(x+t(y-x)) - f(x) &\le t\,[f(y)-f(x)+(1-t)\,\omega(\|y-x\|)\,\|y-x\|] \\ \liminf_{z\to x} f(z) - f(x) &\le \liminf_{t\searrow 0} f(x+t(y-x)) - f(x) \\ &\le \liminf_{t\searrow 0}\, t\,[f(y)-f(x)+(1-t)\,\omega(\|y-x\|)\,\|y-x\|] = 0.\end{aligned}$$

Always from the semiconvexity property, for $0 < t < 1$,

$$\begin{aligned}\frac{f(x+t(y-x))-f(x)}{t} &\le f(y)-f(x)+(1-t)\,\omega(\|y-x\|)\,\|y-x\| \\ \underline{d}_H f(x;y-x) &\le \liminf_{t\searrow 0} \frac{f(x+t(y-x))-f(x)}{t} \\ &\le \liminf_{t\searrow 0}\,[f(y)-f(x)+(1-t)\,\omega(\|y-x\|)\,\|y-x\|] \\ &= f(y)-f(x)+\omega(\|y-x\|)\,\|y-x\|.\end{aligned}$$

(v) From Theorem 3.4(ii). $\square$

Remark 4.2.
The proofs of parts (i), (ii), and the last part of (iii) follow those of P. CANNARSA and C. SINESTRARI [**1**, Thm. 3.2.1, p. 55, and Thm. 2.1.7, p. 33]. Part (iv) applies to the pre- and postcompositions with f under the assumptions of P. CANNARSA and C. SINESTRARI [**1**, Thm. 2.1.12, p. 37]. □

5 ▶ Semidifferential of a Parametrized Extremum

This section deals with the important issue of the semidifferentiability of an infimum (lower envelope) or a supremum (upper envelope) that depends on a set of parameters. We have already seen the rule for the lower and upper envelopes of a finite family of functions in Theorem 3.3(ii). We now generalize it to an infinite family at the price of a few assumptions.

A problem amenable to that formulation is the optimal shape of a column formulated by J. L. LAGRANGE [**1**] in 1770 and later studied by T. CLAUSEN [**1**] in 1849. It consists in finding the best profile of a vertical column to prevent buckling. It is one of the very early optimal design problems. Since Lagrange many authors have proposed solutions, but a complete theoretical and numerical solution was only given in 1992 by S. J. COX and M. L. OVERTON [**1**] using the generalized gradient. Other problems related to columns have been revisited in a series of papers by S. J. COX [**1**, **2**] and S. J. COX and C. M. MCCARTHY [**1**]. This problem can also be tackled by using the Hadamard semidifferential (see M. C. DELFOUR and J.-P. ZOLÉSIO [**1**]). The theorems and explicit formulae given in this section can be directly applied to get the Hadamard semidifferential of the discretized version of the *Euler buckling load* which is nothing but an eigenvalue depending on the shape parameters of the column.

The specialization of this section to the semidifferential of the least and greatest eigenvalues of a parametrized symmetric matrix will be given in section 2.2 of Chapter 4.

5.1 Semidifferential of an Infimum with respect to a Parameter

First, we introduce a simple but important theorem that will be used to get the formula of the semidifferential of an infimum. By combining with another property such as the Lipschitz continuity of the infimum with respect to the parameters, it yields the stronger Hadamard semidifferential of the infimum.

Consider the function

$$G\colon [0,\tau]\times X \to \mathbb{R} \tag{5.1}$$

for some $\tau > 0$ and some subset X of $\mathbb{R}^n$. For each t in $[0,\tau]$, define

$$g(t) \overset{\text{def}}{=} \inf\{G(t,x) : x \in X\} \quad \text{and} \quad X(t) \overset{\text{def}}{=} \{x \in X : G(t,x) = g(t)\}. \tag{5.2}$$

The objective is to find conditions for the existence and characterization of the limit

$$dg(0) \overset{\text{def}}{=} \lim_{t\searrow 0} \frac{g(t)-g(0)}{t} \tag{5.3}$$

when $X(t)$ is not empty for all $t, 0 \le t \le \tau$.

When $X(t) = \{x^t\}$ is a singleton, $0 \le t \le \tau$, and the semiderivative of x

$$\dot{x} = \lim_{t \searrow 0} \frac{x^t - x^0}{t} \tag{5.4}$$

is known, then it is easy to obtain $dg(0)$ under some differentiability of the functional $G(t,x)$ with respect to t and x. But when $\dot{x}$ is not readily available or when the sets $X(t)$ are not singletons, this direct approach fails or becomes very intricate. In this section, we present a theorem that gives an explicit expression for $dg(0)$. Its originality is that the differentiability of x^t is replaced by a sequential semicontinuity assumption on the *set-valued function* $X(t)$ and semicontinuity assumptions on the partial derivative of the functional $G(t,x)$ with respect to the parameter t. In other words, this technique does not require the knowledge of the derivative $\dot{x}$ of the minimizing elements x^t with respect to t.

Theorem 5.1. *Let X be a nonempty subset of $\mathbb{R}^n$, let $\tau > 0$, and let $G\colon [0,\tau[\times X \to \mathbb{R}$. Assume that the following conditions are verified:*

(H1) *for all t, $0 \le t < \tau$, $X(t) \neq \varnothing$;*

(H2) *at every x in $\bigcup_{t\in[0,\tau[} X(t)$,*

$$\begin{cases} \partial_t G(t,x) \overset{\text{def}}{=} \displaystyle\lim_{s\to 0} \frac{G(s,x)-G(t,x)}{s} & \textit{exists at each } t,\ 0<t<\tau, \\ \partial_t G(0,x) \overset{\text{def}}{=} \displaystyle\lim_{s\searrow 0} \frac{G(s,x)-G(0,x)}{s} & \textit{exists}; \end{cases}$$

(H3) *for any sequence $\{t_n\} \subset]0,\tau[$, $t_n \to t_0 = 0$, there exist $x_0 \in X(0)$ and a subsequence $\{t_{n_k}\}$ for which there exists $x_{n_k} \in X(t_{n_k})$ such that*

$$\text{(i) } x_{n_k} \to x^0 \quad \textit{and} \quad \text{(ii) } \liminf_{\substack{k\to\infty \\ t\searrow 0}} \partial_t G(t, x_{n_k}) \ge \partial_t G(0,x^0);$$

(H4) *at every x in $X(0)$,*

$$\limsup_{t\searrow 0} \partial_t G(t,x) \le \partial_t G(0,x).$$

Then there exists $x^0 \in X(0)$ such that

$$dg(0) = \lim_{t\searrow 0} \frac{g(t)-g(0)}{t} = \inf_{x\in X(0)} \partial_t G(0,x) = \partial_t G(0,x^0). \tag{5.5}$$

When $X(0)$ is a singleton $\{x_0\}$, $dg(0) = \partial_t G(0,x_0)$.

Remark 5.1.
In the literature, condition (H3) (i) is known as *sequential semicontinuity* for set-valued functions. It can be expressed in a more aesthetic way using the notion of *upper Painlevé–Kuratowski limit*: given a sequence $\{t_n\}$ such that $t_n \searrow 0$ as $n \to \infty$,

$$\operatorname*{Lim\,sup}_{t_n \searrow 0} X(t_n) \overset{\text{def}}{=} \left\{ x \in X : \liminf_{t_n \searrow 0} d_{X(t_n)}(x) = 0 \right\} \tag{5.6}$$

(see, for instance, J.-P. AUBIN and H. FRANKOWSKA [**1**, Dfn. 1.1.1 and the footnotes, p.17] and R. T. ROCKAFELLAR and R. J.-B. WETS [**1**]). Condition (H3) (i) is equivalent to the following condition:

$$\forall \{t_n\} \text{ such that } t_n \searrow 0, \quad \mathrm{X}(0) \cap \operatorname*{Lim\,sup}_{t_n \searrow 0} X(t_n) \neq \varnothing. \tag{5.7}$$

Indeed, if condition (H3) (i) is verified, then for the $x_0 \in X(0)$ and the subsequence

$$\begin{aligned} 0 \le d_{X(t_{n_k})}(x_0) \le \|x_0 - x_{n_k}\| \to 0 \quad &\Rightarrow \liminf_{t_n \searrow 0} d_{X(t_n)}(x_0) = 0 \\ &\Rightarrow x_0 \in \operatorname*{Lim\,sup}_{t_n \searrow 0} X(t_n). \end{aligned}$$

Conversely, if $x_0 \in X(0) \cap \limsup_{t_n \searrow 0} X(t_n)$,

$$\exists \{t_{n_k}\} \text{ such that } \lim_{k\to\infty} d_{X(t_{n_k})}(x_0) = \liminf_{t_n \searrow 0} d_{X(t_n)}(x_0) = 0.$$

Therefore, for each k, there exists $k' \ge k$ such that

$$d_{X(t_{n_{k'}})}(x_0) < 1/k \text{ and } \exists x_{t_{n_{k'}}} \text{ such that } \|x_0 - x_{t_{n_{k'}}}\| - d_{X(t_{n_{k'}})}(x_0) < 1/k,$$

$\|x_0 - x_{t_{n_{k'}}}\| < 2/k$, and $x_{t_{n_{k'}}} \to x_0$ as $k \to \infty$. □

Remark 5.2.
This theorem, and in particular the last part of property (5.5), extends a former result by B. LEMAIRE [**1**, Thm. 2.1, p. 38] in 1970, where sequential compactness of the set X was assumed. It also completes and extends Theorem 1 in the thesis of J.-P. ZOLÉSIO [**1**] in 1979 and the specific application to the semidifferentiability of the least eigenvalue presented at the NATO Advanced Institute by J.-P. ZOLÉSIO [**2**] in 1980. For a similar theorem in the framework of the generalized gradient the reader is referred to F. H. CLARKE [**2**, sec. 2.8, pp. 85–95] in 1983. □

Proof of Theorem 5.1.[54] (i) We first establish upper and lower bounds to the differential quotient

$$\frac{\Delta g(t)}{t}, \quad \Delta g(t) \overset{\text{def}}{=} g(t) - g(0).$$

Choose arbitrary x_0 in $X(0)$ and x_t in $X(t)$. Then, by definition,

$$\begin{aligned} G(t,x_t) = \quad g(t) &\le G(t,x_0), \\ -G(0,x_t) \le -g(0) &= -G(0,x_0). \end{aligned}$$

Add up the above two inequalities to obtain

$$G(t,x_t) - G(0,x_t) \le \Delta g(t) \le G(t,x_0) - G(0,x_0).$$

[54] This proof is quoted from M. C. DELFOUR and J.-P. ZOLÉSIO [**1**, Thm. 2.1, pp. 524–526].

By assumption (H2), there exist θ_t, $0 < \theta_t < 1$, and α_t, $0 < \alpha_t < 1$, such that

$$G(t,x_t) - G(0,x_t) = t\,\partial_t G(\theta_t t, x_t),$$
$$G(t,x_0) - G(0,x_0) = t\,\partial_t G(\alpha_t t, x_0),$$

and by dividing by $t > 0$,

$$\partial_t G(\theta_t t, x_t) \le \frac{\Delta g(t)}{t} \le \partial_t G(\alpha_t t, x_0). \tag{5.8}$$

(ii) Consider the lower and upper limits of the differential quotient:

$$\underline{d}g(0) = \liminf_{t \searrow 0} \frac{\Delta g(t)}{t}, \quad \overline{d}g(0) = \limsup_{t \searrow 0} \frac{\Delta g(t)}{t}.$$

There exists a sequence $\{t_n : 0 < t_n \le \tau\}$, $t_n \to 0$, such that

$$\lim_{n\to\infty} \frac{\Delta g(t_n)}{t_n} = \underline{d}g(0).$$

By assumption (H3), there exist $x^0 \in X(0)$ and a subsequence $\{t_{n_k}\}$ of $\{t_n\}$ for which there exists $x_{n_k} \in X(t_{n_k})$ such that $x_{n_k} \to x^0$ in $\mathcal{T}_X$ and

$$\liminf_{\substack{t \searrow 0 \\ k\to\infty}} \partial_t G(t, x_{n_k}) \ge \partial_t G(0, x^0).$$

So, from the first part of the estimate (5.8), we have for $t = t_{n_k}$,

$$\partial_t G(\theta_{t_{n_k}} t_{n_k}, x_{n_k}) \le \frac{\Delta g(t_{n_k})}{t_{n_k}}$$

and

$$\partial_t G(0,x^0) \le \liminf_{k\to\infty} \partial_t G(\theta_{t_{n_k}} t_{n_k}, x_{n_k}) \le \lim_{k\to\infty} \frac{\Delta g(t_{n_k})}{t_{n_k}} = \underline{d}g(0).$$

Therefore,

$$\exists x^0 \in X(0), \quad \partial_t G(0,x^0) \le \underline{d}g(0),$$

and

$$\inf_{x\in X(0)} \partial_t G(0,x) \le \partial_t G(0,x^0) \le \underline{d}g(0). \tag{5.9}$$

From the second part of (5.8) and assumption (H4), we also obtain

$$\forall x \in X(0), \quad \partial_t G(0,x) \ge \overline{d}g(0) \quad \Rightarrow \overline{d}g(0) \le \inf_{x\in X(0)} \partial_t G(0,x), \tag{5.10}$$

and necessarily

$$\inf_{x\in X(0)} \partial_t G(0,x) = \underline{d}g(0) = \overline{d}g(0) = \inf_{x\in X(0)} \partial_t G(0,x).$$

In particular, from (5.9) and (5.10),

$$\partial_t G(0,x^0) = dg(0) = \inf_{x\in X(0)} \partial_t G(0,x)$$

and x^0 is a minimizing point of $\partial_t G(0,\cdot)$. □

5.2 Infimum of a Parametrized Quadratic Function

Given an $n \times n$ symmetric matrix A, a vector $a \in \mathbb{R}^n$, and a nonempty subset E of $\mathbb{R}^n$, consider the infimum parametrized by A and a,

$$f(A,a) \overset{\text{def}}{=} \inf_{x \in E} Ax \cdot x + a \cdot x, \tag{5.11}$$

and the set of minimizers

$$X(0) \overset{\text{def}}{=} \{x \in E : Ax \cdot x + a \cdot x = f(A,a)\}. \tag{5.12}$$

Let B be another $n \times n$ symmetric matrix, $b \in \mathbb{R}^n$, and $t \geq 0$. Consider the perturbed problems

$$f(A+t\,B, a+tb) = \inf_{x \in S} [A+t\,B]x \cdot x + [a+t\,b] \cdot x \tag{5.13}$$

and the associated set of minimizers

$$X(t) \overset{\text{def}}{=} \{x \in E : [A+t\,B]x \cdot x + [a+t\,b] \cdot x = f(A+t\,B, a+tb)\}. \tag{5.14}$$

In a first step, Theorem 5.1 is used to find the expression of the semidifferential

$$df(A,a;B,b) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{f(A+t\,B,a+tb) - f(A,a)}{t} \tag{5.15}$$

when the limit exists. In a second step, we show that f is Lipschitzian and conclude that f is Hadamard semidifferentiable.

Theorem 5.2. *Let* Sym_n *be the vector space of all symmetric* $n \times n$ *matrices endowed with the matrix norm*

$$\|A\| \overset{\text{def}}{=} \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Assume that E *is nonempty and compact.*

(i) *The function* $(A,a) \mapsto f(A,a) : \mathrm{Sym}_n \times \mathbb{R}^n \to \mathbb{R}$ *is concave and Lipschitzian:*

$$\forall A, B \in \mathrm{Sym}_n, \forall a, b, \quad |f(B,b) - f(A,a)| \leq \|B - A\| + \|b - a\|.$$

(ii) *For all* $A, B \in \mathrm{Sym}_n$ *and* $a, b \in \mathbb{R}^n$

$$\begin{aligned} d_H f(A,a;B,b) &\overset{\text{def}}{=} \lim_{\substack{t \searrow 0 \\ (W,w) \to (B,b)}} \frac{f(A+tW, a+tw) - f(A,a)}{t} \\ &= \min_{x \in E(A,a)} Bx \cdot x + b \cdot x = \min_{x \in E(A,a)} x \otimes x \cdot\cdot B + b \cdot x, \end{aligned} \tag{5.16}$$

where $E(A,a) \overset{\text{def}}{=} \{x \in X : Ax \cdot x + a \cdot x = f(A,a)\}$,

$$(x \otimes y)_{ij} \overset{\text{def}}{=} x_i\, y_j \tag{5.17}$$

is the tensor product *of x and y, and*

$$X \cdot\cdot Y \overset{\text{def}}{=} \sum_{1 \le i,j \le n} X_{ij} Y_{ij} \tag{5.18}$$

is the double inner product *of two matrices. (Note that $x \otimes x$ is a positive semidefinite matrix of trace equal to $\|x\|^2$.)*

(iii) *Let the maps $\vec{p} \mapsto A(\vec{p}) : \mathbb{R}^k \to \mathrm{Sym}_n$ and $\vec{p} \mapsto a(\vec{p}) : \mathbb{R}^k \to \mathbb{R}^n$ be Hadamard semidifferentiable at $\vec{p}$. Consider the function*

$$\vec{p} \mapsto \ell(\vec{p}) \overset{\text{def}}{=} f(A(\vec{p}), a(\vec{p})) : \mathbb{R}^k \to \mathbb{R}. \tag{5.19}$$

Then for $\vec{v} \in \mathbb{R}^k$,

$$\begin{aligned} d_H \ell(\vec{p};\vec{v}) &= d_H f(A(\vec{p}), a(\vec{p}); d_H A(\vec{p};\vec{v}), d_H a(\vec{p};\vec{v})) \\ &= \min_{x \in E(A(\vec{p}),a(\vec{p}))} x \otimes x \cdot\cdot d_H A(\vec{p};\vec{v}) + x \cdot d_H a(\vec{p};\vec{v}). \end{aligned}$$

Remark 5.3.
Formula (5.16) gives a complete description of the nondifferentiability of $f(A,a)$. When $E(A,a)$ is not a singleton, as the direction (B,b) changes, the pair in $E(A,a)$ achieving the infimum of $Bx \cdot x + b \cdot x$ changes resulting in the nonlinearity of the map $(B,b) \mapsto d_H f(A,a;B,b)$. □

Remark 5.4.
The theorem holds with obvious changes for a supremum. Concavity becomes convexity and the minimum becomes a maximum in the conclusions. □

Proof. (i) For all α, $0 < \alpha < 1$, (A,a), (B,b), and $x \in E$,

$$\begin{aligned} &[\alpha A + (1-\alpha B]x \cdot x + [\alpha a + (1-\alpha b] \cdot x \\ &= \alpha [Ax \cdot x + a \cdot x] + (1-\alpha)[Bx \cdot x + b \cdot x] \\ &\ge \alpha \inf_{x \in E} [Ax \cdot x + a \cdot x] + (1-\alpha) \inf_{x \in E} [Bx \cdot x + b \cdot x] \\ \Rightarrow\ & f(\alpha(A,a) + (1-\alpha(B,b)) \ge \alpha f(A,a) + (1-\alpha) f(B,b) \end{aligned}$$

and the function is concave. For (A,a) and (B,b),

$$\begin{aligned} & Ax \cdot x + a \cdot x = (A-B)x \cdot x + (a-b)x + Bx \cdot x + b \cdot x \\ \Rightarrow\ & Ax \cdot x + a \cdot x \ge -\|A-B\| \, \|x\|^2 - \|a-b\| \, \|x\| + Bx \cdot x + b \cdot x \\ \Rightarrow\ & \inf_{x \in E} Ax \cdot x + a \cdot x \ge -c^2 \|A-B\| - c\,\|a-b\| + \inf_{x \in E} Bx \cdot x + b \cdot x \\ \Rightarrow\ & f(B,b) - f(A,a) \le c^2 \|A-B\| + c\,\|a-b\|, \end{aligned}$$

since the norms of the elements of the compact E are bounded by some constant c. By interchanging the roles of (A,a) and (B,b), we get $f(A,a) - f(B,b) \le c^2\|A-B\| + c\,\|a-b\|$ and $|f(B,b) - f(A,a)| \le c^2\|A-B\| + c\,\|a-b\|$.

(ii) We need only to show that

$$df(A,a;B,b) = \min_{x\in E(A,a)} Bx\cdot x + b\cdot x. \tag{5.20}$$

Then, from the Lipschitz continuity of $f(A,a)$ (see Theorem 3.5 in Chapter 3),

$$d_H f(A,a;B,b) = \min_{x\in E(A,a)} Bx\cdot x + b\cdot x.$$

So, choose $X = E$ and the function

$$G(t,x) \stackrel{\text{def}}{=} [A+t\,B]x\cdot x + [a+t\,b]\cdot x, \quad \partial_t G(t,x) = Bx\cdot x + b\cdot x, \tag{5.21}$$

$$g(t) = f(A+t\,B, a+t\,b) \tag{5.22}$$

and check the assumptions of Theorem 5.1. Assumption (H1) is verified by the Weierstrass theorem since the function $G(t,x)$ is continuous at x and $X = E$ is compact. Since $\partial_t G(t,x) = Bx\cdot x + b\cdot x$ is independent of t, assumptions (H2) and (H4) are verified. For any sequence $\{t_n\}$, $t_n > 0 \to 0$, pick arbitrary points $x_n \in X(t_n)$. Finally, since X is compact and $X(t_n) \subset X$, there exist a subsequence $\{x_{n_k}\}$ and a point $x_0 \in X$ such that $x_{n_k} \to x_0$ and $t_{n_k} \to 0$. In particular, (H3) (ii) is verified since $Bx_{n_k}\cdot x_{n_k} + b\cdot x_{n_k} \to Bx_0\cdot x_0 + b\cdot x_0$. It remains to show that $x_0 \in X(0)$. For each k and any $x^0 \in X(0)$,

$$\begin{aligned} f(A+t_{n_k}B, a+t_{n_k}b) &= [A+t_{n_k}B]x_{n_k}\cdot x_{n_k} + [a+t_{n_k}b]\cdot x_{n_k} \\ &\le [A+t_{n_k}B]x^0\cdot x^0 + [a+t_{n_k}b]\cdot x^0. \end{aligned}$$

By going to the limit,

$$\begin{aligned} &Ax_0\cdot x_0 + a\cdot x_0 \\ &= \lim_{k\to\infty} [A+t_{n_k}B]x_{n_k}\cdot x_{n_k} + [a+t_{n_k}b]\cdot x_{n_k} \\ &\le \lim_{k\to\infty} [A+t_{n_k}B]x^0\cdot x^0 + [a+t_{n_k}b]\cdot x^0 = Ax^0\cdot x^0 + a\cdot x^0 = f(A,a) \\ &\qquad \Rightarrow \exists x_0 \in X, \text{ such that } Ax_0\cdot x_0 + a\cdot x_0 \le f(A,a). \end{aligned}$$

As $f(A,a)$ is the infimum, $x_0 \in X(0)$ and the conclusions of Theorem 5.1 apply.

(iii) The formula follows from the chain rule for Hadamard semidifferentiable functions of Theorem 3.4. □

Theorem 5.2 assumes that the set E is compact, but its conclusions hold under other sets of assumptions. As a second example involving a quadratic function over $\mathbb{R}^n$, we shall need the following result that anticipates on Chapter 5.

Lemma 5.1. *Let A be an $n\times n$ positive definite matrix and $a \in \mathbb{R}^n$. The problem*

$$\inf_{x\in\mathbb{R}^n} q(x), \quad q(x) \stackrel{\text{def}}{=} Ax\cdot x + a\cdot x \tag{5.23}$$

has a unique minimizer $\hat{x}$ that is the solution of the equation

$$2\,A\hat{x} + a = 0. \tag{5.24}$$

Proof. From Chapter 2, q has the growth property in $\mathbb{R}^n$ from Lemma 5.1 and Example 5.8. So, it has a nonempty lower section in $\mathbb{R}^n$ by Theorem 5.4. Finally, since q is continuous as a polynomial function and $U = \mathbb{R}^n$ is closed, there exists a minimizer $\hat{x}$ by Theorem 5.3. By Taylor's theorem (Theorem 2.6) applied to the function $g(t) = q(\hat{x} + t(x - \hat{x}))$, $x \in \mathbb{R}^n$: there exists $\alpha \in]0,1[$ such that $g(1) = g(0) + g'(0) + (1/2)g''(\alpha)$. But $g'(0) = (2A\hat{x} + a)\cdot(x - \hat{x})$ and $g''(\alpha) = 2A(x - \hat{x})\cdot(x - \hat{x})$. Therefore,

$$q(x) = q(\hat{x}) + (2A\hat{x} + a)\cdot(x - \hat{x}) + A(x - \hat{x})\cdot(x - \hat{x}).$$

In particular, for all $t > 0$,

$$\begin{aligned}
&0 \le q(\hat{x} \pm t(x - \hat{x})) - q(\hat{x}) = \pm t(2A\hat{x} + a)\cdot(x - \hat{x}) + t^2 A(x - \hat{x})\cdot(x - \hat{x}) \\
&\Rightarrow 0 \le \frac{q(\hat{x} \pm t(x - \hat{x})) - q(\hat{x})}{t} = \pm(2A\hat{x} + a)\cdot(x - \hat{x}) + tA(x - \hat{x})\cdot(x - \hat{x}).
\end{aligned}$$

By letting t go to zero, $\pm(2A\hat{x} + a)\cdot(x - \hat{x}) \ge 0$ and hence $(2A\hat{x} + a)\cdot(x - \hat{x}) = 0$ for all $x \in \mathbb{R}^n$. Thus, $2A\hat{x} + a = 0$. Since A is positive definite, the solution $\hat{x} = -(1/2)A^{-1}a$ is unique. □

Theorem 5.3. *Assume that A is a symmetric $n \times n$ positive definite matrix.*

(i) *There exists $r > 0$ such that the map $(A', a') \mapsto f(A', a') : B_r(A) \times B_r(a) \to \mathbb{R}$ is well defined, concave, and Lipschitzian.*

(ii) *The function f is Hadamard differentiable at (A, a). For all $B \in \mathrm{Sym}_n$ and $b \in \mathbb{R}^n$,*

$$\begin{aligned}
d_H f(A, a; B, b) &\overset{\text{def}}{=} \lim_{\substack{t \searrow 0 \\ (W,w)\to(B,b)}} \frac{f(A + tW, a + tw) - f(A, a)}{t} \\
&= Bx_0 \cdot x_0 + b \cdot x_0 = x_0 \otimes x_0 \cdot\cdot B + b \cdot x_0,
\end{aligned} \tag{5.25}$$

where $x_0 \in \mathbb{R}^n$ is the unique solution of the equation $2Ax_0 + a = 0$.

(iii) *Let the maps $\vec{p} \mapsto A(\vec{p}) : \mathbb{R}^k \to \mathrm{Sym}_n$ and $\vec{p} \mapsto a(\vec{p}) : \mathbb{R}^k \to \mathbb{R}^n$ be Hadamard semidifferentiable at $\vec{p}$ and assume that $A(\vec{p})$ is positive definite. Consider the function*

$$\vec{p} \mapsto \ell(\vec{p}) \overset{\text{def}}{=} f(A(\vec{p}), a(\vec{p})) : \mathbb{R}^k \to \mathbb{R}. \tag{5.26}$$

Then for $\vec{v} \in \mathbb{R}^k$,

$$\begin{aligned}
d_H \ell(\vec{p}; \vec{v}) &= d_H f(A(\vec{p}), a(\vec{p}); d_H A(\vec{p}; \vec{v}), d_H a(\vec{p}; \vec{v})) \\
&= x_0 \otimes x_0 \cdot\cdot d_H A(\vec{p}; \vec{v}) + x_0 \cdot d_H a(\vec{p}; \vec{v})),
\end{aligned}$$

where $x_0 \in \mathbb{R}^n$ is the unique solution of the equation $2A(\vec{p})x_0 + a(\vec{p}) = 0$.

Proof. The proof is similar to the one of Theorem 5.2 with the following changes. Since A is positive definite, there exists $\alpha > 0$ such that for all $x \in \mathbb{R}^n$, $Ax \cdot x \ge \alpha \|x\|^2$. Therefore, for $A' \in \mathrm{Sym}_n$,

$$\begin{aligned}
&A'x \cdot x = Ax \cdot x + (A' - A)x \cdot x \\
\Rightarrow\ &A'x \cdot x \ge Ax \cdot x - \|A' - A\|\,\|x\|^2 \ge \big(\alpha - \|A' - A\|\big)\,\|x\|^2 > (\alpha/2)\,\|x\|^2
\end{aligned}$$

for all A' such that $\|A'-A\| < \alpha/2$, that is, for all A' in the ball $B_{\alpha/2}(A)$ of radius $\alpha/2$ in A. So, for all $A' \in B_{\alpha/2}(A)$ and $a' \in \mathbb{R}^n$, the infimum of the quadratic function $A'x\cdot x + a'\cdot x$ has a unique minimizer $x' \in \mathbb{R}^n$ that is the unique solution of $2\,A'x' + a = 0$. So the map $(A',a') \mapsto f(A',a') : B_{\alpha/2}(A)\times\mathbb{R}^n \to \mathbb{R}$ is well defined. Moreover, for all $a' \in B_{\alpha/2}(a)$,

$$
\begin{aligned}
&2\alpha\,\|x'\|^2 \le 2\,A'x'\cdot x' = -a'\cdot x' \;\le \|a'\|\,\|x'\| \\
\Rightarrow\;& \|x'\| \le \frac{1}{2\alpha}\|a'\| \le \frac{1}{2\alpha}[\|a\| + \|a-a'\|] < k \overset{\text{def}}{=} \frac{1}{2\alpha}\left[\|a\| + \frac{\alpha}{2}\right].
\end{aligned} \tag{5.27}
$$

Define $X=\mathbb{R}^n$ and $G(t,x) = [A+t\,B]x\cdot x + [a+t\,b]\cdot x$ for $B \in \mathrm{Sym}_n$ and t, $0 \le t < \tau$, where $\tau > 0$ is chosen such that $\tau\,\|B\| < \alpha/2$ in order to have $A+tB \in B_{\alpha/2}(A)$ for all t, $0 \le t < \tau$. As a result, $X(t) = \{x_t\}$ is a singleton unique solution of the equation $2[A+t\,B]x\cdot x + [a+t\,b]\cdot x = 0$.

(i) For all α, $0<\alpha<1$, (A_1,a_1), (A_2,a_2), and $x\in E$,

$$
\begin{aligned}
&[\alpha A_1 + (1-\alpha A_2]x\cdot x + [\alpha a_1 + (1-\alpha a_2]\cdot x \\
&\quad = \alpha\,[A_1x\cdot x + a_1\cdot x] + (1-\alpha)\,[A_2x\cdot x + a_2\cdot x] \\
&\quad \ge \alpha\,\inf_{x\in E}[A_1x\cdot x + a_1\cdot x] + (1-\alpha)\,\inf_{x\in E}[A_2x\cdot x + a_2\cdot x] \\
&\Rightarrow\; f(\alpha(A_1,a_1) + (1-\alpha(A_2,a_2)) \ge \alpha\, f(A_1,a_1) + (1-\alpha)\, f(A_2,a_2)
\end{aligned}
$$

and the function is concave.

For (B,b) and (C,c) in $B_{\alpha/2}(A)\times B_{\alpha/2}(a)$,

$$
\begin{aligned}
&Bx\cdot x + b\cdot x = (B-C)x\cdot x + (b-c)x + Cx\cdot x + c\cdot x \\
&\Rightarrow\; Bx\cdot x + b\cdot x \ge -\|B-C\|\,\|x\|^2 - \|b-c\|\,\|x\| + Cx\cdot x + c\cdot x \\
&\Rightarrow\; Bx\cdot x + b\cdot x \ge -\|B-C\|\,\|x\|^2 - \|b-c\|\,\|x\| + \inf_{x\in E} Cx\cdot x + c\cdot x.
\end{aligned}
$$

If x_B is the minimizing element for the infimum $f(B,b)$, then

$$
\begin{aligned}
f(B,b) = Bx_B\cdot x_B + b\cdot x_B &\ge -\|B-C\|\,\|x_B\|^2 - \|b-c\|\,\|x_B\| + \inf_{x\in E} Cx\cdot x + c\cdot x \\
&\Rightarrow\; f(C,c) - f(B,b) \le \|x_B\|^2\|B-C\| + \|x_B\|\,\|b-c\|.
\end{aligned}
$$

But from (5.27), $\|x_B\| < k$ and

$$
f(C,c) - f(B,b) \le k^2\|B-C\| + k\,\|b-c\|.
$$

By interchanging the roles of (B,b) and (C,c), we get $f(B,b) - f(C,c) \le k^2\|B-C\| + k\,\|c-b\|$ and $|f(B,b) - f(C,c)| \le k^2\|B-C\| + k\,\|c-b\|$.

(ii) The proof is the same as the one of Theorem 5.2 by observing that all the $X(t)$'s, $0 \le t < \tau$, are contained in the compact ball $\overline{B_k(0)}$. In addition, $d_H f(A,a;B,b)$ is linear in (B,b) since $X(0)$ is a singleton. So f is Hadamard and Fréchet differentiable in (A,a).

(iii) The formula follows from the chain rule for Hadamard semidifferentiable functions of Theorem 3.4. □

6 Summary of Semidifferentiability and Differentiability

This section summarizes the various semidifferentials and differentials introduced in this chapter. The basic notions are the semidifferential $df(x;v)$ and the stronger Hadamard semidifferential $d_H f(x;v)$ which plays a central role since the rules of the classical differential calculus remain valid for that family of functions. With the linearity with respect to v, we have the Gateaux and Hadamard differentials, respectively, and the notion of gradient.

$$df(x;v) \overset{\text{def}}{=} \lim_{t\searrow 0} \frac{f(x+tv)-f(x)}{t}$$

f semidifferentiable at x in the direction v

$\forall v \in \mathbb{R}^n$, $df(x;v)$ exists and $v \mapsto df(x;v) : \mathbb{R}^n \to \mathbb{R}^n$ is linear

f Gateaux differentiable at x

$$d_H f(x;v) \overset{\text{def}}{=} \lim_{\substack{t\searrow 0\\ w\to v}} \frac{f(x+tw)-f(x)}{t}$$

f Hadamard semidifferentiable at x in the direction v

$\forall v \in \mathbb{R}^n$, $d_H f(x;v)$ exists and $v \mapsto d_H f(x;v) : \mathbb{R}^n \to \mathbb{R}^n$ is linear

f Hadamard differentiable at x

Hadamard semidifferentiability has an equivalent geometric characterization in terms of paths.

f Hadamard semidifferentiable at x in the direction v $\iff$ $\exists\, g(x,v) \in \mathbb{R}$ such that $\forall h : [0,\infty[\to \mathbb{R}^n$, $h(0) = x$ and $dh(0;+1) = v$, $d(f\circ h)(0;+1)$ exists and $d(f\circ h)(0;+1) = g(x,v)$

Hadamard differentiability coincides with the Fréchet differentiability and also has an equivalent geometric characterization in terms of paths.

f Hadamard differentiable at x

$\iff$ f Fréchet differentiable at x : $\exists\, L(x) : \mathbb{R}^n \to \mathbb{R}$ linear such that $\displaystyle\lim_{v\to 0} \frac{f(x+v)-f(x)-L(x)v}{\|v\|} = 0$

$\iff$ $\exists\, L(x) : \mathbb{R}^n \to \mathbb{R}$ linear such that $\forall h : \mathbb{R} \to \mathbb{R}^n$, $h(0) = x$ and $h'(0)$ exists, $(f\circ h)'(0)$ exists and $(f\circ h)'(0) = L(x)\,h'(0)$

The above notions are sufficient for most problems of practical interest.

In going to arbitrary functions $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $\operatorname{dom} f \neq \varnothing$, each of the two semidifferentials splits into two by replacing the limit by the liminf and the limsup which can both take the value $+\infty$ or $-\infty$.

$$\underline{d} f(x;v) \overset{\text{def}}{=} \liminf_{t \searrow 0} \frac{f(x+tv)-f(x)}{t}$$
Dini lower semidifferential
at $x \in \text{dom}\, f$ in the direction v

$$\underline{d}_H f(x;v) \overset{\text{def}}{=} \liminf_{\substack{t \searrow 0 \\ w \to v}} \frac{f(x+tw)-f(x)}{t}$$
Hadamard lower semidifferential
at $x \in \text{dom}\, f$ in the direction v

$$\overline{d} f(x;v) \overset{\text{def}}{=} \limsup_{t \searrow 0} \frac{f(x+tv)-f(x)}{t}$$
Dini upper semidifferential
at $x \in \text{dom}\, f$ in the direction v

$$\overline{d}_H f(x;v) \overset{\text{def}}{=} \limsup_{\substack{t \searrow 0 \\ w \to v}} \frac{f(x+tw)-f(x)}{t}$$
Hadamard upper semidifferential
at $x \in \text{dom}\, f$ in the direction v

$$\boxed{\underline{d}_H f(x;v) \le \underline{d} f(x;v) \le \overline{d} f(x;v) \le \overline{d}_H f(x;v).}$$

The next two notions necessitate that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite not only at the point $x \in \text{dom}\, f$ but also in a neighborhood $V(x)$ of f in order to make sense of terms of the form $f(y+tv) - f(y)$, that is, $x \in \text{int}\,(\text{dom}\, f)$.

$$\underline{d}_C f(x;v) \overset{\text{def}}{=} \liminf_{\substack{t \searrow 0 \\ y \to x}} \frac{f(y+tv)-f(y)}{t}$$
Clarke lower semidifferential
at $x \in \text{dom}\, f$ in the direction v

$$\overline{d}_C f(x;v) \overset{\text{def}}{=} \limsup_{\substack{t \searrow 0 \\ y \to x}} \frac{f(y+tv)-f(y)}{t}$$
Clarke upper semidifferential
at $x \in \text{dom}\, f$ in the direction v

When f is Lipschitzian at $x \in \text{dom}\, f$, the six lower/upper semidifferentials are finite and

$$\boxed{\underline{d}_C f(x;v) \le \underline{d}_H f(x;v) \le \underline{d} f(x;v) \le \overline{d} f(x;v) \le \overline{d}_H f(x;v) \le \overline{d}_C f(x;v).}$$

To finish, we have also included the intermediary notion of *directional derivative*:

$$\lim_{t \to 0} \frac{f(x+tv)-f(x)}{t}$$
f has a derivative in the direction v at x

which is of limited interest since it hides the more fundamental notion of semidifferential and is not yet the Gateaux differential.

7 Exercises

Exercise 7.1.
Prove that the function (see Figure 3.10)

$$f(x,y) \overset{\text{def}}{=} \begin{cases} \dfrac{xy^2}{x^2+y^4}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

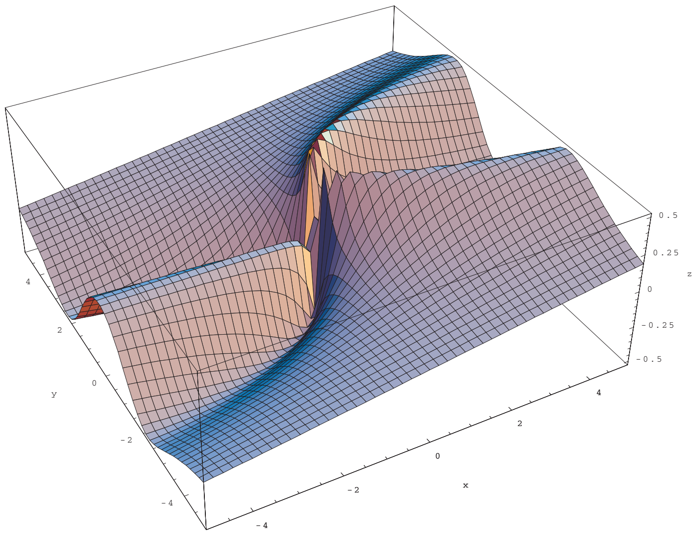


Figure 3.10. *Function of Exercise* 7.1.

is directionally differentiable at $(x,y)=(0,0)$, but is neither Gateaux differentiable nor continuous at $(x,y)=(0,0)$. Note the following properties:

$$x<0 \quad \Rightarrow f(x,y)\le 0 \quad \text{and} \quad x>0 \quad \Rightarrow f(x,y)\ge 0$$
$$f(-x,y)=-f(x,y) \quad \text{and} \quad f(x,-y)=f(x,y). \qquad \square$$

Exercise 7.2.
Let $A:\mathbb{R}^n\to\mathbb{R}^n$ be linear (or an $n\times n$ matrix) and $b\in\mathbb{R}^n$ (or an n-vector). Define

$$f(x)\overset{\text{def}}{=}\frac{1}{2}(Ax)\cdot x+b\cdot x, \quad x\in\mathbb{R}^n.$$

(i) Compute $df(x;v)$ (or the gradient $\nabla f(x)$) and $d^2f(x;v;w)$ (or the Hessian $Hf(x)$).

(ii) Give necessary and sufficient conditions on A,b for the convexity of f.

(iii) Give necessary and sufficient conditions on A,b for the strict convexity of f.

(iv) Are the functions f associated with the following matrices and vectors convex?

$$\text{(a)} \quad A=\begin{bmatrix} 3 & 1\\ -1 & 2\end{bmatrix}, b=\begin{bmatrix} -2\\ 1\end{bmatrix}, \quad \text{and} \quad \text{(b)} \quad A=\begin{bmatrix} 2 & 4\\ 4 & 1\end{bmatrix}, b=\begin{bmatrix} 1\\ 1\end{bmatrix}. \qquad \square$$

Exercise 7.3.
Let $f(x)=|x|^n$, $x\in\mathbb{R}$, $n\ge 1$ an integer.

(i) Determine the values of n for which f is differentiable on $\mathbb{R}$.

(ii) Give the directional derivatives of f at $x = 0$ as a function of $n \geq 1$, if it exists. Otherwise, give the semidifferentials.

(iii) Determine the values of $n \geq 1$ for which f is convex on $\mathbb{R}$. □

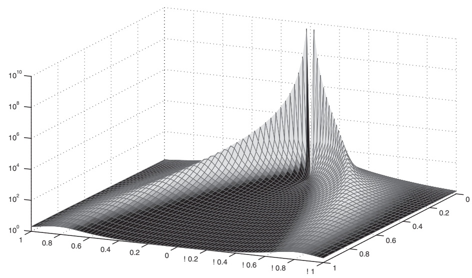
Exercise 7.4.
Given an integer $n \geq 1$, define the function $f : \mathbb{R}^n \to \mathbb{R}$:

$$f(x) \stackrel{\text{def}}{=} \sum_{i=1}^{n} |x_i|. \tag{7.1}$$

(i) Prove that the function f is convex and Lipschitzian on $\mathbb{R}^n$ and give the expression of its semidifferential $df(x;v)$.

(ii) For $n = 2$, compute $df(x;v)$ at the points $x = (1,1)$, $(0,1)$, $(1,0)$, $(0,0)$ as a function of $v = (v_1, v_2)$. In which cases is f Gateaux differentiable? □

Exercise 7.5.
Show that the function $f(x) = \sin x + (1+x)^2$ is convex on the interval $[0,1]$. □



Chapter 4
Optimality Conditions

1 Introduction

This chapter presents several general optimality conditions to characterize the local and global minima of semidifferentiable or differentiable objective functions f with respect to a subset U of $\mathbb{R}^n$. We first consider the unconstrained differentiable case with several examples including the least and greatest eigenvalues of a symmetric matrix. This last example is used to illustrate the general theorem on the semidifferentiability of an infimum with respect to parameters. We then separately consider the cases of a convex U and of an arbitrary U. We give necessary and sufficient optimality conditions when U and f are both convex. When U is arbitrary, we need to introduce the *tangent cone of admissible directions* to locally approximate U and we get a necessary optimality condition. A few technical results on cones and dual cones will be needed. The characterization of the dual cones for sets U that are specified by a finite number of affine functional constraints will be given. They will be used in Chapter 5 to deal with equality constraints (Lagrange multipliers theorem), inequality constraints (Karush–Kuhn–Tucker theorem), and the mixed case.

The chapter is structured as follows. Section 2 is devoted to the unconstrained case of an objective function of class $C^{(2)}$ with several examples. Special attention is given to the generic example of the least and greatest eigenvalues of a symmetric matrix. The explicit expression of their semidifferential with respect to parameters is given using Theorem 5.2 from section 5 in Chapter 3 when the eigenvalue is not simple.

For an abstract constraint set U, we proceed in two steps while relaxing differentiability of the objective function to one of the forms of semidifferentiability introduced in Chapter 3. We first consider the case of a convex U in section 3. Section 3.2 gives a necessary and sufficient condition for a convex Gateaux differentiable objective function and section 3.3 gives conditions for a semidifferentiable objective function. Finally, section 3.4 gives a necessary and sufficient condition for an arbitrary convex function using the lower semidifferential.

In preparation for the case of an arbitrary U, section 4 introduces the notions of *admissible directions* that are half-tangents to U that generate *tangent cones* to U. The properties of those cones are studied in sections 4.2 and 4.3. Section 5 introduces the notions of orthogonality, transposition, and dual cones.

Section 6 gives a necessary optimality condition in terms of the upper Hadamard semidifferential of the objective function for an arbitrary set U. It can also be expressed in the form of an inclusion of the tangent cone of admissible directions into a cone associated with the objective function. This optimality condition leads to the *dual necessary optimality condition* on the duals of those cones in section 6.2. It will be used in Chapter 5 in the differentiable case as well as in the convex case.

Section 7 specializes to sets specified by linear equality or inequality constraints for which a complete characterizations of the tangent cone and its dual are available. The dual cone is given in terms of the *Lagrange multipliers*. This is applied to the *linear programming* problem and to a generalization of Farkas's lemma. At this juncture, the *Lagrangian* is introduced along with its connection with two-person zero-sum games and Fenchel's primal and dual problems of Chapter 2. The constructions and results are extended to the *quadratic programming* problem in section 7.6 and to Fréchet differentiable objective functions in section 7.7. Both convex and nonconvex quadratic objective functions are considered.

In closing this chapter, section 8 very briefly outlines another point of view: the characterization of minimizers via *subdifferentials* that are not single-valued but *set-valued functions*.

2 Unconstrained Differentiable Optimization

By *unconstrained* we mean the fact that the minimizers occur at interior points of U (that is, the "constraints" specified by U are not *active*). In such cases, only the local properties of the function and its derivatives are involved and we recover *Fermat's rule* that says that the gradient of f is zero at the points of U corresponding to local infima. The general case where minimizers can occur on the boundary of U (that is, the "constraints" specified by U are *active*) will be studied later in this chapter. The necessary optimality condition will be specialized to sets U that are characterized by a finite number of equality and inequality constraints in Chapter 5. The sign of the Hessian matrix will complete the characterization of a local extremum by discriminating between a minimum, a maximum, or something else.

2.1 Some Basic Results and Examples

This section gives the main characterizations of local and global minimizers using the gradient and the Hessian matrix. Consider the infimum $\inf f(U)$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ over an open subset U of $\mathbb{R}^n$. The result will apply to the supremum $\sup f(U)$ by replacing f by $-f$ and considering the infimum $\inf -f(U)$.

Definition 2.1.
Let $U, \varnothing \neq U \subset \mathbb{R}^n$.

(i) $f : \mathbb{R}^n \to \mathbb{R}$ has a *global minimum* on U if

$$\exists x \in U \text{ such that } \forall y \in U, \quad f(x) \le f(y). \tag{2.1}$$

(ii) $f : \mathbb{R}^n \to \mathbb{R}$ has a *local minimum* on U if there exists $x \in U$ and a neighborhood $V(x)$ of x such that

$$\forall y \in U \cap V(x), \quad f(x) \le f(y), \tag{2.2}$$

that is, x is a global minimum on $U \cap V(x)$. □

The following theorem gives conditions for the existence of a local minimum. They become necessary and sufficient for quadratic functions.

Theorem 2.1. *Let U, $\emptyset \neq U \subset \mathbb{R}^n$, be open and $f : \mathbb{R}^n \to \mathbb{R}$ be $C^{(2)}$ on U.*

(i) *If f has a local minimum on U, then*

$$\exists x \in U, \quad \nabla f(x) = 0 \quad \text{and} \quad Hf(x) \ge 0. \tag{2.3}$$

(ii) *If*

$$\exists x \in U \quad \text{such that} \quad \nabla f(x) = 0$$

and if there exists a convex neighborhood $V(x) \subset U$ of x such that

$$\forall y \in V(x), \quad Hf(y) \ge 0, \tag{2.4}$$

then f has a minimum at x which is global on $V(x)$ and local on U.

(iii) *If*

$$\exists x \in U, \quad \nabla f(x) = 0 \quad \text{and} \quad Hf(x) > 0, \tag{2.5}$$

then x is a local minimizer of f on U and there exists a neighborhood $V(x)$ of x where x is the unique global minimum of f.

Proof. (i) If x is a local minimizer of f on U, then there exists $r > 0$ such that

$$B_r(x) \subset U \quad \text{and} \quad \forall y \in B_r(x), \quad f(y) \ge f(x).$$

For any $y \in B_r(x)$ and $t \in]0,1[$, $x + t(y-x) \in B_r(x)$ and

$$f(x + t(y-x)) \ge f(x). \tag{2.6}$$

Consider the new function $g(t) = f(x + t(y-x))$. For all $y \in B_r(x)$, $|t| < 1$ and $|s| < 1$, the points $x + s(y-x)$ and $x + t(y-x)$ belong to $B_r(x)$ and for $s \neq t$,

$$\begin{aligned}\frac{g(s)-g(t)}{s-t} &= \frac{f(x+s(y-x)) - f(x+t(y-x))}{s-t} \\ &= \frac{f(x+t(y-x)+(s-t)(y-x)) - f(x+t(y-x))}{s-t}.\end{aligned}$$

Since f is of class $C^{(1)}$ on U, the semidifferential of f at $x+t(y-x)$ in the direction $y-x$ exists and

$$\begin{aligned}g'(t) &= \lim_{\substack{(s-t)\to 0\\ s\neq t}} \frac{g(s)-g(t)}{s-t} \\ &= \lim_{\substack{(s-t)\to 0\\ s\neq t}} \frac{f(x+t(y-x)+(s-t)(y-x)) - f(x+t(y-x))}{s-t} \\ &= df(x+t(y-x); y-x) = \nabla f(x+t(y-x)) \cdot (y-x).\end{aligned}$$

Since f is of class $C^{(2)}$, we get by the same process

$$g^{(2)}(t) = d^2 f(x+t(y-x); y-x; y-x) = Hf(x+t(y-x))(y-x)\cdot(y-x).$$

Therefore, from (2.6) for $t > 0$,

$$\begin{aligned}\frac{g(t)-g(0)}{t} &= \frac{f(x+t(y-x))-f(x)}{t} \geq 0 \\ \Rightarrow g'(0) = \lim_{t\searrow 0}\frac{g(t)-g(0)}{t} \geq 0 &\quad \Rightarrow df(x;y-x) = \nabla f(x)\cdot(y-x) \geq 0\end{aligned}$$

and hence

$$\forall y \in B_r(x), \quad \nabla f(x)\cdot(y-x) \geq 0.$$

For all $v \in \mathbb{R}^n$, $v \neq 0$, and

$$y = x \pm \frac{r}{2}\frac{v}{\|v\|} \in B_r(x) \quad \Rightarrow \pm\frac{r}{2\|v\|}\nabla f(x)\cdot v \geq 0 \quad \Rightarrow \forall v \in \mathbb{R}^n, \nabla f(x)\cdot v = 0$$

and hence $\nabla f(x) = 0$. By Taylor's theorem (Theorem 2.6), for all $y \in B_r(x)$, there exists $\alpha \in]0,1[$ such that

$$g(1) = g(0) + g'(0) + \frac{1}{2}g^{(2)}(\alpha).$$

Using the expression

$$g^{(2)}(\alpha) = d^2 f(x+\alpha(y-x); y-x; y-x) = Hf(x+\alpha(y-x))(y-x)\cdot(y-x),$$

we get

$$f(y) = f(x) + \nabla f(x)\cdot(y-x) + \frac{1}{2}Hf(x+\alpha(y-x))(y-x)\cdot(y-x).$$

Using the inequality $f(y) \geq f(x)$ on $B_r(x)$ and $\nabla f(x) = 0$, we get

$$\forall y \in B_r(x), \quad Hf\,(x+\alpha(y-x))\,(y-x)\cdot(y-x) \geq 0.$$

For all $v \neq 0$ in $\mathbb{R}^n$, substitute $y = x + r\,v/(2\|v\|)$ in the above inequality to get

$$\begin{aligned}\forall 0 \neq v \in \mathbb{R}^n, &\quad Hf\left(x+\alpha\frac{r}{2}\frac{v}{|v|}\right)\frac{v}{|v|}\cdot\frac{v}{|v|} \geq 0 \\ \Rightarrow \forall v \in \mathbb{R}^n, &\quad Hf\left(x+\alpha\frac{r}{2}\frac{v}{|v|}\right)v\cdot v \geq 0.\end{aligned}$$

By continuity of Hf at x, letting r go to 0, we get, for all $v \in \mathbb{R}^n$, $Hf(x)v\cdot v \geq 0$ and the Hessian matrix $Hf(x)$ is positive semidefinite.

(ii) It is sufficient to use Taylor's theorem (Theorem 2.6) as in part (i). For all $y \in V(x)$, there exists $\alpha \in]0,1[$ such that

$$f(y) = f(x) + \nabla f(x)\cdot(y-x) + \frac{1}{2}Hf(x+\alpha(y-x))(y-x)\cdot(y-x).$$

But $\nabla f(x) = 0$ and since $V(x)$ is convex, $x + \alpha(y-x) \in V(x)$ and, by assumption, $Hf(x + \alpha(y-x)) \geq 0$ is positive semidefinite. This yields

$$f(y) \geq f(x), \quad \forall y \in V(x) \subset U.$$

The point $x \in U$ is a global minimizer of f on $V(x)$ and a local one on U.

(iii) Indeed, if $Hf(x) > 0$, then, since f is of class $C^{(2)}$, the Hessian matrix Hf is continuous and there exists an open ball $B_r(x) \subset U$ such that

$$\forall y \in B_r(x), \quad Hf(y) > 0.$$

Since $\nabla f(x) = 0$ and condition (2.4) is verified on $B_r(x)$, x is a global minimum on $B_r(x)$. Moreover, for all $y \in B_r(x)$, $y \neq x$, there exists $\alpha \in]0,1[$ such that

$$\begin{aligned} &f(y) = f(x) + \nabla f(x) \cdot (y-x) + \frac{1}{2} Hf(x + \alpha(y-x))(y-x) \cdot (y-x) \\ &\Longrightarrow \forall y \in B_r(x) \text{ such that } y \neq x, \quad f(y) > f(x). \end{aligned}$$

This proves uniqueness in $B_r(x)$. □

The conditions of Theorem 2.1 are necessary and sufficient for quadratic functions.

Theorem 2.2. *Consider the quadratic function*

$$f(x) \overset{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x + c \tag{2.7}$$

for an $n \times n$ symmetrical matrix A, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

(i) *f has a minimizer in $\mathbb{R}^n$ if and only if*

$$\exists x \in \mathbb{R}^n \text{ such that } \nabla f(x) = Ax + b = 0 \text{ and } A \geq 0. \tag{2.8}$$

(ii) *There exists a unique minimizer of f in $\mathbb{R}^n$ if and only if $A > 0$.*

Proof. (i) The condition is necessary from Theorem 2.1(i) and sufficient from Theorem 2.1(ii) since $Hf(y) = A = Hf(x) \geq 0$ in the neighborhood $V(x) = \mathbb{R}^n$.

(ii) We know form part (i) that there exists a minimizer $x \in \mathbb{R}^n$ if and only if $Ax + b = 0$ and $A \geq 0$. If $A > 0$, then $A \geq 0$, A is invertible, and for all b the equation $Ax + b = 0$ has a unique solution $x = -A^{-1}b$. Conversely, by Taylor's theorem, using the fact that $Hf(x) = A$, we have the exact expansion

$$f(y) - f(x) = \nabla f(x) \cdot (y-x) + \frac{1}{2} A(y-x) \cdot (y-x).$$

If x is the unique minimizer, then $\nabla f(x) = 0$ and

$$\begin{aligned} \forall y \neq x, \quad 0 < f(y) - f(x) &= \nabla f(x) \cdot (y-x) + \frac{1}{2} A(y-x) \cdot (y-x) \\ &= \frac{1}{2} A(y-x) \cdot (y-x). \end{aligned}$$

By choosing $y = x + z$ for any $z \neq 0$, we get

$$\forall y \neq 0, \quad 0 < f(x+z) - f(x) = \frac{1}{2} Az \cdot z \quad \Rightarrow A > 0.$$ □

Example 2.1.
Consider the function $x \mapsto f(x) = x_1^2 + x_2^2 - 2x_1x_2 : \mathbb{R}^2 \to \mathbb{R}$. It is continuous on the closed set $U = \mathbb{R}^n$, but it does not have a bounded lower section. So, to prove existence we must proceed directly. It is readily seen that $f(x) = (x_1 - x_2)^2 \geq 0$ and that 0 is the greatest lower bound of f on $\mathbb{R}^2$ that is achieved on the set of points

$$M(f) = \{(x_1, x_2) : x_1 = x_2\}.$$

We now recover that result directly form part (iii) of the previous theorem. f is $C^{(2)}(\mathbb{R}^2)$ since f is a polynomial function. Its gradient and Hessian matrix are

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x)\right) = (2x_1 - 2x_2, 2x_2 - 2x_1)$$

$$Hf(x) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(x) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dfrac{\partial^2 f}{\partial x_2^2}(x) \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}.$$

The gradient is zero at the points

$$2x_1 - 2x_2 = 0 \text{ and } 2x_2 - 2x_1 = 0 \quad \Rightarrow x_1 = x_2.$$

The Hessian matrix is constant and positive semidefinite:

$$\left(\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 2(v_1 - v_2)v_1 - 2(v_1 - v_2)v_2 = 2(v_1 - v_2)^2 \geq 0.$$

Therefore, $Hf(x) \geq 0$ for all $x \in \mathbb{R}^2$. By Theorem 2.1(ii) with $V(x) = \mathbb{R}^2$, each point of $M(f)$ is a global minimizer of f on $\mathbb{R}^2$. This result can also be obtained directly from Theorem 2.2(i). □

Example 2.2.
Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x_1, x_2) \stackrel{\text{def}}{=} 16x_1^2 + (x_2 - 4)^2.$$

It is $C^{(2)}(\mathbb{R}^2)$ since it is polynomial and $U = \mathbb{R}^2$ is closed. It also has a bounded lower section. Choosing $f(0,3) = 1$,

$$(0,3) \in F_1 \stackrel{\text{def}}{=} \left\{(x_1, x_2) \in \mathbb{R}^2 : 16x_1^2 + (x_2 - 4)^2 \leq 1\right\} \neq \varnothing$$
$$\Rightarrow 4|x_1| \leq 1 \text{ and } |x_2 - 4| \leq 1 \quad \Rightarrow |x_1| \leq 1/4 \text{ and } |x_2| \leq 5.$$

So there are global minimizers on $\mathbb{R}^2$. Compute

$$\nabla f(x) = (32x_1, 2(x_2 - 4)) \quad \text{and} \quad Hf(x) = \begin{bmatrix} 32 & 0 \\ 0 & 2 \end{bmatrix}.$$

The Hessian matrix is constant positive definite. The gradient is zero for $x_1 = 0$ and $x_2 = 4$. So the global minimizer $x = (0,4)$ is unique by Theorem 2.2(ii). □

Example 2.3 (Rosenbrock Function [H. H. ROSENBROCK [**1**] function])**.**
Consider the nonconvex function (see Figure 4.1)

$$(x_1, x_2) \mapsto f(x_1, x_2) \stackrel{\text{def}}{=} 100(x_2 - x_1^2)^2 + (1 - x_1)^2 : \mathbb{R}^2 \to \mathbb{R}.$$

The interest in this example is the occurrence of a steep narrow canyon along the parabola $x_2 = x_1^2$ that creates a challenge for numerical methods.

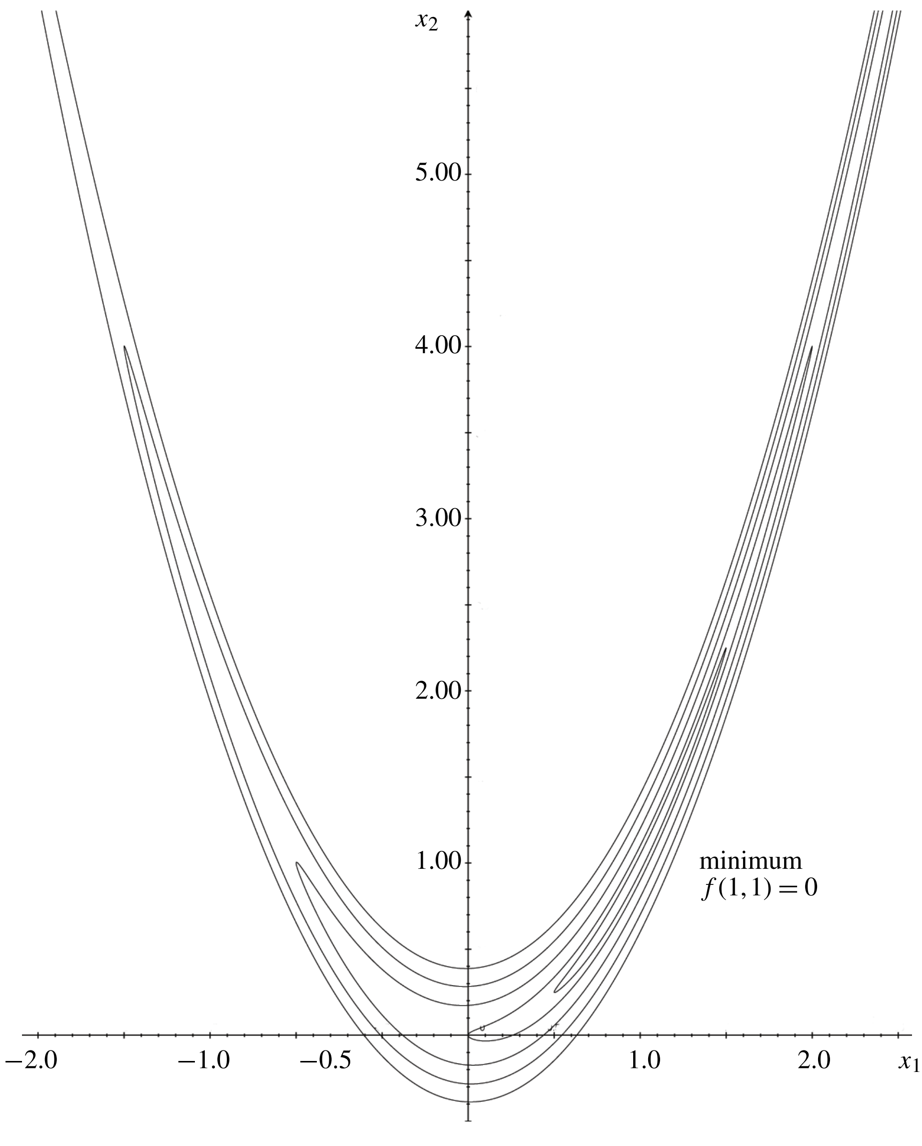


Figure 4.1. *Banana-shaped level sets of the Rosenbrock function at* 0.25*,* 1*,* 4*,* 9*,* 16*.*

It is of class $C^{(2)}$ on $\mathbb{R}^2$ since it is polynomial. It also has a bounded lower section. To see this consider the lower section

$$F_1 \stackrel{\text{def}}{=} \left\{(x_1,x_2) \in \mathbb{R}^2 : 100(x_2 - x_1^2)^2 + (1-x_1)^2 \le 1\right\}.$$

$F_1 \ne \varnothing$ since $(0,0) \in F_1$ and for any $(x_1,x_2) \in F_1$,

$$|1-x_1| \le 1 \text{ and } 10\,|x_2 - x_1^2| \le 1 \quad \Rightarrow\ |x_1| \le 2 \text{ and } 10\,|x_2| \le 41.$$

By Theorem 5.3(i) of Chapter 2, there are global minimizers on $\mathbb{R}^2$. To find them, compute $\nabla f(x)$ and $Hf(x)$:

$$\nabla f(x) = \begin{bmatrix} 200(x_2 - x_1^2)(-2x_1) - 2(1-x_1) \\ 200(x_2 - x_1^2) \end{bmatrix}$$
$$Hf(x) = \begin{bmatrix} -400x_2 + 1200x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}.$$

The zeros of the gradient are solution of the system

$$200(x_2 - x_1^2)(-2x_1) - 2(1-x_1) = 0$$
$$200(x_2 - x_1^2) = 0.$$

From the second identity $x_2 = x_1^2$ and from the first one (by using $x_2 = x_1^2$), $x_1 = 1$ and we get only the point $(x_1,x_2) = (1,1)$. Compute Hf on $(1,1)$,

$$Hf(1,1) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix},$$

and check the positive definiteness of $Hf(1,1)$. Indeed

$$\begin{aligned} Hf(1,1)v\cdot v &= 802v_1^2 - 800v_1v_2 + 200v_2^2 = 200\left[v_2^2 - 4v_1v_2 + \frac{802}{200}v_1^2\right] \\ &= 200[(v_2 - 2v_1)^2 - 4v_1^2 + 4.01v_1^2] = 200[(v_2-2v_1)^2 + 0.01v_1^2] \ge 0. \end{aligned}$$

Therefore, $Hf(1,1) \ge 0$. To show that $Hf(1,1) > 0$, it remains to show that

$$Hf(1,1)v\cdot v = 0 \quad \Rightarrow v = 0.$$

Here, $Hf(1,1)v\cdot v = 0$ implies

$$(v_2 - 2v_1)^2 = 0 \text{ and } v_1^2 = 0 \quad \Rightarrow v_1 = 0 \text{ and } v_2 = 2v_1 \quad \Rightarrow v = 0.$$

Finally, $Hf(1,1) > 0$ and $(1,1)$ is the unique local minimizer of f. Since we have already established the existence of global minimizers of f and that the point $(1,1)$ is the unique point satisfying the conditions $\nabla f(1,1) = 0$ and $Hf(1,1) \ge 0$, it is the unique global minimizer of f on $\mathbb{R}^2$. $\square$

Example 2.4 (Piecewise polynomial approximation of functions)**.**
This is a simple example of a large system where conditions (2.3) of Theorem 2.1 are difficult to check. Let $0 = x_0 < x_1 < \cdots < x_i < x_{i+1} < \cdots < x_n = 1$ be a partition of the interval $[0,1]$. For simplicity, consider the uniform partition $x_i = x_0 + i/n$, $0 \le i \le n$. Denote by $I_i = [x_{i-1}, x_i]$ the subintervals created by this partition. Denote by $P^1(I_i)$ the set of all polynomials of order less than or equal to one on I_i, $1 \le i \le n$, and by

$$P_n^1(0,1) \stackrel{\text{def}}{=} \{f \in C^0([0,1]) \, : \, f(x)|_{I_i} \in P^1(I_i), i = 1, \dots, n\}$$

the space of continuous functions on $[0,1]$ that are $P^1(I_i)$ on each subinterval of the partition $\{x_i\}$ of $[0,1]$.

Consider the following approximation problem: find $p \in P_n^1(0,1)$ that minimizes

$$\frac{1}{2} \int_0^1 |p(x) - g(x)|^2 dx$$

for a large n (for instance $n > 1000$) and a function $g \in C^0([0,1])$.

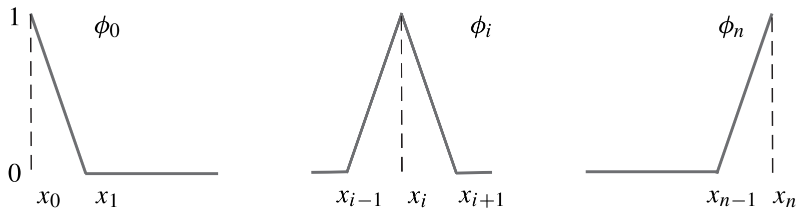


Figure 4.2. *Basis functions* $(\phi_0, \dots, \phi_i, \dots, \phi_n)$ *of* $P_n^1(0,1)$.

Associate with $P_n^1(0,1)$ the basis functions $\{\phi_i\}$ associated with each point x_i, $0 \le i \le n$, of the partition $\{I_i\}$ of the interval $[0,1]$ (see Figure 4.2):

$$\phi_0(x) \stackrel{\text{def}}{=} \begin{cases} 1 - n(x - x_0), & \text{if } x \in [x_0, x_1] \\ 0, & \text{otherwise} \end{cases}$$

$$\phi_i(x) \stackrel{\text{def}}{=} \begin{cases} 1 + n(x - x_i), & \text{if } x \in [x_{i-1}, x_i] \\ 1 - n(x - x_i), & \text{if } x \in [x_i, x_{i+1}]\,, \quad 1 < i < n, \\ 0, & \text{otherwise} \end{cases}$$

$$\phi_n(x) \stackrel{\text{def}}{=} \begin{cases} 1 + n(x - x_n), & \text{if } x \in [x_{n-1}, x_n] \\ 0, & \text{otherwise.} \end{cases}$$

At any point $x \in [0,1]$,

$$\sum_{i=0}^{n} \phi_i(x) = 1$$

and each element $p \in P_n^1(0,1)$ can be expressed in the form

$$p(x) = \sum_{i=0}^{n} p(x_i)\,\phi_i(x);$$

that is, all elements of $P_n^1(0,1)$ are of the general form

$$p(x) = \sum_{i=0}^{n} p_i\,\phi_i(x)$$

for vectors $\vec{p} = (p_0, \dots, p_n) \in \mathbb{R}^{n+1}$. There is a bijection

$$\vec{p} = (p_0, \dots, p_n) \in \mathbb{R}^{n+1} \mapsto p(x) = \sum_{i=0}^{n} p_i\,\phi_i(x) : \mathbb{R}^{n+1} \to P_n^1(0,1)$$
$$p \mapsto (p(x_0), \dots, p(x_n)) : P_n^1(0,1) \to \mathbb{R}^{n+1}.$$

Therefore, the objective function

$$J(\vec{p}) \stackrel{\text{def}}{=} \frac{1}{2}\int_0^1 \left(\sum_{i=0}^{n} p_i\,\phi_i(x) - g(x)\right)^2 dx$$

is defined and continuous on $\mathbb{R}^{n+1}$. Then for each j,

$$\frac{\partial J}{\partial p_j}(\vec{p}) = \int_0^1 \left(\sum_{i=0}^{n} p_i\,\phi_i(x) - g(x)\right)\phi_j(x)\,dx.$$

It is easy to check that

$$\nabla J(\vec{p}) = A\vec{p} - \vec{g} \text{ and } H(\vec{p}) = A,$$

where the matrix A is of dimension $n+1$ and tridiagonal (that is, the elements of A are nonzero on the diagonal and on the two adjacent diagonals):

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & 0 & 0 & 0 & 0 \\ a_{1,0} & a_{1,1} & a_{1,2} & 0 & \ddots & 0 \\ 0 & a_{2,1} & a_{2,2} & a_{2,3} & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \dots & 0 & a_{n,n-1} & a_{n,n} \end{bmatrix}. \tag{2.9}$$

The nonzero elements are given by the following expressions:

$$a_{0,0} = \int_{x_0}^{x_1} |\phi_0(x)|^2 \, dx, \quad a_{0,1} = \int_{x_0}^{x_1} \phi_0(x)\phi_1(x)\,dx,$$

$$\left.\begin{aligned} a_{i,i-1} &= \int_{x_{i-1}}^{x_i} \phi_{i-1}(x)\phi_i(x)\,dx \\ a_{i,i} &= \int_{x_{i-1}}^{x_{i+1}} |\phi_i(x)|^2 \, dx, \\ a_{i,i+1} &= \int_{x_i}^{x_{i+i}} \phi_{i+1}(x)\phi_i(x)\,dx \end{aligned}\right| \quad 1 < i < n,$$

$$a_{n,n-1} = \int_{x_{n-1}}^{x_n} \phi_{n-1}(x)\phi_n(x)\,dx, \quad a_{n,n} = \int_{x_{n-1}}^{x_n} |\phi_n(x)|^2 \, dx.$$

The components of the vector $\vec{g}$ are

$$g_0 = \int_{x_0}^{x_1} g(x)\phi_0(x)\,dx, \quad g_n = \int_{x_{n-1}}^{x_n} g(x)\phi_n(x)\,dx,$$
$$g_i = \int_{x_{i-1}}^{x_{i+1}} g(x)\phi_i(x)\,dx, \quad 0 < i < n.$$

The proof of the positive definiteness of the large matrix (2.9) is left to the reader as an exercise (see Exercise 9.1). Therefore, the equation $A\vec{p} = \vec{g}$ has a unique solution. □

Example 2.5 (Pseudoinverse)**.**
Given a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ (or an $m \times n$ matrix) and a point $c \in \mathbb{R}^m$, the equation

$$Ax = c \tag{2.10}$$

has a solution $\hat{x} \in \mathbb{R}^n$ if and only if $c \in \operatorname{Im} A$. Now consider the minimization problem

$$\inf_{x \in \mathbb{R}^n} f(x), \quad f(x) \stackrel{\text{def}}{=} \|Ax - c\|^2_{\mathbb{R}^m}. \tag{2.11}$$

It is quadratic and, by Theorem 2.2(i), there exists a solution $\hat{x}$ if and only if

$$\nabla f(\hat{x}) = 2\,A^\top(A\hat{x} - c) = 0 \quad \text{and} \quad A^\top A \geq 0. \tag{2.12}$$

The second condition is always verified since

$$\forall v \in \mathbb{R}^n, \quad A^\top A v \cdot v = \|Av\|^2_{\mathbb{R}^m} \geq 0$$

and the first one is verified if and only if

$$\exists \hat{x} \text{ such that } (A^\top A)\hat{x} = A^\top c.$$

As $\operatorname{Im}(A^\top A) = \operatorname{Im} A^\top$, this equation always has a solution. In general, a solution $\hat{x}$ is not unique, but $A\hat{x}$ is unique. If A is injective, the $n \times n$ matrix $A^\top A$ is invertible and there is a unique solution $\hat{x} = (A^\top A)^{-1}A^\top c$. The linear map $A^+ = (A^\top A)^{-1}A^\top : \mathbb{R}^m \to \mathbb{R}^n$ is called the (Moore–Penrose) *pseudoinverse* or the *generalized inverse* of A and $A^+A = I$. However, in general, $f(\hat{x}) \neq 0$, that is, $A\hat{x} \neq c$.

Another approach, the so-called *Tikhonov regularization*,[1] is to consider the following *regularized problem* for some $\varepsilon > 0$:

$$\inf_{x\in\mathbb{R}^n} f(x) + \varepsilon\|x\|^2_{\mathbb{R}^n}, \quad f(x) \overset{\text{def}}{=} \|Ax - c\|^2_{\mathbb{R}^m}. \tag{2.13}$$

It is quadratic and, by Theorem 2.2(ii), there exists a unique solution $\hat{x}_\varepsilon$ if and only if the matrix $A^\top A + \varepsilon I$ is positive definite. This is verified since

$$\forall x, \quad (A^\top A + \varepsilon I)x \cdot x \geq \varepsilon\|x\|^2_{\mathbb{R}^m}. \tag{2.14}$$

Hence, the following equation has a unique solution $\hat{x}_\varepsilon$:

$$\nabla f(\hat{x}_\varepsilon) + 2\varepsilon\,\hat{x}_\varepsilon = 2A^\top(A\hat{x}_\varepsilon - c) + 2\varepsilon\,\hat{x}_\varepsilon = 0 \quad \Rightarrow x_\varepsilon = [A^\top A + \varepsilon I]^{-1}A^\top c.$$

By taking the inner product of the equation $(A^\top A + \varepsilon I)\hat{x}_\varepsilon = A^\top c$ with $\hat{x}_\varepsilon$, we obtain the following estimates:

$$\begin{aligned}&\|A\hat{x}_\varepsilon\|^2 + \varepsilon\,\|\hat{x}_\varepsilon\|^2 = c \cdot A\hat{x}_\varepsilon \leq \|c\|\,\|A\hat{x}_\varepsilon\|\\ &\Rightarrow \|A\hat{x}_\varepsilon\| \leq \|c\| \quad \text{and} \quad \varepsilon\,\|\hat{x}_\varepsilon\|^2 \leq \|c\|^2.\end{aligned}$$

So there exist a sequence $\varepsilon_n \searrow 0$ and a point $y \in \mathbb{R}^n$ such that $Ax_{\varepsilon_n} \to Ay$ and $\varepsilon_n\,\hat{x}_{\varepsilon_n} \to 0$ as $n \to \infty$. Coming back to the equation for $\hat{x}_{\varepsilon_n}$,

$$0 = (A^\top A + \varepsilon_n I)\hat{x}_{\varepsilon_n} - A^\top c \to A^\top A\,y - A^\top c$$

and y is indeed a solution of the initial minimization problem since

$$\begin{aligned}&\forall z \in \mathbb{R}^n, \quad \|A\,\hat{x}_{\varepsilon_n} - c\|^2 \leq \|A\,\hat{x}_{\varepsilon_n} - c\|^2 + \varepsilon_n\,\|\hat{x}_{\varepsilon_n}\|^2 \leq \|A\,z - c\|^2 + \varepsilon_n\,\|z\|^2\\ &\Rightarrow \forall z \in \mathbb{R}^n, \quad \|A\,y - c\|^2 \leq \|A\,z - c\|^2.\end{aligned}$$

If $[A^\top A]^{-1}$ exists, then $[A^\top A + \varepsilon I]^{-1}A^\top \to A^+$ as ε goes to zero. □

2.2 Least and Greatest Eigenvalues of a Symmetric Matrix

Let A be a symmetric $n \times n$ matrix. Its eigenvalues are real numbers λ such that $\det[A - \lambda I] = 0$. Denote by $\sigma(A)$ the set of eigenvalues of A. Equivalently,

$$\lambda \in \sigma(A) \quad \iff \quad \exists x_\lambda \neq 0 \text{ such that } [A - \lambda I]x_\lambda = 0. \tag{2.15}$$

A nonzero solution x_λ is called an *eigenvector*. The *eigensubspace* corresponding to $\lambda \in \sigma(A)$ is defined as

$$E(\lambda) \overset{\text{def}}{=} \left\{x \in \mathbb{R}^n : [A - \lambda I]x = 0\right\}. \tag{2.16}$$

It is a linear subspace[2] of $\mathbb{R}^n$ that contains all the eigenvectors plus 0.

[1] Andrey Nikolayevich Tychonoff (1906–1993).

[2] Cf. Definition 8.1(i) of Chapter 2.

Define the function

$$x \mapsto f(x) \stackrel{\text{def}}{=} \frac{Ax \cdot x}{\|x\|^2} : \mathbb{R}^n \backslash \{0\} \to \mathbb{R} \tag{2.17}$$

also known as the *Rayleigh's quotient*.[3] This function is $C^{(\infty)}(\mathbb{R}^n \backslash \{0\})$ and at any point $x \neq 0$,

$$\begin{aligned}
\frac{1}{2} df(x;v) &= \frac{Ax \cdot v \, \|x\|^2 - x \cdot v \, Ax \cdot x}{\|x\|^4} \\
&= \frac{1}{\|x\|} \left(A \frac{x}{\|x\|} \cdot v - \frac{x}{\|x\|} \cdot v \, A \frac{x}{\|x\|} \cdot \frac{x}{\|x\|} \right) \\
&= \frac{1}{\|x\|} \left([A - f(x) I] \frac{x}{\|x\|} \right) \cdot v
\end{aligned}$$

$$\Rightarrow \boxed{\exists x \neq 0 \text{ such that } \nabla f(x) = 0 \quad \Longleftrightarrow \quad \exists x \neq 0 \text{ such that } [A - f(x) I] x = 0.}$$

As a result, for any nonzero $x_\lambda \in E(\lambda)$, $f(x_\lambda) = \lambda$ and $\nabla f(x_\lambda) = 0$. Conversely, if there exists $x \neq 0$ such that $\nabla f(x) = 0$, then $f(x)$ is an eigenvalue of A.

Consider the problems:

$$\inf_{0 \neq x \in \mathbb{R}^n} f(x) \quad \text{and} \quad \sup_{0 \neq x \in \mathbb{R}^n} f(x). \tag{2.18}$$

Notice that

$$\forall x \neq 0, \quad f(x) = A \frac{x}{\|x\|} \cdot \frac{x}{\|x\|} = f\left(\frac{x}{\|x\|} \right).$$

As a result,

$$\inf_{0 \neq x \in \mathbb{R}^n} f(x) = \inf_{x \in S} Ax \cdot x \quad \text{and} \quad \sup_{0 \neq x \in \mathbb{R}^n} f(x) = \sup_{x \in S} Ax \cdot x,$$

where $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ is the compact unit sphere in $\mathbb{R}^n$. By the Weierstrass theorem, there exists $x_* \in \mathbb{R}^n$ such that $\|x_*\| = 1$ and

$$\lambda_* \stackrel{\text{def}}{=} \inf_{x \in S} Ax \cdot x = Ax_* \cdot x_* = \frac{Ax_* \cdot x_*}{\|x_*\|^2} = \inf_{0 \neq x \in \mathbb{R}^n} \frac{Ax \cdot x}{\|x\|^2} \tag{2.19}$$

and there exists $x^* \in \mathbb{R}^n$ such that $\|x^*\| = 1$ and

$$\lambda^* \stackrel{\text{def}}{=} \sup_{x \in S} Ax \cdot x = Ax^* \cdot x^* = \frac{Ax^* \cdot x^*}{\|x^*\|^2} = \sup_{0 \neq x \in \mathbb{R}^n} \frac{Ax \cdot x}{\|x\|^2}. \tag{2.20}$$

The sets of solutions are

$$E_* \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n \backslash \{0\} : f(y) = \inf_{0 \neq x \in \mathbb{R}^n} f(x)\} \tag{2.21}$$

$$E^* \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n \backslash \{0\} : f(y) = \sup_{0 \neq x \in \mathbb{R}^n} f(x)\}. \tag{2.22}$$

[3] John William Strutt, 3rd Baron Rayleigh, better known as Lord Rayleigh (1842–1919).

Since the function f is of class $C^{(2)}$ on the open set $\mathbb{R}^n\backslash\{0\}$, conditions (2.3) of Theorem 2.1(i) yield the following characterization of the minimizers and the maximizers of f:

$$\begin{aligned}
&\forall x \in E_*, \quad \nabla f(x) = 0 \quad \text{and} \quad Hf(x) \geq 0 \quad \Rightarrow E(\lambda_*) = E_* \cup \{0\}\\
&\forall x \in E^*, \quad \nabla f(x) = 0 \quad \text{and} \quad Hf(x) \leq 0 \quad \Rightarrow E(\lambda^*) = E^* \cup \{0\}.
\end{aligned}$$

Thus, for any eigenvalue $\lambda \in \sigma(A)$, we have $\lambda_* \leq \lambda \leq \lambda^*$. This raises an interesting question. Since $\nabla f(x_\lambda) = 0$, is x_λ a local minimum or maximum of f when $\lambda_* < \lambda < \lambda^*$? To check that, we compute the Hessian matrix:

$$\begin{aligned}
&\frac{1}{2} d^2 f(x; v; w)\\
&= \frac{1}{\|x\|^8} \Big\{ \|x\|^4 \Big[Aw \cdot v \, \|x\|^2 + Ax \cdot v \, 2x \cdot w - w \cdot v \, Ax \cdot x - x \cdot v \, 2Ax \cdot w \Big]\\
&\qquad\qquad - 2 \, \|x\|^2 x \cdot w \Big[Ax \cdot v \, \|x\|^2 - x \cdot v \, Ax \cdot x \Big] \Big\}.
\end{aligned}$$

For $w = v$ and x such that $\nabla f(x) = 0$, we have $Ax = f(x)x$ and

$$\begin{aligned}
&\frac{1}{2} d^2 f(x; v; v)\\
&= \frac{1}{\|x\|^6} \Big\{ \|x\|^2 \Big[Av \cdot v \, \|x\|^2 - v \cdot v \, Ax \cdot x \Big] - 2x \cdot v \Big[Ax \cdot v \, \|x\|^2 - x \cdot v \, Ax \cdot x \Big] \Big\}\\
&= \frac{1}{\|x\|^2} \Big\{ Av \cdot v - f(x) \, \|v\|^2 \Big\}\\
&\qquad\qquad \Rightarrow \frac{\|x\|^2}{2} Hf(x) v \cdot v = Av \cdot v - f(x) \, \|v\|^2\\
&\qquad \Rightarrow \boxed{\forall v \neq 0, \quad \frac{\|x\|^2}{2} Hf(x) v \cdot v = (f(v) - f(x)) \, \|v\|^2.}
\end{aligned}$$

In particular, for our λ, $\lambda_* < \lambda < \lambda^*$, and $x_\lambda \neq 0$ such that $\nabla f(x_\lambda) = 0$, we get

$$\text{with } v = x_*, \quad \frac{\|x_\lambda\|^2}{2} Hf(x_\lambda) x_* \cdot x_* = (f(x_*) - f(x_\lambda)) \, \|x_*\|^2 < 0 \tag{2.23}$$

$$\text{with } v = x^*, \quad \frac{\|x_\lambda\|^2}{2} Hf(x_\lambda) x^* \cdot x^* = (f(x^*) - f(x_\lambda)) \, \|x^*\|^2 > 0 \tag{2.24}$$

and x_λ is neither a local minimum nor a local maximizer of f.

2.3 ▶ Hadamard Semidifferential of the Least Eigenvalue

In this section, we apply the results of section 5 of Chapter 3 to compute the Hadamard semidifferential of the least eigemvalue of a symmetric matrix with respect to a set of parameters (see also Remark 5.2 of Chapter 3).

Recall from section 2.2 that for an $n \times n$ symmetric matrix, the least eigenvalue of A is given by

$$\lambda(A) = \inf_{x \in S} Ax \cdot x \tag{2.25}$$

and the minimizers are given by

$$X(0) \stackrel{\text{def}}{=} \{x \in S : [A - \lambda(A)I]x = 0\} \neq \varnothing. \tag{2.26}$$

Let B be another $n \times n$ symmetric matrix and $t \geq 0$. From the previous result, the problem

$$\lambda(A + t\,B) = \inf_{x \in S} [A + t\,B]x \cdot x \tag{2.27}$$

also has minimizers given by

$$X(t) \stackrel{\text{def}}{=} \{x \in S : \{[A + t\,B] - \lambda([A + t\,B])I\}x = 0\} \neq \varnothing. \tag{2.28}$$

We first use Theorem 5.1 to prove the existence of the limit

$$d\lambda(A; B) \stackrel{\text{def}}{=} \lim_{t \searrow 0} \frac{\lambda(A + t\,B) - \lambda(A)}{t} \tag{2.29}$$

and explicitly compute its expression. We next show that $\lambda(A)$ is Lipschitzian and concave and we conclude that $\lambda(A)$ is Hadamard semidifferentiable.

Theorem 2.3. *Let* Sym_n *be the vector space of all symmetric $n \times n$ matrices endowed with the matrix norm*

$$\|A\| \stackrel{\text{def}}{=} \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

(i) *The function* $A \mapsto \lambda(A) : \text{Sym}_n \to \mathbb{R}$ *is concave and Lipschitz continous:*

$$\forall A, B \in \text{Sym}_n, \quad |\lambda(b) - \lambda(A)| \leq \|B - A\|.$$

(ii) *For all* $A, B \in \text{Sym}_n$,

$$\begin{aligned} d_H\lambda(A; B) &\stackrel{\text{def}}{=} \lim_{\substack{t \searrow 0 \\ W \to B}} \frac{\lambda(A + t\,W) - \lambda(A)}{t} \\ &= \min_{x \in X(0)} Bx \cdot x = \min_{x \in X(0)} x \otimes x \cdot\cdot B, \end{aligned} \tag{2.30}$$

where $X(0) = \{x \in S : [A - \lambda(A)\,I]x = 0\}$, *and*

$$(x \otimes y)_{ij} = x_i\, y_j,\ i, j = 1, \ldots, n, \quad \text{and} \quad X \cdot\cdot Y = \sum_{1 \leq i, j \leq n} X_{ij}\, Y_{ij}$$

are, respectively, the tensor product $x \otimes y$ *of two vectors x and y and the* double inner product $X \cdot\cdot Y$ *of two matrices X and Y.*

(iii) *If $\lambda(A)$ is simple, that is, the dimension of its associated eigensubspace is one, then $\lambda(A)$ is Hadamard and Fréchet differentiable and*

$$d_H\lambda(A;B) = Bx\cdot x = x\otimes x \cdot\cdot B, \quad \forall x \in X(0). \tag{2.31}$$

(Note that $x\otimes x$ is a positive semidefinite matrix of trace equal to one.)

(iv) *Let the map $\vec{p}\mapsto A(\vec{p}) : \mathbb{R}^k \to \mathrm{Sym}_n$ be Hadamard semidifferentiable and consider the function*

$$\vec{p}\mapsto \ell(\vec{p}) \overset{\text{def}}{=} \lambda(A(\vec{p})) : \mathbb{R}^k \to \mathbb{R}. \tag{2.32}$$

Then

$$d_H\ell(\vec{p};\vec{v}) = d_H\lambda(A(\vec{p}); d_H A(\vec{p};\vec{v})) = \inf_{x\in E(\vec{p})} x\otimes x \cdot\cdot d_H A(\vec{p};\vec{v}),$$

where $E(\vec{p})$ are the eigenvectors of norm one associated with $\lambda(A(\vec{p}))$.

Remark 2.1.
Formula (2.30) gives a complete description of the nondifferentiability of $\lambda(A)$. When $\lambda(A)$ is not simple, the dimension of its associated eigensubspace is at least two. So, as the direction B changes, the eigenvector achieving the infimum of $Bx\cdot x$ changes, resulting in the nonlinearity of the map $B\mapsto d_H\lambda(A;B)$. □

Remark 2.2.
The theorem holds with obvious changes for the greatest eigenvalue of the matrix. The concavity becomes convexity and the minimum becomes a maximum in the conclusions. □

Proof. (i)–(ii) From Theorem 5.2(i) and (ii) in Chapter 3.

(iii) When $\lambda(A)$ is simple, $X(0) = \{\pm x_0\}$, and $Bx_0\cdot x_0 = B(-x_0)\cdot(-x_0)$ (or $x_0\otimes x_0 = (-x_0)\otimes(-x_0)$), so that the infimum can be dropped and the mapping $B\mapsto Bx_0\cdot x_0 = x_0\otimes x_0\cdot\cdot B : \mathrm{Sym}_n \to \mathbb{R}$ is linear.

(iv) From Theorem 5.2(iii) in Chapter 3. □

3 Optimality Conditions for U Convex

Section 3.2 gives a necessary and sufficient optimality condition for a Gateaux differentiable convex function and section 3.3 gives a necessary optimality condition for a semidifferentiable function that is not necessarily convex. Finally, section 3.4 gives a necessary and sufficient condition for an arbitrary convex function using the lower semidifferential.

3.1 Cones

After the linear and affine subspaces of section 8 in Chapter 2, the third important convex set is the cone in 0. When U is specified by a number of affine equalities, U is an affine subspace. When U is specified by affine inequalities, U is the translation of a cone in 0. The set of admissible directions at a point $x\in\overline{U}$ will also be a cone in 0.

Definition 3.1. (i) $C \subset \mathbb{R}^n$ is a *cone in* 0 if

$$\forall \lambda > 0, \forall x \in C, \quad \lambda x \in C. \tag{3.1}$$

(ii) $C \subset \mathbb{R}^n$ is a *convex cone in* 0 if C is a cone in 0 and C is convex. □

Several examples of cones in 0 are given in Figures 4.3, 4.4, and 4.5.

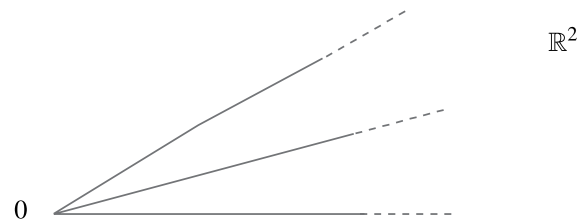


Figure 4.3. *A nonconvex closed cone in* 0*.*

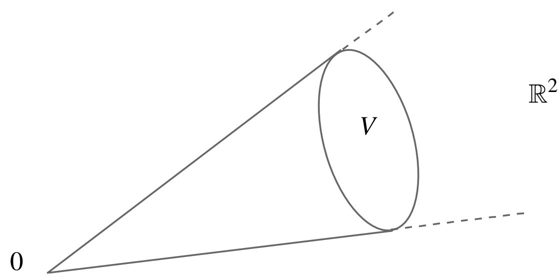


Figure 4.4. *Closed convex cone in* 0 *generated by the closed convex set* V*.*

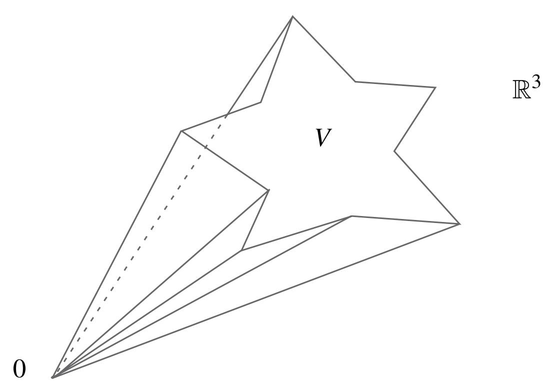


Figure 4.5. *Closed convex cone in* 0 *generated in* $\mathbb{R}^3$ *by a nonconvex set* V*.*

Theorem 3.1. *$U \subset \mathbb{R}^n$ is a convex cone (in 0) if and only if*

$$\forall \lambda > 0, \ \forall x \in U, \ \lambda x \in U \tag{3.2}$$

$$\forall x_1, x_2 \in U, \ x_1 + x_2 \in U. \tag{3.3}$$

Proof. ($\Rightarrow$) By definition of a cone, we have (3.2). By convexity,

$$\forall x_1, x_2 \in U, \ \frac{1}{2}x_1 + \frac{1}{2}x_2 \in U$$

and by (3.2),

$$x_1 + x_2 = 2\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right) \in U.$$

($\Leftarrow$) From (3.2), U is a cone. For all x_1 and x_2 in U and $\lambda \in]0,1[$,

$$\lambda x_1 \in U \quad \text{and} \quad (1-\lambda)x_2 \in U.$$

By using (3.3), we finally get

$$\lambda x_1 + (1-\lambda)x_2 \in U.$$

By definition, U is convex and a fortiori a convex cone. □

3.2 Convex Gateaux Differentiable Objective Function

It is now possible to completely characterize the minimum of a convex Gateaux differentiable function over a convex set U. The condition will be further specialized to the linear and affine subspaces and to convex cones in 0.

Theorem 3.2. *Let $\varnothing \neq U \subset \mathbb{R}^n$ be convex and $f : \mathbb{R}^n \to \mathbb{R}$ be convex and Gateaux differentiable on U. There exists a minimizer of f in U if and only if*

$$\exists x \in U \text{ such that } \forall y \in U, \quad \nabla f(x) \cdot (y - x) \geq 0. \tag{3.4}$$

Since U is a convex subset of $\mathbb{R}^n$, condition (3.4) means that U is contained in the closed half-space defined by (see Figure 4.6)

$$\{y \in \mathbb{R}^n : \nabla f(x) \cdot (y - x) \geq 0\}.$$

Proof. If condition (3.4) is verified, then we know, by Theorem 4.1 of Chapter 3, that for a convex Gateaux differentiable function,

$$\begin{gathered}\forall y \in U, \quad f(y) - f(x) \geq \nabla f(x) \cdot (y - x) \geq 0 \\ \Rightarrow \forall y \in U, \quad f(y) \geq f(x).\end{gathered}$$

So the point $x \in U$ minimizes f with respect to U. Conversely, there exists $x \in U$ that minimizes f in U, then, by convexity of f, for all $y \in U$ and $t \in]0,1]$,

$$f(x + t(y-x)) - f(x) \geq 0 \quad \Rightarrow \quad \frac{f(x + t(y-x)) - f(x)}{t} \geq 0 \tag{3.5}$$

and since f is Gateaux differentiable, by going to the limit,

$$\nabla f(x) \cdot (y - x) = df(x; y - x) = \lim_{t \searrow 0} \frac{f(x + t(y-x)) - f(x)}{t} \geq 0 \tag{3.6}$$

that completes the proof. □

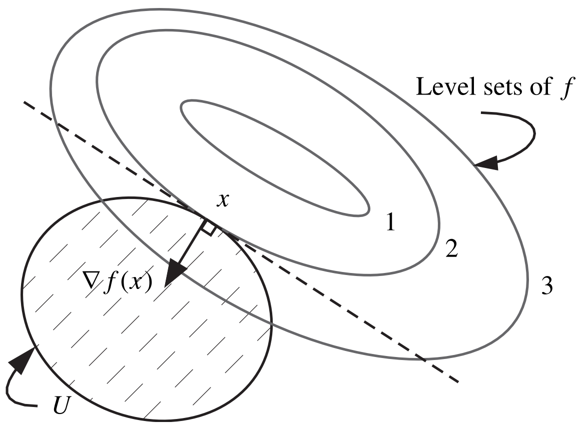


Figure 4.6. *Convex set U tangent to the level set of f through $x \in U$.*

Corollary 1. *Let the assumptions of Theorem* 3.2 *hold.*

(i) *If $U = S$, a linear subspace, then condition* (3.4) *is equivalent to*

$$\exists x \in S \text{ such that } \forall y \in S, \quad \nabla f(x) \cdot y = 0. \tag{3.7}$$

In particular, $\nabla f(x) \in S^{\perp}$.

(ii) *If $U = A$, an affine subspace, then condition* (3.4) *is equivalent to*

$$\exists x \in A \text{ such that } \forall y \in A, \quad \nabla f(x) \cdot (y - x) = 0 \tag{3.8}$$

or, equivalently, to

$$\exists x \in A \text{ such that } \forall y \in S, \quad \nabla f(x) \cdot y = 0, \tag{3.9}$$

where S is the linear subspace associated with the affine subspace A. In particular, $\nabla f(x) \in S^{\perp}$ (see Figure 4.7*).*

(iii) *If $U = C$, a convex cone in* 0*, then condition* (3.4) *is equivalent to*

$$\exists x \in C \text{ such that } \nabla f(x) \cdot x = 0 \text{ and } \forall y \in C, \quad \nabla f(x) \cdot y \geq 0. \tag{3.10}$$

Remark 3.1.
Equality (3.7), which must be verified for each $y \in S$, is the prototype of a *variational equation* that implies here that $\nabla f(x) \in S^{\perp}$ and not necessarily the equation $\nabla f(x) = 0$. Hence, a variational equation is a weaker form of equation. By analogy, inequality (3.10), which must be verified for each $y \in C$, is referred to as a *variational inequality*. The terminology *variational inequation* would have been more appropriate, but the word "inequation" does not exist in English. □

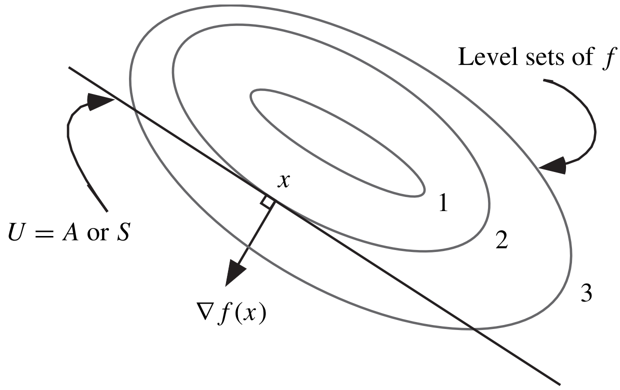


Figure 4.7. *Tangency of the affine subspace A or the linear subspace S to a level set of f.*

Proof. (i) If $U = S$, a linear subspace, then for all $y \in S$, $x \pm y \in S$. By substituting into (3.4), we get

$$\forall y \in S, \quad \pm\nabla f(x) \cdot y \geq 0 \quad \Rightarrow \nabla f(x) \cdot y = 0.$$

Conversely, we have

$$\nabla f(x) \cdot x = 0 \text{ and } \forall y \in S, \nabla f(x) \cdot y = 0 \quad \Rightarrow \forall y \in S, \nabla f(x) \cdot (y - x) = 0 \geq 0.$$

(ii) Since $U = A$, an affine subspace, for all $y \in A$ and $\alpha \in \mathbb{R}$, $\alpha y + (1-\alpha)x \in A$ and

$$\alpha \nabla f(x) \cdot (y - x) = \nabla f(x) \cdot (\alpha y + (1-\alpha)x - x) \geq 0$$

and this implies that $\nabla f(x) \cdot (y - x) = 0$ for all $y \in A$. Since A is affine, $S = A - x$ is a linear subspace by Theorem 8.1 of Chapter 2 and for all $s \in S$, $\nabla f(x) \cdot s = 0$. Therefore, $\nabla f(x) \in S^\perp$. Conversely, $S = A - x$ and for all $y \in A$, $\nabla f(x) \cdot (y - x) = 0 \geq 0$.

(iii) If $U = C$, a convex cone in 0, then for all $y \in C$, $x + y \in C$ and

$$\forall y \in C, \quad \nabla f(x) \cdot y = \nabla f(x) \cdot (y + x - x) \geq 0.$$

For $y = 2x$ and $y = x/2$, we get

$$\nabla f(x) \cdot (2x - x) \geq 0 \text{ and } \nabla f(x) \cdot \left(\frac{x}{2} - x\right) \geq 0 \quad \Rightarrow \nabla f(x) \cdot x = 0.$$

Conversely,

$$\forall y \in C, \quad \nabla f(x) \cdot y \geq 0 \text{ and } \nabla f(x) \cdot x = 0 \quad \Rightarrow \forall y \in C, \quad \nabla f(x) \cdot (y - x) \geq 0.$$

This completes the proof. □

Remark 3.2.
When U is given by

$$\left\{x \in \mathbb{R}^n : g_j(x) \leq 0, \quad 1 \leq j \leq m\right\}$$

for convex (and continuous) functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, then U is convex (and closed). Moreover, if the objective function f satisfies the assumptions of Theorem 3.2, then condition (3.4) is necessary and sufficient. $\square$

From part (i) of the corollary, we get a useful necessary and sufficient condition for quadratic objective functions that completes Theorem 2.2(i).

Theorem 3.3. *Consider the quadratic function*

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x + c \tag{3.11}$$

for an $n \times n$ symmetrical matrix A, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$. The following conditions are equivalent:

$$\text{(a)} \qquad \exists \hat{x} \in \mathbb{R}^n, \quad f(\hat{x}) = \inf_{x \in \mathbb{R}^n} f(x), \tag{3.12}$$

$$\text{(b)} \qquad \inf_{x \in \mathbb{R}^n} f(x) > -\infty, \tag{3.13}$$

$$\text{(c)} \qquad \exists \hat{x} \in \mathbb{R}^n \text{ such that } A\hat{x} + b = 0 \tag{3.14}$$

$$\text{and } \forall v \in \mathbb{R}^n, \quad Av \cdot v \ge 0. \tag{3.15}$$

Condition (3.14) *is equivalent to*

$$\forall y \in \operatorname{Ker} A, \quad b \cdot y = 0. \tag{3.16}$$

Condition (3.15) *is equivalent to the convexity of f on $\mathbb{R}^n$.*

Proof. From Theorem 2.2(i), (a) $\iff$ (c). Obviously (c) $\Rightarrow$ (a) $\Rightarrow$ (b). To prove (b) $\Rightarrow$ (c), we proceed by contradiction.

If (3.14) is not true, then $b \notin \operatorname{Im} A$. Since the image of A is a linear subspace of $\mathbb{R}^n$, it is closed. From part (i) of the Corollary 1, there exists $\hat{x} \in \mathbb{R}^n$ such that

$$0 < \|A\hat{x} + b\|^2 = \inf_{x \in \mathbb{R}^n} \|Ax + b\|^2$$

$$\forall x \in \mathbb{R}^n, \quad (A\hat{x} + b) \cdot Ax = 0 \quad \Rightarrow A\hat{x} + b \in [\operatorname{Im} A]^\perp.$$

Given $n \in \mathbb{N}$, set $x_n = -n(A\hat{x} + b)$:

$$f(x_n) = c - 2n\, b \cdot (A\hat{x} + b) + n^2 \underbrace{A(A\hat{x} + b) \cdot \underbrace{(A\hat{x} + b)}_{\in [\operatorname{Im} A]^\perp}}_{=0}$$

$$\Rightarrow \; f(x_n) = c - 2n\, b \cdot (A\hat{x} + b) = c - 2n\, \|A\hat{x} + b\|^2 + 2n\, A\hat{x} \cdot \underbrace{(A\hat{x} + b)}_{\in \operatorname{Im} A]^\perp}$$

$$= c - 2n\, b \cdot (A\hat{x} + b) = c - 2n\, \|A\hat{x} + b\|^2 \to -\infty$$

as $n \to +\infty$ since $A\hat{x} + b \neq 0$. This contradicts condition (3.13).

If condition (3.15) is not verified, then there exists $w \neq 0$ such that $Aw \cdot w < 0$. For each $n \in \mathbb{N}$, let $x_n = \hat{x} + nw$. Then, as n goes to infinity,

$$f(x_n) = c + 2b \cdot x_n + Ax_n \cdot x_n = f(\hat{x}) + 2n(A\hat{x} + b) \cdot w + n^2\, Aw \cdot w \to -\infty$$

since $Aw \cdot w$ is strictly negative. This contradiction yields (3.13).

Since $\operatorname{Im} A$ is closed, condition (3.14) is equivalent to

$$b \in \operatorname{Im} A = (\operatorname{Ker} A)^{\perp} \quad \Rightarrow \forall w \in \operatorname{Ker} A, \quad b \cdot w = 0.$$

Finally, condition (3.15) is equivalent to the convexity of f on $\mathbb{R}^n$ by Theorem 4.3(i) of Chapter 3. □

Example 3.1 ($U = S$, a linear subspace)**.**
Consider the objective function

$$f(x_1, x_2, x_3) = \frac{1}{2}\left[(x_1 - 1)^2 + (x_2 - 2)^2 + (x_3 - 3)^2\right]$$

to minimize with respect to the set

$$U = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}.$$

U is a linear subspace of $\mathbb{R}^3$. Since f is polynomial, it is infinitely differentiable and its gradient and Hessian matrix are given by

$$\nabla f(x) = \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \\ x_3 - 3 \end{bmatrix} \quad \text{and} \quad Hf(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The function f is then strictly convex on $\mathbb{R}^3$ (see Exercise 7.2 of Chapter 3). By Corollary 1(i), the minima $x = (x_1, x_2, x_3)$ of f with respect to U (if they exist) are characterized by

$$x_1 + x_2 + x_3 = 0$$

$$\text{and} \quad \begin{cases} \forall y = (y_1, y_2, y_3) \text{ such that } y_1 + y_2 + y_3 = 0 \\ \nabla f(x) \cdot y = (x_1 - 1)y_1 + (x_2 - 2)y_2 + (x_3 - 3)y_3 = 0. \end{cases}$$

By eliminating x_3 and y_3, this condition becomes

$$\forall y_1, y_2 \in \mathbb{R}, \quad (x_1 - 1)y_1 + (x_2 - 2)y_2 + (x_1 + x_2 + 3)(y_1 + y_2) = 0.$$

By factoring, we get

$$\begin{aligned} &\forall y_1, y_2 \in \mathbb{R}, \quad (2x_1 + x_2 + 2)y_1 + (2x_2 + x_1 + 1)y_2 = 0 \\ &\Rightarrow \quad 2x_1 + x_2 + 2 = 0 \text{ and } 2x_2 + x_1 + 1 = 0 \\ &\Rightarrow \quad x_1 = -1, x_2 = 0 \text{ and } x_3 = -x_1 - x_2 = 1. \end{aligned}$$ □

Example 3.2 ($U = A$, an affine subspace)**.**
Consider the objective function

$$f(x_1, x_2, x_3) = \frac{1}{2}\left[x_1^2 + x_2^2 + x_3^2\right]$$
$$A = \{(x_1, x_2, x_3) : x_1 + 2x_2 + 3x_3 = 4\}.$$

It is easy to check that A is an affine subspace of $\mathbb{R}^3$ and that the associated linear subspace is $S = \{(y_1, y_2, y_3) : y_1 + 2y_2 + 3y_3 = 0\}$. Since f is polynomial, it is infinitely differentiable and its gradient and Hessian matrix are given by

$$\nabla f(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{and} \quad Hf(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The function f is then strictly convex in $\mathbb{R}^3$ (see Exercise 7.2 of Chapter 3). By Corollary 1(ii), all minima $x = (x_1, x_2, x_3)$ of f with respect to A (if they exist) are characterized by

$$x_1 + 2x_2 + 3x_3 = 4$$

$$\text{and} \quad \begin{cases} \forall y = (y_1, y_2, y_3) \text{ such that } y_1 + 2y_2 + 3y_3 = 0 \\ \nabla f(x) \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3 = 0. \end{cases}$$

By eliminating x_3 and y_3, those conditions become

$$\forall y_1,\, y_2 \in \mathbb{R}, \quad x_1 y_1 + x_2 y_2 + \frac{1}{9}(4 - x_1 - 2x_2)(-y_1 - 2y_2) = 0.$$

By factoring, we get

$$\begin{aligned} \forall y_1,\, y_2 \in \mathbb{R}, \quad & \left[x_1 - \frac{1}{9}(4 - x_1 - 2x_2)\right] y_1 + \left[x_2 - \frac{2}{9}(4 - x_1 - 2x_2)\right] y_2 = 0 \\ \Rightarrow \quad & x_1 - \frac{1}{9}(4 - x_1 - 2x_2) = 0 \text{ and } x_2 - \frac{2}{9}(4 - x_1 - 2x_2) = 0 \\ \Rightarrow \quad & x_1 - \frac{4}{9} + \frac{1}{9}x_1 + \frac{2}{9}x_2 = 0 \text{ and } x_2 - \frac{8}{9} + \frac{2}{9}x_1 + \frac{4}{9}x_2 = 0 \\ \Rightarrow \quad & -2 + 5x_1 + x_2 = 0 \text{ and } -8 + 2x_1 + 13x_2 = 0 \\ \Rightarrow \quad & x_1 = \frac{2}{7}, \quad x_2 = \frac{4}{7}, \quad x_3 = \frac{6}{7}. \end{aligned}$$

□

Example 3.3 ($U = C$, a closed convex cone in 0)**.**
Consider the objective function f and the cone C:

$$f(x) \stackrel{\text{def}}{=} 2x_1^2 + 4x_2^2 - x_1 - x_2 \tag{3.17}$$

$$C \stackrel{\text{def}}{=} \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0 \text{ and } x_2 \geq 0\}. \tag{3.18}$$

Since f is polynomial, it is infinitely differentiable and its gradient and Hessian matrix are given by

$$\nabla f(x) = \begin{bmatrix} 4x_1 - 1 \\ 8x_2 - 1 \end{bmatrix} \quad \text{and} \quad Hf(x) = \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}. \tag{3.19}$$

The function f is then strictly convex on $\mathbb{R}^2$. It is readily seen that the set C is a closed convex cone in 0. Then by Corollary 1(iii),

$$x_1 \geq 0, \;\; x_2 \geq 0 \tag{3.20}$$

$$(4x_1 - 1)x_1 + (8x_2 - 1)x_2 = 0 \tag{3.21}$$

$$\forall y_1 \geq 0, \, \forall y_2 \geq 0, \quad (4x_1 - 1)y_1 + (8x_2 - 1)y_2 \geq 0. \tag{3.22}$$

Hence, by choosing $y_2 = 0$,

$$4x_1 - 1 \geq 0 \tag{3.23}$$

and by choosing $y_1 = 0$,

$$8x_2 - 1 \geq 0. \tag{3.24}$$

From (3.20), (3.21), (3.23), and (3.24)

$$\begin{aligned} x_1 \geq 0, \ \ 4x_1 - 1 \geq 0 \ \text{ and } \ x_1(4x_1 - 1) = 0 \\ x_2 \geq 0, \ \ 8x_2 - 1 \geq 0 \ \text{ and } \ x_2(8x_2 - 1) = 0. \end{aligned}$$

If $x_1 = 0$, then $4x_1 - 1 = -1 \not\geq 0$. Therefore, $x_1 = 1/4$. On the other hand, if $x_2 = 0$, then $8x_2 - 1 = -1 \not\geq 0$. Therefore, $x_2 = 1/8$. We will also be able to use Karush–Kuhn–Tucker theorem with its multipliers, but in that case, Corollary 1(iii) is more direct. We simultaneously get the characterization of the minimizers and their existence. □

Example 3.4 (Mixed case of equalities and inequalities)**.**
Consider the objective function

$$\begin{aligned} f(x_1,x_2,x_3) &\stackrel{\text{def}}{=} \frac{1}{2}\left[x_1^2 + x_2^2 + x_3^2\right] \\ U &\stackrel{\text{def}}{=} \{(x_1,x_2,x_3) : 3x_1 + 2x_2 + x_3 = 4 \text{ and } x_1 \leq 0\}. \end{aligned}$$

Since f is polynomial, it is infinitely differentiable and its gradient and Hessian matrix are given by

$$\nabla f(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{and} \quad Hf(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.25}$$

The function f is then strictly convex on $\mathbb{R}^3$. It is easy to check that the set U is closed and convex. Therefore, all minima $x = (x_1,x_2,x_3)$ of f with respect to U (if they exist) are characterized by

$$3x_1 + 2x_2 + x_3 = 4 \quad \text{and} \quad x_1 \leq 0$$

$$\text{and} \quad \begin{cases} \forall y = (y_1,y_2,y_3) \text{ such that } 3y_1 + 2y_2 + y_3 = 4 \quad \text{and} \quad y_1 \leq 0, \\ \nabla f(x) \cdot (y - x) = x_1(y_1 - x_1) + x_2(y_2 - x_2) + x_3(y_3 - x_3) \geq 0. \end{cases}$$

By eliminating x_3 and y_3, this condition becomes: for all $y_1 \leq 0$ and all $y_2 \in \mathbb{R}$,

$$x_1(y_1 - x_1) + x_2(y_2 - x_2) - (4 - 3x_1 - 2x_2)(3(y_1 - x_1) + 2(y_2 - x_2)) \geq 0.$$

By factoring, we get for all $y_1 \leq 0$ and all $y_2 \in \mathbb{R}$,

$$\begin{aligned} [x_1 - 3(4 - 3x_1 - 2x_2)](y_1 - x_1) + [x_2 - 2(4 - 3x_1 - 2x_2)](y_2 - x_2) \geq 0 \\ [10x_1 + 6x_2 - 12](y_1 - x_1) + [5x_2 + 6x_1 - 8](y_2 - x_2) \geq 0. \end{aligned}$$

This last inequality can be decoupled since y_2 is an unconstrained variable. This implies

$$5x_2 + 6x_1 - 8 = 0$$
$$10x_1 + 6x_2 - 12 \le 0$$
$$[10x_1 + 6x_2 - 12]x_1 = 0 \text{ and } x_1 \le 0.$$

Either $x_1 = 0$ and from the first equation $5x_2 = 8$, or (x_1, x_2) is solution of the system

$$\begin{bmatrix} 6 & 5 \\ 10 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix} \quad \Rightarrow x_1 = \frac{6}{7} \quad x_2 = \frac{4}{7}.$$

Since $x_1 \le 0$, this last solution is not acceptable. The last one remaining is $(x_1, x_2) = (0, 8/5)$ for which the condition

$$10x_1 + 6x_2 - 12 = 6\frac{8}{5} - 12 = 12\left(\frac{4}{5} - 1\right) \le 0$$

is verified. Moreover, $x_3 = 4 - 3x_1 - 2x_2 = 4 - 2 \times 8/5 = 4/5$. □

3.3 Semidifferentiable Objective Function

We now characterize the minimum of a semidifferentiable function. The condition will be necessary and sufficient for a convex function on a convex set. When f is not convex, the level sets of f are not convex, but we still get the tangency of the set U to the level set of f through the minimizer as shown in Figure 4.8.

Theorem 3.4. *Let U, $\varnothing \ne U \subset \mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}$ be a function such that $df(x; y - x)$ exists for all x and y in U.*

(i) *If f has a minimizer on U, then*

$$\exists x \in U, \forall y \in U, \quad df(x; y - x) \ge 0. \tag{3.26}$$

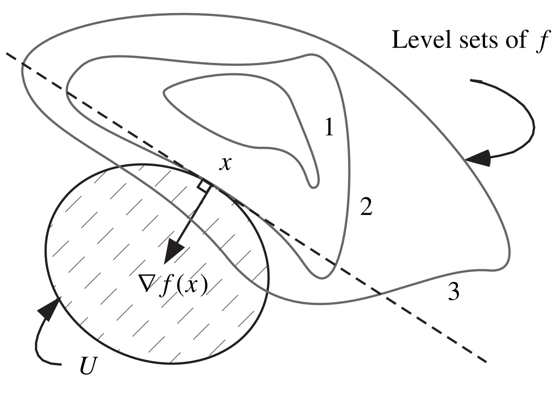


Figure 4.8. *Tangency of U to a level set of the function f at $x \in U$.*

(ii) *If there exists $x \in U$ such that* (3.26) *is verified and if there exists a neighborhood $V(x)$ of x such that f is convex on $U \cap V(x)$, then*

$$\forall y \in V(x) \cap U, \quad f(y) \geq f(x), \tag{3.27}$$

that is, x is a local minimizer of f on U.

Proof. (i) If f has a local minimum at $x \in U$, then there exists a ball $B_\varepsilon(x) = \{y \in \mathbb{R}^n : \|y - x\| < \varepsilon\}$ of radius ε, $0 < \varepsilon \leq 1$, such that

$$\forall y \in U \cap B_\varepsilon(x), \quad f(y) \geq f(x).$$

Given $y \in U$, choose $\bar{t} = \min\{\varepsilon/(2|y-x|), 1\}$ so that for all t, $0 < t < \bar{t}$, we have $ty + (1-t)x \in U \cap B_\varepsilon(x)$. Hence,

$$\forall t \in]0, \bar{t}], \quad f(x + t(y-x)) - f(x) = f(ty + (1-t)x) - f(x) \geq 0.$$

Dividing by $t > 0$, for all $y \in U$,

$$\forall t \text{ such that } 0 < t \leq \bar{t}, \quad \frac{f(x + t(y-x)) - f(x)}{t} \geq 0$$

and going to the limit as t goes to 0,

$$\forall y \in U, \quad df(x; y-x) \geq 0.$$

(ii) Since f is convex on $U \cap V(x)$,

$$\forall y \in U \cap V(x), \forall t \in]0,1] \quad f(ty + (1-t)x) \leq tf(y) + (1-t)f(x)$$

and since f is semidifferentiable at x in the direction $y - x$,

$$f(x + t(y-x)) - f(x) \leq t[f(y) - f(x)] \quad \Rightarrow df(x; y-x) \leq f(y) - f(x).$$

So, if (3.26) is verified, then

$$\forall y \in U \cap V(x), \ f(y) - f(x) \geq df(x; y-x) \geq 0$$

and x is a local minimizer of f on U. □

3.4 ▶ Arbitrary Convex Objective Fonction

Recall that we can associate with a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a set U, the function f_U so that

$$\inf f(U) = \inf f_U(\mathbb{R}^n) \quad \text{and} \quad \operatorname{argmin} f(U) = \operatorname{argmin} f_U(\mathbb{R}^n). \tag{3.28}$$

If $\operatorname{dom} f_U = \varnothing$, then

$$f(x) = +\infty \text{ on } U \quad \text{and} \quad \operatorname{argmin} f(U) = U. \tag{3.29}$$

If $\operatorname{dom} f_U \neq \varnothing$, then

$$\operatorname{argmin} f(U) = \operatorname{argmin} f_U(\mathbb{R}^n) \subset \operatorname{dom} f_U = U \cap \operatorname{dom} f, \tag{3.30}$$

but $\operatorname{argmin} f(U)$ might be empty.

Theorem 3.5. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\text{dom } f \neq \varnothing$, and U, $\varnothing \neq U \subset \mathbb{R}^n$, be convex. The following conditions are equivalent:*

(i) *there exists $x \in U \cap \text{dom } f$ such that*

$$f(x) = \inf f(U); \tag{3.31}$$

(ii) *there exists $x \in U \cap \text{dom } f$ such that*

$$\forall y \in U, \quad \underline{d} f(x; y-x) \geq 0; \tag{3.32}$$

(iii) *there exists $x \in \text{dom } f_U$ such that*

$$\forall y \in U, \quad \underline{d} f_U(x; y-x) \geq 0. \tag{3.33}$$

Proof. (i) $\Rightarrow$ (ii) Since x is a minimizer and f and U are convex, for each $y \in U$ and all $0 < t < 1$, $x + t(y-x) \in U$ and

$$\begin{aligned} 0 &\leq \frac{f(x+t(y-x)) - f(x)}{t} \\ \Rightarrow 0 &\leq \liminf_{t \searrow 0} \frac{f(x+t(y-x)) - f(x)}{t} = \underline{d} f(x; y-x). \end{aligned}$$

(ii) $\Rightarrow$ (iii) By definition of f_U, we have $f_U \geq f$. As $\text{dom } f_U = U \cap \text{dom } f$, at $x \in U \cap \text{dom } f$, $f_U(x) = f(x)$ and $\underline{d} f_U(x; y-x) \geq \underline{d} f(x; y-x) \geq 0$.

(iii) $\Rightarrow$ (i) For $x \in \text{dom } f_U$, all $y \in U$, and all $0 < t < 1$, we have, by convexity of f_U,

$$\begin{aligned} \frac{f_U(x+t(y-x)) - f_U(x)}{t} &\leq f_U(y) - f_U(x) \\ \Rightarrow f_U(y) - f_U(x) &\geq \liminf_{t \searrow 0} \frac{f_U(x+t(y-x)) - f_U(x)}{t} = \underline{d} f_U(x; y-x) \geq 0 \\ &\Rightarrow \forall y \in U, \quad f_U(y) - f_U(x) \geq 0. \end{aligned}$$

Since $\text{dom } f_U = U \cap \text{dom } f$, $f_U = f$ on U and for all $y \in U$, $f(y) \geq f(x)$. Therefore, the point $x \in U \cap f$ is a minimizer of f with respect to U. $\square$

Example 3.5.
Consider the convex lsc function in Example 9.1 of Chapter 2: $f(x,y) = x^2/y$ for $y > 0$, $f(0,0) = 0$, and $f(x,y) = +\infty$ elsewhere. We have shown that its domain $\text{dom } f = \{(x,y) : x \in \mathbb{R} \text{ and } y > 0\} \cup \{(0,0)\}$ is made up of interior points $\{(x,y) : x \in \mathbb{R} \text{ and } y > 0\}$ and the boundary point $(0,0)$. So, $\text{dom } f$ is neither open nor closed, f is Lipschitzian continuous at interior points, but f is not Lipschitzian continuous at $(0,0)$.

It is easy to check that the set of minimizers of f over $\mathbb{R}^2$ is $\text{argmin} f(\mathbb{R}^2) = \{(0,y) : y \geq 0\}$. As f is Fréchet differentiable at interior points of $\text{dom } f$,

$$\begin{gathered} \forall y > 0, \quad \nabla f(x,y) = \frac{x}{y^2} \begin{bmatrix} 2y \\ -x \end{bmatrix} \quad \Rightarrow \forall y > 0, \quad \nabla f(0,y) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Rightarrow \forall y > 0, \forall (x', y'), \quad \underline{d} f(0,y); (x',y') - (0,y)) = \nabla f(0,y) \cdot (x', y'-y) = 0 \end{gathered}$$

and, from part (ii), the points $\{(0,y) : y > 0\}$ are minimizers. For the other interior points of the form (x,y), $x \neq 0$, $y > 0$, we have the direction $(x,2y) \in \text{dom } f$,

$$\forall y > 0, x \neq 0, \quad \underline{d} f(x,y);(x,2y)-(x,y)) = -\frac{x^2}{y} < 0,$$

and, from part (ii), the points $\{(x,y) : x \neq 0 \text{ and } y > 0\}$ are not minimizers. At the boundary point $(0,0)$, we have, by definition,

$$\underline{d} f((0,0);(0,0)) = df((0,0);(0,0)) = 0.$$

On the other hand, for all $(x,y) \in \text{dom } f$ such that $(x,y) \neq (0,0)$, we have $y > 0$. Therefore, for $t > 0$, $(0,0)+t(x,y) \in \text{dom } f$, $(tx,ty) \neq (0,0)$, and $ty > 0$. The differential quotient is then given by

$$\frac{f((0,0)+t(x,y)) - f(0,0)}{t} = \frac{1}{t}\frac{(tx)^2}{ty} = \frac{x^2}{y} \quad \Rightarrow \underline{d} f((0,0);(y-0,x-0)) \geq 0.$$

As a result,

$$\forall (x,y) \in \text{dom } f, \quad \underline{d} f((0,0);(y-0,x-0)) \geq 0$$

and condition (ii) is verified in $(0,0)$. Finally, the condition of part (ii) of Theorem 3.5 yields $\text{argmin} f(\mathbb{R}^2) = \{(0,y) : y \geq 0\}$. □

4 Admissible Directions and Tangent Cones to U

4.1 Set of Admissible Directions or Half-Tangents

To obtain a necessary optimality condition for an arbitrary U, we need a local approximation of U at the points $x \in \overline{U}$. It will take the form of a set of *half-tangents* to U at boundary points of U. Indeed, consider a connected set U and a local minimizer $x \in U$ of f with respect to U: there exists a neighborhood $V(x)$ of x such that $f(y) \geq f(x)$ for all $y \in U \cap V(x)$. Each path or trajectory $t \mapsto h(t) : [0,t_0) \to \mathbb{R}^n$ contained in $U \cap V(x)$ ending

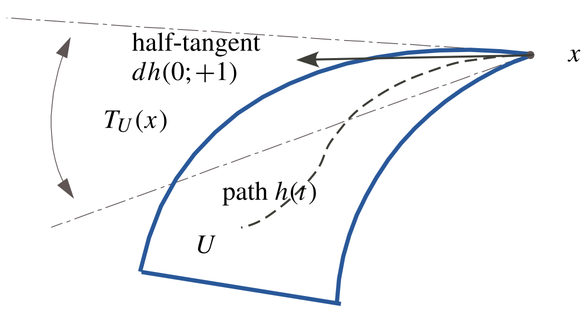


Figure 4.9. *Half-tangent $dh(0;+1)$ to the path $h(t)$ in U at the point $h(0) = x$.*

at $h(0) = x$ will verify $f(h(t)) \geq f(h(0)) = f(x)$. If f is Fréchet differentiable at $h(0) = x$ and if the path is semidifferentiable at $t = 0^+$, that is, $dh(0;+1)$ exists, then

$$0 \leq \lim_{t \searrow 0} \left. \frac{f(h(t)) - f(h(0))}{t} \right|_{t=0^+} = d(f \circ h)(0;+1) = \nabla f(x) \cdot dh(0;+1)$$

by the chain rule for the composition of two functions. Therefore, the scalar product of the gradient of f with the half-tangent $dh(0;+1)$ at $t = 0^+$ to all trajectories h in $U \cap V(x)$ is positive or null. The magnitude of the velocity along paths h can be increased or decreased by introducing the new parametrization $h_\lambda(t) = h(\lambda t)$ of the path for some $\lambda > 0$. We obtain $dh_\lambda(0;+1) = \lambda dh(0;+1)$ and hence the whole half line tangent to the path h at x. This generates a cone of half-tangents at x to the set U. The following definition extends this notion of tangent cone to arbitrary sets U (not necessarily connected).

Definition 4.1.
Let $U \subset \mathbb{R}^n$ and $x \in \overline{U}$.

(i) $h \in \mathbb{R}^n$ is an *admissible direction* for U at x (or *half-tangent* to U in x) if there exists a sequence $\{t_n > 0\}$, $t_n \searrow 0$ as $n \to \infty$, for which

$$\forall n, \quad \exists x_n \in U \text{ such that } \lim_{n \to \infty} \frac{x_n - x}{t_n} = h. \tag{4.1}$$

(ii) Denote by $T_U(x)$ the set[4] [5] [6] of all admissible directions of U at x.

The case of a point $x \notin \overline{U}$ is of no interest since the distance from x to $\overline{U}$ is strictly positive and $T_U(x)$ would be empty. □

Remark 4.1.
If there exists a function $t \mapsto x(t) : [0, t_0) \to U$, $t_0 > 0$, such that $x(0) = x$ and $dx(0;+1) = h$, then $h \in T_U(x)$. Indeed, by assumption,

$$\lim_{t \searrow 0} \frac{x(t) - x}{t} = h. \tag{4.2}$$

It is sufficient to choose the sequences $t_n = t_0/(2n)$ and $x_n = x(t_0/(2n))$. □

Example 4.1.
Definition 4.1 extends the notion of admissible direction to sets U that are not connected or have no interior. For instance, consider the set

$$U \stackrel{\text{def}}{=} \left\{ x_n \in \mathbb{R}^2 : x_n \stackrel{\text{def}}{=} \left(\frac{1}{n}, \frac{1}{n^2} \right), \forall n \in \mathbb{N}, n \geq 1 \right\}$$

---

[4] $T_U(x)$ seems to have been independently introduced in 1930 by G. Bouligand [**1**] and F. Severi [**1**] in the same volume 9 of the "Annales de la Société Polonaise de Mathématiques (Krakóv)" started in 1921 by Stanislaw Zaremba. Bouligand calls it *contingent cone* from the latin *contingere* that means to touch on all sides. In Bouligand's paper, an admissible direction is called a *half-tangent* and in M. R. Hestenes [**1**, p. 264] a *sequential tangent*. $T_U(x)$ arises in many contexts such as in Geometry where it is the tangent space at x to U or in *Viability Theory* for ordinary differential equations (see M. Nagumo [**1**] or J.-P. Aubin and A. Cellina [**1**, p. 174 and p. 180]).

[5] Georges Louis Bouligand (1889–1979).

[6] Francesco Severi (1879–1961).

that has a limit point in $x_0 = (0,0)$. Moreover, $\overline{U} = \{x_0\} \cup U$. For each $n \geq 1$, $T_U(x_n) = \{(0,0)\}$ since x_n is an isolated point of U for which only the path $h(t) = x_0$ remains in U as the size of the neighborhood goes to zero. But, for the accumulation point x_0, by choosing $t_n = \|x_n\| = \sqrt{1/n^2 + 1/n^4}$ that goes to zero, we get

$$\frac{x_n - x_0}{t_n} = \frac{\left(\frac{1}{n}, \frac{1}{n^2}\right) - (0,0)}{\sqrt{\frac{1}{n^2} + \frac{1}{n^4}}} = \frac{\left(1, \frac{1}{n}\right)}{\sqrt{1 + \frac{1}{n^2}}} \to (1,0).$$

Choosing next $t_n^\lambda = \lambda t_n$ for all $\lambda > 0$, we get a half line generated by $(1,0)$ which, by definition, belongs to $T_U(0,0)$. □

We always have $0 \in T_U(x)$ for the sequences $t_n = 1/n$ and $x_n = x$. For points in the interior $\operatorname{int} U$ of U,

$$\boxed{\forall x \in \operatorname{int} U, \quad T_U(x) = \mathbb{R}^n.} \tag{4.3}$$

Since $x \in \operatorname{int} U$, there exists $r > 0$ such that $B_r(x) \subset U$. For $0 \neq h \in \mathbb{R}^n$, choose $t_0 > 0$ such that $t_0\,|h| < r$ and, for $n \geq 1$, $t_n = t_0/n$ and $x_n = x + t_n h$. Hence,

$$|x_n - x| = t_n |h| = t_0\,|h|/n < r \quad \Rightarrow x_n \in B_r(x) \subset U$$
$$\left|\frac{x_n - x}{t_n} - h\right| = 0 \quad \Rightarrow h \in T_U(x).$$

Therefore, $T_U(x) = \mathbb{R}^n$ and, in particular, $T_{\mathbb{R}^n}(x) = \mathbb{R}^n$.

As a consequence, if $T_U(x) \neq \mathbb{R}^n$, then $x \in \partial U$, the boundary of U. However, there are examples such as Figure 4.10 where $T_U(x) = \mathbb{R}^n$ for a boundary point. Figures 4.11, 4.12, 4.13, and 4.14 display $T_U(x)$ for different choices of U and points $x \in \partial U$.

When $U = A$, an affine subspace of $\mathbb{R}^n$, there is a unique linear subspace S such that for all $x \in A$, $A = x + S$. Therefore, for all $h \in S$ and $t > 0$, $x(t) = x + th \in A$, $(x(t) - x)/t = h \in T_A(x)$, and $S \subset T_A(x)$. Conversely, for $h \in T_A(x)$, there exist $x_n \in A$ and $t_n > 0$ such that $(x_n - x)/t_n \to h$. But $x_n - x \in A - x = S$ and $1/t_n(x_n - x) \in S$. In the limit, $h \in \overline{S} = S$ and $T_A(x) \subset S$. In conclusion, for all points x in an affine subspace A of $\mathbb{R}^n$, $T_A(x) = A - x = S$.

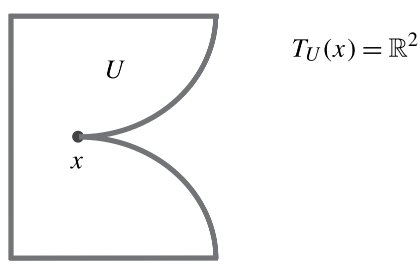


Figure 4.10. *Cusp at $x \in \partial U$: U and $T_U(x)$.*

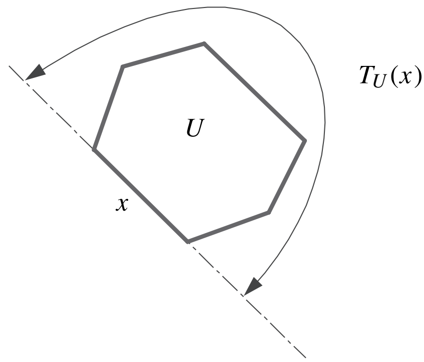


Figure 4.11. *First example: U and $T_U(x)$.*

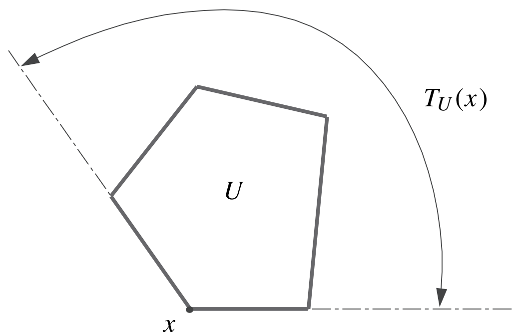


Figure 4.12. *Second example: U and $T_U(x)$.*

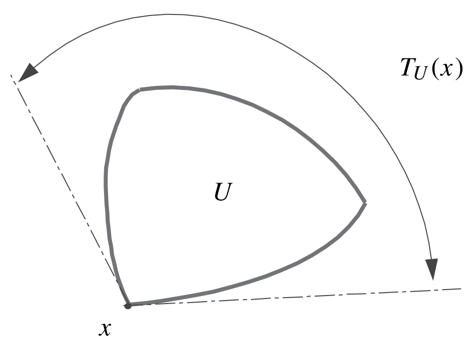


Figure 4.13. *Third example: U and $T_U(x)$.*

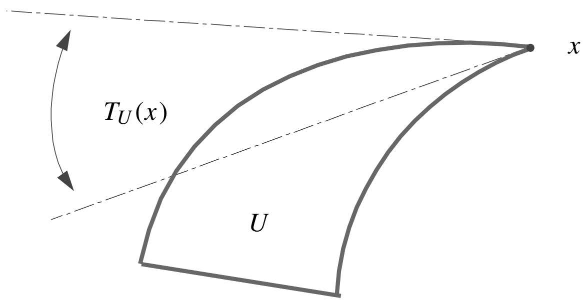


Figure 4.14. *Fourth example: U and $T_U(x)$.*

4.2 Properties of the Tangent Cones $T_U(x)$ and $S_U(x)$

We now give the properties of $T_U(x)$.

Theorem 4.1. *Let $\varnothing \neq U \subset \mathbb{R}^n$ and $x \in \overline{U}$.*

(i) *$T_U(x)$ is a closed cone in 0 and*

$$T_U(x) = \left\{ h \in \mathbb{R}^n : \liminf_{t \searrow 0} \frac{d_U(x+tv)}{t} = 0 \right\} \subset S_U(x), \tag{4.4}$$

where

$$S_U(x) \stackrel{\text{def}}{=} \text{closure}\,\{\lambda(y-x) : \forall \lambda > 0, \forall y \in U\} = \overline{\mathbb{R}^+(U-x)} \tag{4.5}$$

and $\mathbb{R}^+ \stackrel{\text{def}}{=} \{\lambda \in \mathbb{R} : \forall \lambda > 0\}$.

(ii) *$T_{\overline{U}}(x) = T_U(x)$ and $S_{\overline{U}}(x) = S_U(x)$.*

(iii) *If U is convex, then $T_U(x)$ is convex and*

$$T_U(x) = S_U(x). \tag{4.6}$$

Proof. (i) Given $\lambda > 0$ and $h \in T_U(x)$, we want to show that $\lambda h \in T_U(x)$. By definition of $T_U(x)$, there exists a sequence $\{t_n > 0\}$, $t_n \searrow 0$ as $n \to \infty$, for which

$$\forall n, \ \exists x_n \in U \ \text{and} \ \lim_{n\to\infty} \frac{x_n - x}{t_n} = h.$$

Choose $\bar{t}_n = t_n / \lambda > 0$ and $\bar{x}_n = x_n$. By construction, $\bar{x}_n \in U$ and

$$\frac{\bar{x}_n - x}{\bar{t}_n} = \frac{x_n - x}{t_n/\lambda} = \lambda \left[\frac{x_n - x}{t_n} \right].$$

By going to the limit, we get

$$\bar{t}_n \to 0 \ \text{and} \ \lim_{n\to\infty} \frac{\bar{x}_n - x}{\bar{t}_n} = \lambda \lim_{n\to\infty} \left[\frac{x_n - x}{t_n} \right] = \lambda h.$$

By definition, $\lambda h \in T_U(x)$ and $T_U(x)$ is a cone in 0.

For all $t > 0$, $y \in U$, and $h \in T_U(x)$,

$$\frac{\|y - x - th\|}{t} \geq \frac{d_U(x+th)}{t} \geq 0.$$

Since $h \in T_U(x)$, there exist sequences $t_n \searrow 0$ and $\{y_n\} \subset U$ such that $(y_n - x)/t_n \to h$. Hence, for any $\varepsilon > 0$ and $\rho > 0$, there exist t_n, $0 < t_n < \rho$, and $y_n \in U$ such that

$$\begin{aligned}
\varepsilon > \left\| \frac{y_n - x}{t_n} - h \right\| \geq \inf_{0<t<\rho} \inf_{y \in U} \left\| \frac{y - x - th}{t} \right\| = \inf_{0<t<\rho} \frac{d_U(x+th)}{t} \geq 0 \\
\Rightarrow \forall \varepsilon > 0, \forall \rho > 0, \quad 0 \leq \inf_{0<t<\rho} \frac{d_U(x+th)}{t} < \varepsilon \\
\Rightarrow \forall \varepsilon > 0, \quad 0 \leq \liminf_{t \searrow 0} \frac{d_U(x+th)}{t} = \sup_{\rho>0} \inf_{0<t<\rho} \frac{d_U(x+th)}{t} < \varepsilon
\end{aligned}$$

and the lim inf is 0. Conversely, consider h such that

$$0 = \liminf_{t \searrow 0} \frac{d_U(x+th)}{t} = \lim_{\rho \searrow 0} \inf_{0<t<\rho} \inf_{y \in U} \left\| \frac{y - x - th}{t} \right\|$$

and construct the following sequence: let $0 < \rho_1 < 1$ be such that

$$\inf_{\substack{0<t<\rho_1 \\ y \in U}} \left\| \frac{y - x - th}{t} \right\| < \frac{1}{2^2}.$$

At step n, let $0 < \rho_n < 1/n$ be such that

$$\inf_{\substack{0<t<\rho_n \\ y \in U}} \left\| \frac{y - x - th}{t} \right\| < \frac{1}{2^{n+1}}.$$

Associate with each n the points t_n, $0 < t_n < \rho_n$, and $y_n \in U$ such that

$$\begin{aligned}
\inf_{\substack{0<t<\rho_n \\ y \in U}} \left\| \frac{y - x - th}{t} \right\| \leq \left\| \frac{y_n - x - t_n h}{t_n} \right\| < \inf_{\substack{0<t<\rho_n \\ y \in U}} \left\| \frac{y - x - th}{t} \right\| + \frac{1}{2^{n+1}} \\
\Rightarrow \left\| \frac{y_n - x}{t_n} - h \right\| < \frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} = \frac{1}{2^n} \quad \text{and} \quad 0 < t_n < 1/n.
\end{aligned}$$

So, by definition, $h \in T_U(x)$.

To show that $T_U(x)$ is closed, consider h in the closure of $T_U(x)$ and show that $h \in T_U(x)$. Let $\{h_n\} \subset T_U(x)$ be a sequence such that $h_n \to h$ and

$$\liminf_{t \searrow 0} \frac{d_U(x+th_n)}{t} = 0.$$

But, since d_U is Lipschitz of constant 1,

$$\begin{aligned}
0 \leq \frac{d_U(x+th)}{t} &\leq \frac{d_U(x+th_n)}{t} + \frac{\|d_U(x+th) - d_U(x+th_n)\|}{t} \\
&\leq \frac{d_U(x+th_n)}{t} + \|h - h_n\| \\
\Rightarrow 0 \leq \liminf_{t \searrow 0} \frac{d_U(x+th)}{t} &\leq \liminf_{t \searrow 0} \frac{d_U(x+th_n)}{t} + \|h - h_n\| = \|h - h_n\|.
\end{aligned}$$

So, as n goes to infinity, the liminf is 0 and $h \in T_U(x)$.

Property (4.4) is true by definition of an element $h \in T_U(x)$: there exists a sequence $\{t_m > 0 : m \geq 1\}$ such that $t_m \to 0$ as m goes to infinity and, for each m, there exists a point $x_m \in U$ such that

$$\frac{x_m - x}{t_m} \to h \quad \Rightarrow \frac{1}{t_m}(x_m - x) \in \mathbb{R}^+(U - x) \quad \Rightarrow h \in S_U(x) = \overline{\mathbb{R}^+(U - x)}$$

and necessarily $T_U(x) \subset S_U(x)$.

(ii) By definition, $U \subset \overline{U}$ implies $T_U(x) \subset T_{\overline{U}}(x)$ and $S_U(x) \subset S_{\overline{U}}(x)$. Let $h \in T_{\overline{U}}(x)$. For each $m \geq 1$, there exist sequences t_m, $0 < t_m < 1/m$, and $y_m \in \overline{U}$ such that

$$\left\| \frac{y_m - x}{t_m} - h \right\| < 1/(2m).$$

Since $y_m \in \overline{U}$, there exists $x_m \in U$ such that

$$\|x_m - y_m\| < t_m/(2m).$$

So we have constructed sequences $\{t_m\}$, $0 < t_m < 1/m$, and $\{x_m\} \subset U$ such that

$$\forall m \geq 1, \quad \left\| \frac{x_m - x}{t_m} - h \right\| \leq \left\| \frac{y_m - x}{t_m} - h \right\| + \left\| \frac{y_m - x_m}{t_m} \right\| < 1/m.$$

Therefore, $h \in T_U(x)$ and $T_{\overline{U}}(x) \subset T_U(x)$. For the other cone, it is sufficient to note that $\mathbb{R}^+(\overline{U} - x) \subset \mathbb{R}^+\overline{(U - x)} \subset \overline{\mathbb{R}^+(U - x)}$.

(iii) From Theorem 7.1(i) of Chapter 2, the closure of a convex subset of $\mathbb{R}^n$ is convex. In the light of part (i), it is now sufficient to show that $S_U(x) \subset T_U(x)$ when U is convex for an arbitrary $x \in \overline{U}$ (convex by Theorem 7.1(i) of Chapter 2). We proceed with the following construction. For each $y \in U$ and $\lambda > 0$, choose $t_0 = \min\{1, \lambda^{-1}\}$ and for each $t \in]0, t_0[$, $x(t) = x + t\lambda(y - x)$. Since $0 < t\lambda < 1$ and $\overline{U}$ is convex, $x(t) = (1 - t\lambda)x + t\lambda y \in \overline{U}$. Finally,

$$\frac{x(t) - x}{t} = \lambda(y - x) \quad \text{(independent of } t\text{)}.$$

From Remark 4.1, for all $\lambda > 0$ and $y \in U$, $\lambda(y - x) \in T_{\overline{U}}(x)$. Hence,

$$\mathbb{R}^+(U - x) = \{\lambda(y - x) : \lambda > 0, y \in U\} \subset T_{\overline{U}}(x) = T_U(x)$$

from part (ii). By combining this result with the one of part (i), we get

$$\mathbb{R}^+(U - x) \subset T_U(x) \subset \overline{\mathbb{R}^+(U - x)} = S_U(x)$$

and since, always from part (i), $T_U(x)$ is closed, we get

$$\overline{\mathbb{R}^+(U - x)} \subset T_U(x) \subset \overline{\mathbb{R}^+(U - x)} \quad \Rightarrow T_U(x) = S_U(x).$$

We also know from part (i) that $T_U(x)$ is a closed cone in 0. It remains to show that it is convex. This amounts to prove that $C = \mathbb{R}^+(U - x)$ is convex, since $S_U(x)$ is the closure of C and the closure of a convex is convex, from Theorem 7.1(i) of Chapter 2.

Going back to Theorem 3.1, C satisfies the first condition. Indeed, for all $\lambda > 0$ and $s \in C$, there exist $\alpha > 0$ and $y \in U$ such that $s = \alpha(y-x)$. For $\lambda > 0$, this implies $\lambda s = \lambda(\alpha(y-x)) = (\lambda\alpha)(y-x) \in \mathbb{R}^+(U-x) = C$. As for the second condition, we must prove that for all y_1 and y_2 in C, the sum belongs to C. There exist $\lambda_1 > 0$, $\lambda_2 > 0$, x_1, and x_2 in U such that $y_1 = \lambda_1(x_1-x)$ and $y_2 = \lambda_2(x_2-x)$. Since $\lambda_1+\lambda_2 > 0$, then, by convexity of U,

$$\begin{aligned} y_1+y_2 &= \lambda_1(x_1-x)+\lambda_2(x_2-x) \\ &= (\lambda_1+\lambda_2)\left\{\left[\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\right]-x\right\} \in \mathbb{R}^+(U-x). \end{aligned}$$

From Theorem 3.1 we have that C is convex and that its closure $S_U(x)$ is also convex. □

4.3 ▶ Clarke's and Other Tangent Cones

Bouligand's cone is not the only approximation cone that can be used in the context of a necessary optimality condition. A fairly complete description of such cones can be found in the book of J.-P. AUBIN and H. FRANKOWSKA [**1**, Chap. 4]. For instance, the *Clarke tangent cone*

$$C_U(x) \overset{\text{def}}{=} \left\{ h \in \mathbb{R}^n : \limsup_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th)}{t} = 0 \right\}, \tag{4.7}$$

the *intermediate or adjacent cone*

$$T_U^\flat(x) \overset{\text{def}}{=} \left\{ h \in \mathbb{R}^n : \lim_{t \searrow 0} \frac{d_U(x+th)}{t} = 0 \right\}, \tag{4.8}$$

and the *Dubovitskiĭ–Miljutin cone*[7]

$$\begin{aligned} D_U(x) &\overset{\text{def}}{=} \left\{ h \in \mathbb{R}^n : \exists B_\varepsilon(h), \varepsilon > 0, \forall k \in B_\varepsilon(h) \text{ and } \forall 0 < t < \varepsilon, \quad x+tk \in U \right\} \\ &\Rightarrow \forall x \in \operatorname{int} U, \quad D_U(x) = \mathbb{R}^n \quad \text{and} \quad \forall x \in \partial U, \quad D_U(x) = \complement T_{\complement U}(x). \end{aligned}$$

So $D_U(x)$ is an open cone and $D_U(x) \subset T_U^\flat(x) \subset T_U(x)$.

Theorem 4.2. *Let $U \subset \mathbb{R}^n$ be nonempty and $x \in \overline{U}$.*

(i) *$C_U(x)$, $T_U^\flat(x)$, and $T_U(x)$ are closed cones in* 0 *and*

$$C_U(x) \subset T_U^\flat(x) \subset T_U(x). \tag{4.9}$$

Moreover, $C_{\overline{U}}(x) = C_U(x)$, $T_{\overline{U}}^\flat(x) = T_U^\flat(x)$, and $T_{\overline{U}}(x) = T_U(x)$.

(ii) *$C_U(x)$ is a closed convex cone in* 0.

(iii) *If U is convex, $C_U(x) = T_U^\flat(x) = T_U(x)$.*

[7]Cf. for instance, A. JA. DUBOVITSKIĬ and A. A. MILJUTIN [**1**] and A. A. MILJUTIN [**1**].

Proof. (i) By definition, $T_U^\flat(x) \subset T_U(x)$. Let $h \in C_U(x)$. Since $d_U(x') = 0$ for $x' \in \overline{U}$,

$$\limsup_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th) - d_U(x')}{t} = \limsup_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th)}{t} = 0$$

and, since the function d_U is positive,

$$0 = \limsup_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th) - d_U(x')}{t} \geq \liminf_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th) - d_U(x')}{t} \geq 0$$
$$\Rightarrow \lim_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th) - d_U(x')}{t} = 0.$$

The limsup can be replaced by a lim in the definition of $C_U(x)$. Moreover, since $x \in \overline{U}$,

$$0 = \limsup_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th) - d_U(x')}{t} \geq \limsup_{t \searrow 0} \frac{d_U(x+th) - d_U(x)}{t}$$
$$\liminf_{t \searrow 0} \frac{d_U(x+th) - d_U(x)}{t} \geq \liminf_{\substack{t \searrow 0 \\ x' \in \overline{U} \to x}} \frac{d_U(x'+th) - d_U(x')}{t} = 0$$
$$\Rightarrow dd_U(x;h) = \lim_{t \searrow 0} \frac{d_U(x+th) - d_U(x)}{t} = 0.$$

We have already proved that $T_U(x)$ is a closed cone in 0. The proof is essentially the same for $T_U^\flat(x)$ and $C_U(x)$. They are cones in 0 by positive homogeneity of the lim and limsup and they are closed since d_U is uniformly Lipschitzian.

(ii) For $h, k \in C_U(x)$ and $\lambda \in]0,1[$, and $x' \to x$, $t \searrow 0$,

$$\begin{aligned} 0 &\leq \frac{d_U(x'+t(\lambda h + (1-\lambda)k)) - d_U(x')}{t} \\ &\leq \frac{d_U(x'+t\lambda h + t(1-\lambda)k) - d_U(x'+t\lambda h)}{t} + \frac{d_U(x'+t\lambda h) - d_U(x')}{t}. \end{aligned}$$

Now since $t\lambda \searrow 0$ and $t(1-\lambda) \searrow 0$,

$$\begin{aligned} 0 &\leq \frac{d_U(x'+t(\lambda h + (1-\lambda)k)) - d_U(x')}{t} \\ &\leq (1-\lambda)\frac{d_U(x'+t\lambda h + t(1-\lambda)k) - d_U(x'+t\lambda h)}{t(1-\lambda)} + \lambda\frac{d_U(x'+t\lambda h) - d_U(x')}{t\lambda}. \end{aligned}$$

After making the changes of variables take the limsup,

$$\begin{aligned}
\Rightarrow 0 &\le \limsup_{\substack{t\searrow 0\\ x'\in\overline{U}\to x}} \frac{d_U(x'+t(\lambda h+(1-\lambda)k))-d_U(x')}{t}\\
&\le (1-\lambda)\limsup_{\substack{t\searrow 0\\ x'\in\overline{U}\to x}} \frac{d_U(x'+t\lambda h+t(1-\lambda)k)-d_U(x'+t\lambda h)}{t(1-\lambda)}\\
&\quad+\lambda\limsup_{\substack{t\searrow 0\\ x'\in\overline{U}\to x}} \frac{d_U(x'+t\lambda h)-d_U(x')}{t\lambda}.
\end{aligned}$$

Going to any $x'\to x$ on the right-hand side,

$$\begin{aligned}
0 &\le \limsup_{\substack{t\searrow 0\\ x'\in\overline{U}\to x}} \frac{d_U(x'+t(\lambda h+(1-\lambda)k))-d_U(x')}{t}\\
&\le (1-\lambda)\limsup_{\substack{t\searrow 0\\ x'\in\overline{U}\to x}} \frac{d_U(x'+tk)-d_U(x')}{t}+\lambda\limsup_{\substack{t\searrow 0\\ x'\in\overline{U}\to x}} \frac{d_U(x'+th)-d_U(x')}{t}=0.
\end{aligned}$$

Therefore, $C_U(x)$ is convex.

(iii) In view of Theorem 4.1(i), it is sufficient to prove that $\mathbb{R}^+(U-x)\subset C_U(x)$. From this and the fact that $C_U(x)$ is closed,

$$\overline{\mathbb{R}^+(U-x)}\subset C_U(x)\subset S_U(x)=\overline{\mathbb{R}^+(U-x)}.$$

Given $h\in\mathbb{R}^+(U-x)$, there exist $y\in U$ and $\tau>0$ such that $h=(y-x)/\tau$. Then $y=x+\tau h\in\overline{U}$ and, since $\overline{U}$ is convex,

$$\forall t\in[0,\tau],\quad x+th=\left(1-\frac{t}{\tau}\right)x+\frac{t}{\tau}(x+th)\in\overline{U}.$$

For any sequences $\{x_n\}\subset\overline{U}$, $x_n\to x$ and $t_n\searrow 0$,

$$x_n+t_nh=\underbrace{\left(1-\frac{t_n}{\tau}\right)x_n+\frac{t_n}{\tau}\,y}_{y_n\in\overline{U}}+t_n\frac{x_n-x}{\tau}.$$

Using the fact that d_U is Lipschitzian and that $d_U(y_n)=0$,

$$0\le\frac{d_U(x_n+t_nh)}{t_n}\le\frac{d_U(y_n)}{t_n}+\left\|\frac{x_n-x}{\tau}\right\|=\left\|\frac{x_n-x}{\tau}\right\|\to 0$$

and $h\in C_U(x)$. □

5 Orthogonality, Transposition, and Dual Cones

5.1 Orthogonality and Transposition

Recall the definition of the orthogonal set to a subset U of $\mathbb{R}^n$ (see (8.4) in Chapter 2):

$$U^\perp \stackrel{\text{def}}{=} \left\{y \in \mathbb{R}^n \text{ such that } y \cdot x = 0,\ \forall x \in U\right\}.$$

We have a first set of properties.

Theorem 5.1. *Let $U \subset \mathbb{R}^n$.*

(i) *$U^\perp$ is a linear subspace of $\mathbb{R}^n$ and $U \subset (U^\perp)^\perp$.*

(ii) *For any pair $U_1 \subset U_2$, we have $U_2^\perp \subset U_1^\perp$.*

(iii) $U^\perp = (\operatorname{span} U)^\perp$.[8]

(iv) *If $U = S$, a nonempty linear subspace of $\mathbb{R}^n$, then*

$$(S^\perp)^\perp = S.$$

(v) *For any $U \subset \mathbb{R}^n$, $(U^\perp)^\perp = \operatorname{span} U$.*

Proof. (i) Indeed, for all $x \in U$ and for all y_1 and y_2 in $U^\perp$ and α, β in $\mathbb{R}$,

$$(\alpha y_1 + \beta y_2) \cdot x = \alpha y_1 \cdot x + \beta y_2 \cdot x = 0$$

and $\alpha y_1 + \beta y_2 \in U^\perp$. Let $x \in U$. By definition of $U^\perp$, for all $y \in U^\perp$, $y \cdot x = 0$. But, by definition of $(U^\perp)^\perp$, $x \in (U^\perp)^\perp$ and $U \subset (U^\perp)^\perp$.

(ii) For all $h \in U_2^\perp$, we have $h \cdot x = 0$ for all $x \in U_2$. Since $U_1 \subset U_2$, we get $h \in U_1^\perp$.

(iii) Since $U \subset \operatorname{span} U$, we have from (ii) $(\operatorname{span} U)^\perp \subset U^\perp$. Conversely, consider a linear combination $\alpha x + \beta y$ of two elements x, y of U. We have

$$\forall h \in U^\perp, \quad h \cdot (\alpha x + \beta y) = \alpha\, h \cdot x + \beta\, h \cdot y = \alpha 0 + \beta\, 0 = 0$$

and $h \in (\operatorname{span} U)^\perp$.

(iv) From part (i) we know that $S \subset (S^\perp)^\perp$. In general, the converse is not true, but it is true for a linear subspace. Given an arbitrary $z \in (S^\perp)^\perp$, consider the following minimization problem:

$$\inf_{x \in S} f(x), \quad f(x) \stackrel{\text{def}}{=} \frac{1}{2}\|x - z\|^2.$$

This problem has a solution $x_0 \in S$. Indeed, f has the growth property at infinity (Theorem 5.5, Chapter 2). So it has a compact lower section in S (Theorem 5.4, Chapter 2). Finally, f is also continuous in $\mathbb{R}^n$ and S is closed in $\mathbb{R}^n$. Then, by Theorem 3.1, we have the existence of a minimizer $x_0 \in S$. We now prove that $x_0 = z$. Since f is convex (its Hessian matrix is the identity matrix) and of class $C^{(1)}$ on $\mathbb{R}^n$. By Corollary 1(i) to Theorem 3.2, the minimizer x_0 is completely characterized by (3.7):

$$0 = \nabla f(x_0) \cdot h = (x_0 - z) \cdot h, \quad \forall h \in S. \tag{5.1}$$

[8] Given $U \subset \mathbb{R}^n$, $\operatorname{span} U$ is the smallest linear subspace containing U.

Therefore, $x_0 - z \in S^\perp$. Since $x_0 \in S$ and, by assumption, $z \in (S^\perp)^\perp$, we get

$$\underbrace{(x_0 - z)}_{\in S^\perp} \cdot \underset{\in S}{x_0} = 0 \text{ and } \underbrace{(x_0 - z)}_{\in S^\perp} \cdot \underset{\in (S^\perp)^\perp}{z} = 0 \quad \Rightarrow \|x_0 - z\|^2 = 0.$$

Therefore, $z = x_0 \in S$ and $(S^\perp)^\perp \subset S$. This completes the proof.

(v) follows from parts (iii) and (iv). □

Consider the linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ (that can be represented by an $m \times n$ matrix). Denote by Im A and Ker A the *image* and the *kernel* of A

$$\text{Im } A \overset{\text{def}}{=} \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ such that } Ax = y\} \subset \mathbb{R}^m \tag{5.2}$$

$$\text{Ker } A \overset{\text{def}}{=} \{x \in \mathbb{R}^n : Ax = 0\} \subset \mathbb{R}^n. \tag{5.3}$$

The set Ker A is closed by continuity of A and a linear subspace; the set Im A is closed by linearity and the fact that $\mathbb{R}^m$ is finite dimensional.

Define the *transposed* or *adjoint* map $A^\top : \mathbb{R}^m \to \mathbb{R}^n$ of $A : \mathbb{R}^n \to \mathbb{R}^m$ by the following process. For all $y \in \mathbb{R}^m$, the map $x \mapsto y \cdot Ax : \mathbb{R}^n \to \mathbb{R}$ is linear. Therefore, there exists a unique vector $a(y) \in \mathbb{R}^n$ such that

$$\forall x \in \mathbb{R}^n, \quad a(y) \cdot x = y \cdot Ax$$

and this induces a map

$$y \mapsto A^\top y \overset{\text{def}}{=} a(y) : \mathbb{R}^m \to \mathbb{R}^n.$$

that is linear. Indeed, for all α, β in $\mathbb{R}$ and y_1, y_2 in $\mathbb{R}^m$, we have

$$\begin{aligned}\forall x \in \mathbb{R}^n, \quad a(\alpha y_1 + \beta y_2) \cdot x = (\alpha y_1 + \beta y_2) \cdot Ax &= \alpha\, y_1 \cdot Ax + \beta\, y_2 \cdot Ax \\ &= \alpha\, a(y_1) \cdot x + \beta\, a(y_2) \cdot x \\ &= [\alpha\, a(y_1) + \beta\, a(y_2)] \cdot x \\ \Rightarrow a(\alpha y_1 + \beta y_2) = \alpha\, a(y_1) + \beta\, a(y_2).&\end{aligned}$$

By construction, $A^\top$ verifies the identity

$$\forall x \in \mathbb{R}^n, y \in \mathbb{R}^m, \quad y \cdot Ax = A^\top y \cdot x. \tag{5.4}$$

If $A_{ij} = Ae^n_i \cdot e^m_j$ is the $n \times m$ matrix associated with A for the canonical orthonormal bases $\{e^n_i : 1 \le i \le n\}$ and $\{e^m_j : 1 \le j \le m\}$ of $\mathbb{R}^n$ and $\mathbb{R}^m$, then the $m \times n$ matrix associated with $A^\top$ is given by $(A^\top)_{ij} = A_{ji}$.

Theorem 5.2. *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be linear.*

(i) *Then*

$$(A^\top)^\top = A \tag{5.5}$$

$$[\text{Im } A]^\perp = \text{Ker}(A^\top) \tag{5.6}$$

$$[\text{Ker } A]^\perp = \text{Im}(A^\top). \tag{5.7}$$

(ii) *The image AU of a (closed) convex set U is (closed) convex; the image AU of a (closed) cone U in 0 is a (closed) cone in 0.*

Proof. (i) (a) For all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$,

$$Ax \cdot y = x \cdot A^\top y = (A^\top)^\top x \cdot y.$$

Therefore,

$$\forall y \in \mathbb{R}^m, \quad Ax \cdot y = (A^\top)^\top x \cdot y, \quad \Rightarrow \forall x \in \mathbb{R}^n, \quad Ax = (A^\top)^\top x, \quad \Rightarrow A = (A^\top)^\top.$$

(b) By definition,

$$[\text{Im } A]^\perp = \{y \in \mathbb{R}^m : Ax \cdot y = 0, \ \forall x \in \mathbb{R}^n\}.$$

But

$$0 = Ax \cdot y = x \cdot A^\top y, \ \forall x \in \mathbb{R}^n \quad \Rightarrow A^\top y = 0 \quad \Rightarrow y \in \text{Ker } A^\top.$$

Conversely,

$$y \in \text{Ker } A^\top \quad \Rightarrow A^\top y = 0 \quad \Rightarrow \forall x \in \mathbb{R}^n, \ x \cdot A^\top y = 0$$

and

$$0 = x \cdot A^\top y = Ax \cdot y, \ \forall x \in \mathbb{R}^n \quad \Rightarrow y \in [\text{Im } A]^\perp.$$

(c) From (5.6) by replacing A by $A^\top$ and by using identity (5.5),

$$[\text{Im } A^\top]^\perp = \text{Ker}\,(A^\top)^\top = \text{Ker } A \quad \Rightarrow ([\text{Im } A^\top]^\perp)^\perp = (\text{Ker } A)^\perp.$$

Finally, by Theorem 5.1, since Im $A^\top$ is a linear subspace $(\text{Im } A^\top)^{\perp\perp} = \text{Im } A^\top$.

(ii) The image of a convex (cone) is a convex (cone) by linearity. For the closure, introduce the *quotient space* $\mathbb{R}^n/\,\text{Ker}\,A$ of *equivalence classes* $[x] = x + \text{Ker}\,A$ for each $x \in \mathbb{R}^n$. Since A is linear, $A[x] = A(x + \text{Ker}\,A) = \{Ax\}$ and the map

$$[x] \mapsto \tilde{A}[x] \overset{\text{def}}{=} Ax : \mathbb{R}^n/\,\text{Ker}\,A \to \text{Im}\,A$$

is well defined, linear, and bijective. Therefore, $\tilde{A}$ and $\tilde{A}^{-1}$ are continuous and the image of any closed set $[U]$ is closed. Since $AU = \tilde{A}[U]$, AU is closed. □

5.2 Dual Cones

We now weaken the notion of orthogonality by requiring only the positivity instead of the equality to zero.

Definition 5.1.
Let $U \subset \mathbb{R}^n$.

(i) Associate with U the set

$$U^* \overset{\text{def}}{=} \left\{ y \in \mathbb{R}^n : y \cdot x \geq 0, \ \forall x \in U \right\}. \tag{5.8}$$

If $U = \varnothing$, then $U^* = \mathbb{R}^n$. The set U^* will be called the *dual* of U.

(ii) Define the *bidual* U^{**} of U as $(U^*)^*$. □

The set U^* turns out to be a closed convex cone in 0. For this reason, we shall refer to U^* as the *dual cone* of U even if U is not a cone in 0. The dual cone U^{**} of U^* will be called the *bidual* cone of U.

Theorem 5.3. *Given $U \subset \mathbb{R}^n$, the set U^* is a closed convex cone in* 0.

Proof. We first show that for all $\lambda > 0$ and $h \in U^*$, $\lambda h \in U^*$. Indeed, by definition,

$$h \cdot x \geq 0, \ \ \forall x \in U \ \text{ and } \ \lambda > 0 \quad \Rightarrow (\lambda h) \cdot x = \lambda h \cdot x \geq 0.$$

Therefore, U^* is a cone in 0. To prove the convexity, use Theorem 3.1. It is sufficient to check that (3.3) is verified. For all h_1 and h_2 in U^*,

$$(h_1 + h_2) \cdot x = h_1 \cdot x + h_2 \cdot x \geq 0, \ \ \forall x \in U$$

since each term is positive for all $x \in U$. Finally, to prove that U^* is closed, we show that all limit points h of U^* belong to U^*. Indeed, if

$$\exists \{h_n\} \subset U^* \ \text{ such that } \ h_n \to h,$$

then, by continuity of the inner product, for all x in U,

$$0 \leq h_n \cdot x \to h \cdot x \quad \Rightarrow \forall x \in U, \quad h \cdot x \geq 0.$$

This proves that U^* is closed and completes the proof. □

Example 5.1.
Let $U = \mathbb{R}^n$. By definition, $h \in (\mathbb{R}^n)^*$ if

$$\forall x \in \mathbb{R}^n, \quad h \cdot x \geq 0 \quad \Rightarrow h \cdot (\pm x) \geq 0 \quad \Rightarrow \pm h \cdot x \geq 0 \quad \Rightarrow h \cdot x = 0.$$

Since this is verified for all $x \in \mathbb{R}^n$, $h = 0$ and $(\mathbb{R}^n)^* = \{0\}$. □

Example 5.2.
Let $U = B_r(0)$ for $r > 0$. By definition, $h \in (B_r(0))^*$ if

$$\forall x \in B_r(0), \quad h \cdot x \geq 0 \quad \Rightarrow h \cdot (\pm x) \geq 0 \quad \Rightarrow \pm h \cdot x \geq 0 \quad \Rightarrow h \cdot x = 0.$$

Moreover, for all $0 \neq x \in \mathbb{R}^n$,

$$\frac{r}{2} \frac{x}{\|x\|} \in B_r(0) \text{ and } h \cdot \left(\frac{r}{2} \frac{x}{\|x\|} \right) = 0 \quad \Rightarrow \forall x \in \mathbb{R}^n, \quad h \cdot x = 0$$

and necessarily $h = 0$ and $(B_r(0))^* = \{0\}$. □

Example 5.3.
Let $U = \{(x, y) : y \geq 0\}$. By definition, $h = (x^*, y^*) \in U^*$ if

$$\begin{gathered} \forall x \in \mathbb{R}, \forall y \geq 0, \quad x^* x + y^* y \geq 0 \\ \Rightarrow \forall x \in \mathbb{R}, \quad x^* x \geq 0 \text{ and } \forall y \geq 0, \quad y^* y \geq 0 \quad \Rightarrow x^* = 0 \text{ and } y^* \geq 0. \end{gathered}$$

Conversely, if $x^* = 0$ and $y^* \geq 0$, then for all $x \in \mathbb{R}$ and all $y \geq 0$,

$$x^* x + y^* y = y^* y \geq 0 \quad \Rightarrow (x^*, y^*) \in U^*$$

and $U^* = \{(x^*, y^*) : y^* \geq 0 \text{ and } x^* = 0\}$. □

Example 5.4.
Let $U = \{(x,y) : x \geq 0 \text{ and } y \geq 0\}$. By definition, $(x^*, y^*) \in U^*$ if

$$\begin{gathered}\forall x \geq 0, \forall y \geq 0, \quad x^* x + y^* y \geq 0 \\ \Rightarrow \forall x \geq 0, \quad x^* x \geq 0 \text{ and } \forall y \geq 0, \quad y^* y \geq 0 \quad \Rightarrow x^* \geq 0 \text{ and } y^* \geq 0.\end{gathered}$$

Conversely, if $x^* \geq 0$ and $y^* \geq 0$, then for all $x \geq 0$ and all $y \geq 0$,

$$x^* x + y^* y \geq 0 \quad \Rightarrow (x^*, y^*) \in U^*$$

and $U^* = \{(x^*, y^*) : y^* \geq 0 \text{ and } x^* \geq 0\} = U$. □

We shall need the following properties.

Theorem 5.4. (i) *For any subsets* U_1 *and* U_2 *of* $\mathbb{R}^n$*,*

$$U_1 \subset U_2 \quad \Rightarrow U_2^* \subset U_1^*. \tag{5.9}$$

(ii) *Given* $U \subset \mathbb{R}^n$*,*

$$(\overline{U})^* = U^*, \quad (\mathbb{R}^+ U)^* = U^*, \quad (\text{co}\, U)^* = U^*, \quad (\overline{\mathbb{R}^+ \text{co}\, U})^* = U^*, \tag{5.10}$$

where co U *is the* convex hull *of* U*:*

$$\text{co}\, U \stackrel{\text{def}}{=} \{\lambda x + (1-\lambda) y : \forall x, y \in U, \lambda \in [0,1]\}. \tag{5.11}$$

Remark 5.1.
Given $U \subset \mathbb{R}^n$ such that $0 \in \text{int}\, U$, then $U^* = \{0\}$. Indeed, there exists a ball $B_r(0) \subset U$. Thence, $U^* \subset B_r(0)^* = \{0\}$ by Example 5.2. □

Proof of Theorem 5.4. (i) For all $h \in U_2^*$,

$$h \cdot x \geq 0, \ \forall x \in U_2$$

and since $U_1 \subset U_2$,

$$h \cdot x \geq 0, \ \forall x \in U_1 \quad \Rightarrow \ h \in U_1^*.$$

(ii) (a) ($(\overline{U})^* = U^*$) From part (i), $U \subset \overline{U}$ implies $(\overline{U})^* \subset U^*$. If $h \in U^*$ and $x \in \overline{U}$, then there exists a sequence $\{x_n\} \subset U$ such that $x_n \to x$. Since $h \in U^*$, then

$$h \cdot x_n \geq 0, \quad \forall n,$$

and, by continuity of the scalar product, by going to the limit,

$$h \cdot x \geq 0 \quad \Rightarrow \forall h \in U^*, \ h \cdot x \geq 0, \ \forall x \in \overline{U} \quad \Rightarrow h \in \overline{U}^*.$$

(b) ($U^* = (\mathbb{R}^+ U)^*$) From part (i), $U \subset \mathbb{R}^+ U$ implies $(\mathbb{R}^+ U)^* \subset U^*$. Given $h \in U^*$, for all $x \in U$ and $\lambda > 0$,

$$h \cdot (\lambda x) = \lambda h \cdot x \geq 0 \quad \Rightarrow h \in (\mathbb{R}^+ U)^*.$$

(c) ($U^* = (\text{co } U)^*$) From part (i), $U \subset \text{co } U$ implies $(\text{co } U)^* \subset U^*$. Given $x^* \in U^*$, for all x and y in U and $\lambda \in [0,1]$,

$$x^* \cdot (\lambda x + (1-\lambda)y) = \lambda x^* \cdot x + (1-\lambda) x^* \cdot y \geq 0 \quad \Rightarrow x^* \in (\text{co } U)^*.$$

From the first three identities, we get the fourth one

$$(\overline{\mathbb{R}^+ \text{co } U})^* = (\mathbb{R}^+ \text{co } U)^* = (\text{co } U)^* = U^*. \qquad \square$$

The duality does not distinguish the subset U of $\mathbb{R}^n$ from the closed convex cone in 0 generated by U.

Theorem 5.5. (i) *For any $U \subset \mathbb{R}^n$, $U \subset (U^*)^*$ and if $U = C$, a nonempty closed convex cone in* 0*, then*

$$C = (C^*)^*. \tag{5.12}$$

For all $U \subset \mathbb{R}^n$,

$$(U^*)^* = \overline{\mathbb{R}^+ \text{co } U}. \tag{5.13}$$

(ii) *Let $U_1 \subset \mathbb{R}^n$ and let U_2 be a closed convex cone in* 0 *in $\mathbb{R}^n$. Then*

$$U_1 \subset U_2 \iff U_2^* \subset U_1^*. \tag{5.14}$$

Proof. (i) If $x \in U$, then by definition of U^*,

$$\forall x^* \in U^*, \; x^* \cdot x \geq 0 \quad \Rightarrow x \in (U^*)^* \quad \Rightarrow U \subset (U^*)^*.$$

Now assume that $U = C$, a closed convex cone in 0. Given $z \in (C^*)^*$, consider the following minimization problem:

$$\inf_{x \in C} f(x), \quad f(x) \stackrel{\text{def}}{=} \frac{1}{2}\|x - z\|^2.$$

This problem has a unique solution $x_0 \in C$. Indeed, f is continuous and C is closed and convex. Moreover, f has the growth property at infinity (Theorem 5.5, Chapter 2). So it has a bounded lower section in C (Theorem 5.4, Chapter 2). Then by Theorem 5.3 of Chapter 2, we have the existence of a minimizer $x_0 \in C$. Since f is of class C^2 on $\mathbb{R}^n$ and its Hessian matrix is the identity matrix, f is strictly convex and $x_0 \in C$ is unique. Moreover, since C is a closed convex cone in 0, the minimizer $x_0 \in C$ is completely characterized by (3.10) of Corollary 1(iii) to Theorem 3.2:

$$\nabla f(x_0) \cdot x_0 = (x_0 - z) \cdot x_0 = 0 \quad \text{and} \quad \forall x \in C, \nabla f(x_0) \cdot x = (x_0 - z) \cdot x \geq 0. \tag{5.15}$$

We now show that $x_0 = z$. By definition of C^*, the second inequality in (5.15) implies that

$$(x_0 - z) \in C^*. \tag{5.16}$$

On the other hand, from the first equality in (5.15),

$$\|x_0 - z\|^2 = (x_0 - z) \cdot x_0 - (x_0 - z) \cdot z = -(x_0 - z) \cdot z \tag{5.17}$$

and since $z \in (C^*)^*$ and $(x_0 - z) \in C^*$,

$$(x_0 - z) \cdot z \geq 0 \quad \Rightarrow \|x_0 - z\|^2 \leq 0 \quad \Rightarrow z = x_0 \in C. \tag{5.18}$$

As a result, $(C^*)^* \subset C$ and this completes the first part of the proof of (i).

Finally, from the last identity in (5.10) of Theorem 5.4, $(\overline{\mathbb{R}^+ \operatorname{co} U})^* = U^*$ and since $\overline{\mathbb{R}^+ \operatorname{co} U}$ is a closed convex cone in 0, $(U^*)^* = ((\overline{\mathbb{R}^+ \operatorname{co} U})^*)^* = \overline{\mathbb{R}^+ \operatorname{co} U}$.

(ii) This is a consequence of (i) since $(U_1^*)^* \supset U_1$, $(U_2^*)^* = U_2$, and

$$U_1 \subset U_2 \quad \Rightarrow U_2^* \subset U_1^* \quad \Rightarrow U_1 \subset (U_1^*)^* \subset (U_2^*)^* = U_2 \quad \Rightarrow U_1 \subset U_2, \tag{5.19}$$

by Theorem 5.4(i). □

As an application of the previous theorems on dual cones, consider a set U specified by one linear equality or inequality constraint.

Theorem 5.6. (i) *If S is a linear subspace of $\mathbb{R}^n$, then $S^* = S^\perp$, where*

$$S^\perp = \{h \in \mathbb{R}^n : h \cdot x = 0, \quad \forall x \in S\}. \tag{5.20}$$

(ii) *Given $a \in \mathbb{R}^n$ consider the convex cones in* 0

$$\begin{gathered} U_1 = \{x \in \mathbb{R}^n : a \cdot x = 0\}, \;\; U_2 = \{x \in \mathbb{R}^n : a \cdot x \geq 0\}, \\ U_3 = \{x \in \mathbb{R}^n : a \cdot x > 0\}. \end{gathered} \tag{5.21}$$

The dual cones are given by

$$U_1^* = \{\lambda a : \lambda \in \mathbb{R}\}, \;\; U_2^* = \{\lambda a : \forall \lambda \geq 0\}, \;\; U_3^* = \begin{Bmatrix} U_2^*, & \text{if } a \neq 0 \\ \mathbb{R}^n, & \text{if } a = 0 \end{Bmatrix} \tag{5.22}$$

(by definition, they are closed convex cones in 0*).*

Proof. (i) By definiton of S^*, $S^\perp \subset S^*$. For all $h \in S^*$,

$$h \cdot x \geq 0, \quad \forall x \in S.$$

Since S is a linear subspace, $-x \in S$ and

$$-(h \cdot x) = h \cdot (-x) \geq 0, \quad \forall x \in S.$$

By combining the two inequalities, we get

$$h \cdot x = 0, \;\; \forall x \in S \quad \Rightarrow h \in S^\perp \quad \Rightarrow S^\perp = S^*.$$

(ii) For U_1, we have by (i) $U_1^* = U_1^\perp$ since U_1 is a linear subspace. Introduce the linear map $x \mapsto Ax = a \cdot x : \mathbb{R}^n \to \mathbb{R}$. Note that $U_1 = \operatorname{Ker} A$ and, by Theorem 5.2,

$$U_1^\perp = [\operatorname{Ker} A]^\perp = \operatorname{Im} A^\top.$$

Here, the transposed map $A^\top : \mathbb{R} \to \mathbb{R}^n$ is easily characterized:

$$\forall x \in \mathbb{R}^n,\ \forall \lambda \in \mathbb{R}, \quad A^\top \lambda \cdot x = \lambda A x = \lambda(a \cdot x) = (\lambda a) \cdot x$$
$$\Rightarrow A^\top \lambda = \lambda a \quad \Rightarrow \operatorname{Im} A^\top = \{\lambda a : \lambda \in \mathbb{R}\}.$$

For U_2, observe that $U_1 \subset U_2$ and, by Theorem 5.4(i), $U_2^* \subset U_1^* = \{\lambda a : \lambda \in \mathbb{R}\}$. If $a = 0$, then $U_2^* = \{0\}$. If $a \neq 0$, then for all $h \in U_2^*$, there exists $\lambda \in \mathbb{R}$ such that $h = \lambda a$ and

$$\lambda(a \cdot x) = (\lambda a) \cdot x = h \cdot x \geq 0, \ \ \forall x \in U_2.$$

In particular, since $0 \neq a \in U_2$, by picking $x = a$, we have $\lambda \|a\|^2 \geq 0$ and hence $\lambda \geq 0$.

Conversely, if $h = \lambda a$ for $\lambda \geq 0$,

$$\forall x \ \text{ such that } \ a \cdot x \geq 0$$

and we have

$$(\lambda a) \cdot x = \lambda(a \cdot x) \geq 0 \quad \Rightarrow \lambda a \in U_2^*.$$

For U_3, we have two cases:

$$a = 0 \quad \Rightarrow U_3 = \varnothing \quad \Rightarrow \ U_3^* = \mathbb{R}^n$$

or

$$a \neq 0 \quad \Rightarrow \overline{U}_3 = U_2 \quad \Rightarrow U_3^* = (\overline{U}_3)^* = U_2^*.$$

Hence the result follows by Theorem 5.4(ii). □

6 Necessary Optimality Conditions for U Arbitrary

6.1 Necessary Optimality Condition

We now turn to the characterization of a local minimizer $x \in U$ with respect to an arbitrary set U that will be approximated by the Bouligand cone $T_U(x)$ at x. We proceed in two steps: first a Hadamard semidifferentiable objective function and then a general function.

6.1.1 Hadamard Semidifferentiable Objective Function

It is now quite remarkable to see the natural complementarity between the Bouligand's cone, $T_U(x)$, and the Hadamard semidifferential $d_H f(x; v)$ of the objective function f in the context of a local minimizer x of f on U.

Theorem 6.1. *Let $\varnothing \neq U \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, and $x \in U \cap \operatorname{dom} f$ be a local minimizer of f with respect to U.*

(i) *If f is Hadamard semidifferentiable at x, then*

$$\forall h \in T_U(x), \quad d_H f(x; h) \geq 0 \tag{6.1}$$

or, equivalently,

$$T_U(x) \subset C_f(x) \stackrel{\text{def}}{=} \left\{ h \in \mathbb{R}^n : \overline{d}_H f(x; h) \geq 0 \right\}. \tag{6.2}$$

(ii) *If f is Fréchet (Hadamard) differentiable at x, then*

$$\forall h \in T_U(x), \quad \nabla f(x) \cdot h \geq 0, \tag{6.3}$$

or, equivalently,

$$T_U(x) \subset C_f(x) = \{h \in \mathbb{R}^n : \nabla f(x) \cdot h \geq 0\}. \tag{6.4}$$

If $x \in \operatorname{int} U$, then $T_U(x) = \mathbb{R}^n$ and condition (6.4) *reduces to*

$$\nabla f(x) = 0. \tag{6.5}$$

The cone $T_U(x)$ can be replaced by $C_U(x)$, $T_U^\flat(x)$, or $D_U(x)$. In general, they yield weaker necessary optimality conditions.

Remark 6.1.
As $T_U(x)$, $C_U(x)$, $T_U^\flat(x)$, and $D_U(x)$ are cones in 0, the set $C_f(x)$ is also a *cone* in 0. We shall see later in section 6.2 and in Chapter 5 how to use the notion of dual cone to obtain the theorem of Lagrange multipliers and the Karush–Kuhn–Tucker's theorem as well as the mixed case of simultaneous equality and inequality constraints. □

Proof of Theorem 6.1. (i) Let $h \in T_U(x)$. By Definition 4.1, there exist sequences $\{t_n > 0\}$, $t_n \searrow 0$ as $n \to \infty$, and $\{x_n\} \subset U$ such that

$$\lim_{n\to\infty} \frac{x_n - x}{t_n} = h.$$

As x is a local minimizer of f on U, there exists a neighborhood of x and hence a ball $B_\eta(x)$ of radius $\eta > 0$ such that

$$\forall y \in U \cap B_\eta(x), \quad f(y) \geq f(x).$$

Since x_n converges to x, there exists $N > 0$ such that

$$\forall n \geq N, \quad x_n \in U \cap B_\eta(x) \quad \Rightarrow \forall n \geq N, \quad f(x_n) \geq f(x).$$

But

$$\frac{f(x + t_n \frac{x_n - x}{t_n}) - f(x)}{t_n} = \frac{f(x_n) - f(x)}{t_n} \geq 0$$

and since $(x_n - x)/\varepsilon_n \to h$ as $t_n \searrow 0$ and $d_H f(x;h)$ exists, we get

$$d_H f(x;h) = \lim_{n\to\infty} \frac{f(x + t_n \frac{x_n - x}{t_n}) - f(x)}{t_n} \geq 0.$$

(ii) From the equivalences of Theorem 3.1 of Chapter 3, a Fréchet differentiable function is Hadamard differentiable.

Finally, in view of the fact that $C_U(x) \subset T_U^\flat(x) \subset T_U(x)$ and $D_U(x) \subset T_U^\flat(x) \subset T_U(x)$ (see section 4.3), the cone $T_U(x)$ can be replaced in Theorem 6.1 by $C_U(x)$, $T_U^\flat(x)$, or $D_U(x)$ to produce weaker necessary optimality conditions. □

6.1.2 ► Arbitrary Objective Function

Theorem 6.2. *Let $\varnothing \neq U \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, and $x \in U \cap \operatorname{dom} f$ be a local minimizer of f with respect to U.*

(i) *If x is not an isolated point of U, then*

$$\forall h \in T_U(x), \quad \overline{d}_H f(x;h) \geq 0 \tag{6.6}$$

or, equivalently,

$$T_U(x) \subset C_f(x) \stackrel{\text{def}}{=} \left\{h \in \mathbb{R}^n : \overline{d}_H f(x;h) \geq 0\right\}. \tag{6.7}$$

(ii) *If f is Lipschitzian at $x \in \operatorname{int}(\operatorname{dom} f)$, then*

$$\forall h \in T_U(x), \quad \overline{d}_C f(x;h) \geq \overline{d}_H f(x;h) \geq 0. \tag{6.8}$$

(iii) *If f is semiconvex and $x \in \operatorname{int}(\operatorname{dom} f)$, then*

$$\forall h \in T_U(x), \quad \overline{d}_C f(x;h) = d_H f(x;h) \geq 0. \tag{6.9}$$

The cone $T_U(x)$ can be replaced by $C_U(x)$, $T_U^\flat(x)$, or $D_U(x)$. In general, they yield weaker necessary optimality conditions.

Remark 6.2.
In general, for $x \in U \cap \operatorname{dom} f$ and $y \in U$,

$$\underline{d}_H f(x;y-x) \leq \underline{d} f(x;y-x) \leq \overline{d} f(x;y-x) \leq \overline{d}_H f(x;y-x).$$

When $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, $\operatorname{dom} f \neq \varnothing$, is convex, there are two cases. If $x \in \operatorname{int}(\operatorname{dom} f)$, we get from part (ii) of Theorem 6.2,

$$\underline{d} f(x;y-x) = d_H f(x;y-x) = \overline{d}_H f(x;y-x) \geq 0.$$

So, from the implication (ii) $\Rightarrow$ (i) of Theorem 3.5, our necessary optimality condition is also sufficient. However, for all $x \in \partial(\operatorname{dom} f) \cap \operatorname{dom} f$, we have

$$\overline{d}_H f(x;y-x) = +\infty > 0,$$

Hence, the condition of part (i) of Theorem 6.2 is verified by all the points of $U \cap \partial(\operatorname{dom} f) \cap \operatorname{dom} f$ but they are not necessarily all minimizers. The condition of part (i) is indeed necessary but not always sufficient. The explanation behind that is that, by using the local approximation $T_U(x)$ of U at the point x, one cannot require as much of the semidifferential. It is necessary to go from $\underline{d}_H f(x;y-x)$ to $\overline{d}_H f(x;y-x)$ as can be seen from the proof of the theorem. □

Proof of Theorem 6.2. (i) Note that, if x was an isolated point of U, then $T_U(x)$ would be $\{0\}$. Let $h \in T_U(x)$. By Definition 4.1, there exist sequences $\{t_n > 0\}$, $t_n \searrow 0$ as $n \to \infty$, and $\{x_n\} \subset U$ such that

$$\lim_{n\to\infty} \frac{x_n - x}{t_n} = h.$$

Since a limsup is involved, the elements of the sequence of quotients must also be different from h. If $h = 0$, we use the fact that x is not an isolated point of U; that is, for all $r > 0$, $B'_r(x) \cap U \neq \varnothing$. Choose the sequence $t_n = 1/n$ and some $x_n \in B'_{t_n^2}(x) \cap U$. By construction, $x_n \neq x$ and

$$0 \neq \left\| \frac{x_n - x}{t_n} \right\| < \frac{t_n^2}{t_n} = t_n \to 0 \quad \Rightarrow \frac{x_n - x}{t_n} \to 0.$$

If $h \neq 0$, the sequence $\{t_n\}$ can be modified in such a way that $(x_n - x)/t'_n \neq h$. If $(x_n - x)/t_n - h \neq 0$, keep $t'_n = t_n$. If $(x_n - x)/t_n - h = 0$, choose $t'_n = t_n/(1 + 1/n)$ for which

$$\frac{x_n - x}{t'_n} - h = \frac{t_n}{t'_n}\left[\frac{x_n - x}{t_n} - h\right] + \left(\frac{t_n}{t'_n} - 1\right) h = \frac{1}{n} h$$
$$\Rightarrow \frac{x_n - x}{t'_n} - h \neq 0 \quad \text{and} \quad \frac{x_n - x}{t'_n} \to h.$$

As x is a local minimizer of f on U, there exists a neighborhood of x and hence a ball $B_\eta(x)$ of radius $\eta > 0$ such that

$$\forall y \in U \cap B_\eta(x), \quad f(y) \geq f(x).$$

Since x_n converges to x, there exists $N > 0$ such that

$$\forall n \geq N, \quad x_n \in U \cap B_\eta(x) \quad \Rightarrow \forall n \geq N, \quad f(x_n) \geq f(x).$$

But

$$\frac{f(x + t_n \frac{x_n - x}{t_n}) - f(x)}{t_n} = \frac{f(x_n) - f(x)}{t_n} \geq 0$$

and since $(x_n - x)/\varepsilon_n \neq h$, as $t_n \searrow 0$ and $(x_n - x)/t_n \to h$, we get

$$\overline{d}_H f(x;h) \geq \limsup_{n \to \infty} \frac{f(x + t_n \frac{x_n - x}{t_n}) - f(x)}{t_n} \geq 0.$$

(ii) If x is an isolated point of U, then $T_U(x) = \{0\}$ and, from Theorem 3.1 of Chapter 3, $\overline{d}_C f(x;0) = \overline{d}_H f(x;0) = 0$ since f is Lipschitzian at x. Otherwise, since f is Lipschitzian at x, by the same theorem, for all $h \in \mathbb{R}^n$, $\overline{d}_C f(x;h) \geq \overline{d}_H f(x;h)$ and the result follows from part (i).

(iii) From Theorem 4.11 in section 4.4 of Chapter 3. □

6.2 Dual Necessary Optimality Condition

From Theorem 6.1 or 6.2, the necessary optimality condition can be written as

$$T_U(x) \subset C_f(x) \tag{6.10}$$

which implies the *dual necessary optimality condition*

$$C_f(x)^* \subset T_U(x)^* \tag{6.11}$$

on the dual cones (Theorem 5.4). Moreover, we know that if $C_f(x)$ is a closed convex cone in 0, the necessary condition (6.10) is equivalent to the dual condition (6.11) (Theorem 5.5(ii)). So, in general, the dual condition will be weaker than the *primal* condition (6.10) since it is equivalent to the weaker condition

$$T_U(x) \subset T_U(x)^{**} \subset C_f(x)^{**} = \overline{\text{co}}\, C_f(x). \tag{6.12}$$

Unfortunately, $C_f(x)$ is not always a closed convex cone.

Lemma 6.1. (i) *$C_f(x)$ is a cone in* 0.

(ii) *If for each $v \in \mathbb{R}^n$ the semidifferential $d_H f(x;v)$ exists, then the mapping $v \mapsto d_H f(x;v)$ is continuous and $C_f(x)$ is a closed cone in* 0.

Proof. (i) Since

$$\forall \alpha > 0, \quad \overline{d}_H f(x;\alpha v) = \alpha \overline{d}_H f(x;v) \geq 0,$$

$C_f(x)$ is a cone in 0.

(ii) To prove that $C_f(x)$ is closed, it is sufficient to prove that all Cauchy sequences $\{v_n\} \subset C_f(x)$ converge in $C_f(x)$. Let v be the limit of the sequence $\{v_n\}$ in $\mathbb{R}^n$. Since, by assumption, $d_H f(x;v)$ exists, we get by definition,

$$\begin{aligned} d_H f(x;v) &= \lim_{\substack{v_n \to v \\ t \searrow 0}} \frac{f(x+t\,v_n) - f(x)}{t} \\ &= \lim_{v_n \to v} \lim_{t \searrow 0} \frac{f(x+t\,v_n) - f(x)}{t} = \lim_{v_n \to v} d_H f(x;v_n). \end{aligned}$$

Hence, the map $v \mapsto d_H f(x;v)$ is continuous. But, since for all n, $d_H f(x;v_n) \geq 0$, by going to the limit, we get $d_H f(x;v) \geq 0$ and $v \in C_f(x)$. So, the cone $C_f(x)$ is closed. □

We summarize our observations.

Theorem 6.3. *Let $U \subset \mathbb{R}^n$, $x \in U$, and $f : \mathbb{R}^n \to \mathbb{R}$.*

(i) $T_U(x) \subset C_f(x) \quad \Rightarrow C_f(x)^* \subset T_U(x)^* \quad \Rightarrow T_U(x) \subset C_f(x)^{**} = \overline{\text{co}}\, C_f(x)$.

(ii) *If $C_f(x)$ is a closed convex cone in* 0 *(in particular, this is true when f is Fréchet differentiable in x or when f is concave on a neighborhood of x), then*

$$T_U(x) \subset C_f(x) \quad \Longleftrightarrow \quad C_f(x)^* \subset T_U(x)^*.$$

(iii) *When f is Fréchet differentiable at x, then $C_f(x)$ is a closed convex cone in* 0 *(in fact a closed half-space) and*

$$C_f(x)^* = \{\lambda \nabla f(x) : \forall \lambda \geq 0\}.$$

The necessary optimality condition (6.10) *is equivalent to*

$$\{\lambda \nabla f(x) : \forall \lambda \geq 0\} \subset T_U(x)^* \tag{6.13}$$

or simply

$$\nabla f(x) \in T_U(x)^*. \tag{6.14}$$

If $x \in \text{int}\, U$, then condition (6.14) *implies that $\nabla f(x) = 0$.*

Remark 6.3.
Be cautious, as Figure 4.14 shows, in general

$$T_U(x) \subset \operatorname{co} T_U(x) \subsetneq T_{\operatorname{co} U}(x).$$

This implies that

$$C_f(x)^* \subset T_U(x)^* = (\operatorname{co} T_U(x))^* \supsetneq T_{\operatorname{co} U}(x)^*.$$

So we still get a dual condition $C_f(x)^* \subset (\operatorname{co} T_U(x))^*$ by convexifying $T_U(x)$, but the relation goes in the wrong direction if we try to convexify U since, in general, $C_f(x)^* \not\subset T_{\operatorname{co} U}(x)^*$. □

Proof. (i) By Theorem 5.4, $U_1 \subset U_2$ implies $U_2^* \subset U_1^*$ that implies $U_1^{**} \subset U_2^{**}$ that implies $U_1 \subset U_2^{**} = \overline{\operatorname{co}}\, U_2$.

(ii) By Theorem 5.5(ii), if U_2 is a closed convex cone in 0, $U_1 \subset U_2$ is equivalent to $U_2^* \subset U_1^*$.

(iii) When f is Fréchet differentiable at x, $d_H f(x;v) = \nabla f(x) \cdot v$, $C_f(x) = \{h \in \mathbb{R}^n : \nabla f(x) \cdot h \geq 0\}$ is a closed convex cone in 0, and, by Theorem 5.6, $C_f(x)^* = \{\lambda \nabla f(x) : \lambda \geq 0\}$. If there exists $r > 0$ such that $B_r(x_0) \subset U$, then $\mathbb{R}^n = T_{B_r(x_0)}(x_0) \subset T_U(x_0)$ and $T_U(x_0)^* = \{0\}$. Thence, $\nabla f(x_0) = 0$. □

It is readily seen that it is the f and not U that determines if there is equivalence of the primal and dual necessary optimality conditions. So, it would be interesting to characterize the functions f for which $C_f(x)$ is a closed convex cone in 0.

Recalling from Theorem 4.2 that $C_U(x)$ is a closed convex cone in 0 and that $C_U(x) \subset T_U(x)$, we have, from Theorem 6.1 or 6.2(i), $C_U(x) \subset T_U(x) \subset C_f(x)$. This yields the weaker necessary optimality condition

$$C_U(x) \subset C_f(x) \quad \Rightarrow C_f(x)^* \subset C_U(x)^*. \tag{6.15}$$

7 Affine Equality and Inequality Constraints

7.1 Characterization of $T_U(x)$

As a first step in the characterization of the dual of $T_U(x)$, it is useful to consider the case where U is defined by a finite number of affine constraints of the equality or inequality type.

Let $n \geq 1$, $m \geq 1$, and $k \geq 1$ be integers, $a_1, a_2, \ldots, a_m$, $b_1, b_2, \ldots, b_k$ vectors of $\mathbb{R}^n$, and $\alpha_1, \alpha_2, \ldots, \alpha_m$, $\beta_1, \beta_2, \ldots, \beta_k$ real numbers. Consider the closed convex set

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \begin{array}{l} a_i \cdot x \leq \alpha_i,\ 1 \leq i \leq m \\ b_j \cdot x = \beta_j,\ 1 \leq j \leq k \end{array} \right\} \tag{7.1}$$

and the set of indices of *active inequality constraints*

$$I(x) \stackrel{\text{def}}{=} \{i : 1 \leq i \leq m \text{ and } a_i \cdot x = \alpha_i\}. \tag{7.2}$$

The other indices correspond to *inactive* constraints. We say that U is a *polyhedral* convex set since it is the intersection of a finite collection of closed half-spaces by expressing each equality constraint as two inequalities.

Theorem 7.1. *Let U and $I(x)$ be as specified by* (7.1) *and* (7.2). *Then, for each $x \in U$,*

$$T_U(x) = \left\{ h \in \mathbb{R}^n : \begin{array}{ll} a_i \cdot h \le 0, & i \in I(x) \\ b_j \cdot h = 0, & 1 \le j \le k \end{array} \right\} \tag{7.3}$$

is a closed convex cone in 0.

Proof. (i) If $h \in T_U(x)$, there exists a sequence $\{x_n\} \subset U$ and $\varepsilon_n \searrow 0$ such that $(x_n - x)/\varepsilon_n \to h$. Hence, as $a_i \cdot x = \alpha_i$ for all i,

$$\begin{aligned} 0 &= \alpha_i - \alpha_i \ge a_i \cdot x_n - a_i \cdot x = a_i \cdot (x_n - x), \quad \forall i \in I(x) \\ 0 &= \beta_j - \beta_j = b_j \cdot x_n - b_j \cdot x = b_j \cdot (x_n - x), \quad \forall j, 1 \le j \le k. \end{aligned}$$

The direction of the above inequalities are not changed by dividing both sides by $\varepsilon_n > 0$:

$$\begin{cases} 0 \ge a_i \cdot \dfrac{x_n - x}{\varepsilon_n}, & \forall i \in I(x) \\ 0 = b_j \cdot \dfrac{x_n - x}{\varepsilon_n}, & \forall j, 1 \le j \le k \end{cases} \Rightarrow \begin{cases} 0 \ge a_i \cdot h, & \forall i \in I(x) \\ 0 = b_j \cdot h, & \forall j, 1 \le j \le k. \end{cases}$$

(ii) Denote by U_0 the right-hand side of identity (7.3). For $h \in U_0$, consider the path $x(\varepsilon) = x + \varepsilon h$, $\varepsilon > 0$. As $h = (x(\varepsilon) - x)/\varepsilon$, it is sufficient to prove that for ε sufficiently small, $x(\varepsilon) \in U$. For $i \notin I(x)$, $a_i \cdot x < \alpha_i$ and

$$\alpha \overset{\text{def}}{=} \sup_{\substack{1 \le i \le m \\ i \notin I(x)}} a_i \cdot x - \alpha_i < 0, \qquad c \overset{\text{def}}{=} 1 + \sup_{\substack{1 \le i \le m \\ i \notin I(x)}} \|a_i\|.$$

For $\varepsilon_0 = |\alpha|/(2(c\,\|h\| + 1)$ and $0 < \varepsilon < \varepsilon_0$,

$$\begin{aligned} \forall i \notin I(x), \quad a_i \cdot x(\varepsilon) - \alpha_i &= a_i \cdot x - \alpha_i + \varepsilon\, a_i \cdot h \\ &< \alpha + \varepsilon\, \|a_i\|\, \|h\| \le \alpha + \varepsilon\, c\, \|h\| \\ &< \alpha + \frac{|\alpha|}{2(c\,\|h\| + 1)}\, c\, \|h\| \le \frac{\alpha}{2} < 0 \\ \forall i \in I(x), \quad a_i \cdot x(\varepsilon) - \alpha_i &= a_i \cdot x - \alpha_i + \varepsilon\, a_i \cdot h \le 0 \\ \forall j, \quad b_j \cdot x(\varepsilon) - \beta_j &= b_j \cdot x - \beta_j + \varepsilon\, b_j \cdot h = 0. \end{aligned}$$

Therefore, $x(\varepsilon) \in U$ for $\varepsilon < \varepsilon_0$. □

7.2 Dual Cones for Linear Constraints

We have seen in Theorem 7.1 that the cone $T_U(x)$ associated with a U specified by affine equalities or inequalities is a closed convex cone in 0 characterized by linear equalities or inequalities. Therefore, in order to characterize $T_U(x)^*$, it is now sufficient to characterize the dual cones U^* of sets U specified by linear equalities or inequalities.

Let $m \ge 1$ be an integer and $\{a_1, \ldots, a_m\}$ a sequence of vectors in $\mathbb{R}^n$. Consider the following sets specified by a finite number of linear equality and inequality constraints:

$$U_0 \overset{\text{def}}{=} \left\{ x \in \mathbb{R}^n : a_j \cdot x = 0,\ j = 1, \ldots, m \right\}, \tag{7.4}$$

$$U_+ \overset{\text{def}}{=} \left\{ x \in \mathbb{R}^n : a_j \cdot x \ge 0,\ j = 1, \ldots, m \right\}. \tag{7.5}$$

It is easy to check that U_0 is a linear subspace and that U_+ is a closed convex cone in 0. Let $\{e_j^m : 1 \leq j \leq m\}$ be the canonical orthonormal basis in $\mathbb{R}^m$ and recall that, for all $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$,

$$y = \sum_{j=1}^{m} y_j e_j^m \quad \text{and} \quad y \cdot e_j^m = y_j, \quad j = 1, \ldots, m.$$

Consider the linear map $A : \mathbb{R}^n \to \mathbb{R}^m$:

$$\boxed{x \mapsto Ax \stackrel{\text{def}}{=} \sum_{j=1}^{m} (a_j \cdot x) e_j^m.} \tag{7.6}$$

Let $A^\top : \mathbb{R}^m \to \mathbb{R}^n$ be its adjoint map: for all $\lambda \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$,

$$\begin{aligned} A^\top \lambda \cdot x = \lambda \cdot Ax = \lambda \cdot \sum_{j=1}^{m} (a_j \cdot x) e_j^m &= \sum_{j=1}^{m} (\lambda \cdot e_j^m)(a_j \cdot x) = \left[\sum_{j=1}^{m} (\lambda \cdot e_j^m) a_j \right] \cdot x \\ \Rightarrow \forall \lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m, &\quad \boxed{A^\top \lambda = \sum_{j=1}^{m} \lambda \cdot e_j^m \, a_j = \sum_{j=1}^{m} \lambda_j \, a_j.} \end{aligned} \tag{7.7}$$

It is easy to check that

$$\boxed{U_0 = \operatorname{Ker} A.}$$

In the other case, introduce the *positive cone* in $\mathbb{R}^m$

$$\boxed{\mathbb{R}_+^m \stackrel{\text{def}}{=} \left\{ y = (y_1, \ldots, y_m) \in \mathbb{R}^m : y_j \geq 0, \quad 1 \leq j \leq m \right\}.} \tag{7.8}$$

This definition is completely equivalent to

$$\mathbb{R}_+^m = \left\{ \sum_{j=1}^{m} \lambda_j \, e_j^m : \lambda_j \geq 0, \quad 1 \leq j \leq m \right\}. \tag{7.9}$$

By using $A : \mathbb{R}^n \to \mathbb{R}^m$ defined by (7.6), we get

$$\boxed{\forall j, 1 \leq j \leq m, \quad a_j = A^\top e_j^m} \quad \text{and} \quad \boxed{U_+ = \left\{ x \in \mathbb{R}^n : Ax \in \mathbb{R}_+^m \right\},}$$

since

$$\begin{aligned} Ax \in \mathbb{R}_+^m \quad &\Rightarrow \forall j, \; x \cdot a_j = x \cdot A^\top e_j^m = Ax \cdot e_j^m \geq 0 \quad \Rightarrow x \in U_+ \\ x \in U_+ \quad &\Rightarrow \forall j, \; Ax \cdot e_j^m = x \cdot A^\top e_j^m = x \cdot a_j \geq 0 \quad \Rightarrow Ax \in \mathbb{R}_+^m. \end{aligned}$$

Theorem 7.2. *Let $m \geq 1$ and $n \geq 1$ be two integers and $\{a_1, \ldots, a_m\}$ a sequence of vectors in $\mathbb{R}^n$.*

(i) *The set*

$$U_0 \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : a_j \cdot x = 0,\, j = 1, \ldots, m\right\} \tag{7.10}$$

is a linear subspace of $\mathbb{R}^n$ and its dual is

$$U_0^* = \left\{ \sum_{j=1}^{m} \lambda_j\, a_j \,:\, \lambda_j \in \mathbb{R},\, j = 1, \ldots, m \right\}. \tag{7.11}$$

(ii) *The set*

$$U_+ \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : a_j \cdot x \geq 0,\, j = 1, \ldots, m\right\} \tag{7.12}$$

is a closed convex cone in 0 and its dual is

$$U_+^* = \left\{ \sum_{j=1}^{m} \lambda_j\, a_j \,:\, \lambda_j \geq 0,\, j = 1, \ldots, m \right\}. \tag{7.13}$$

(iii) *Let $\mathbb{R}^m_+$ be the positive cone in $\mathbb{R}^m$ defined by (7.8) and $A : \mathbb{R}^n \to \mathbb{R}^m$ a linear map. Then $(\operatorname{Ker} A)^* = \operatorname{Im} A^\top$ and the dual U_+^* of*

$$U_+ \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n \,:\, Ax \in \mathbb{R}^m_+\right\} \tag{7.14}$$

is given by

$$U_+^* = \left\{A^\top y \,:\, y \in \mathbb{R}^m_+\right\} = A^\top \mathbb{R}^m_+. \tag{7.15}$$

Equivalently, U_+ and U_+^* are given by (7.12) and (7.13) by choosing $a_j = A^\top e_j^m$, $1 \leq j \leq m$.

Proof. (i) By using the linear map A defined in (7.6) constructed from the sequence $\{a_1, \ldots, a_m\}$, we have shown that $U_0 = \operatorname{Ker} A$. By Theorems 5.6(i) and 5.2, we have $U_0^* = U_0^\perp = \operatorname{Im} A^\top$. From the characterization (7.7) of $A^\top$, we get

$$U_0^* = \operatorname{Im} A^\top = \left\{ \sum_{j=1}^{m} \lambda_j\, a_j \,:\, \lambda_j \in \mathbb{R},\, j = 1, \ldots, m \right\}.$$

(ii) Using A the set U_+ is given by (7.14). Moreover, in the light of (7.9),

$$\mathbb{R}^m_+ = \left\{ \sum_{j=1}^{m} \lambda_j\, e_j^m \,:\, \lambda_j \geq 0,\, 1 \leq j \leq m \right\}$$

$$A^\top \mathbb{R}^m_+ = \left\{ \sum_{j=1}^m \lambda_j\, A^\top e^m_j \,:\, \lambda_j \geq 0,\, 1 \leq j \leq m \right\}$$
$$= \left\{ \sum_{j=1}^m \lambda_j\, a_j \,:\, \lambda_j \geq 0,\, 1 \leq j \leq m \right\}.$$

So it is sufficient to prove part (iii) for an arbitrary linear map $A : \mathbb{R}^n \to \mathbb{R}^m$.

(iii) We first prove that $A^\top \mathbb{R}^m_+ \subset U^*_+$. For $h \in A^\top \mathbb{R}^m_+$, there exists $\lambda \in \mathbb{R}^m_+$ such that $h = A^\top \lambda$. For all x in U_+, we have $Ax \cdot e^m_j \geq 0$, j, $1 \leq j \leq m$. Then

$$\forall x \in U_+, \quad h \cdot x = A^\top \lambda \cdot x = \lambda \cdot Ax = \sum_{j=1}^m \underbrace{\lambda \cdot e^m_j}_{\geq 0}\, \underbrace{Ax \cdot e^m_j}_{\geq 0} \geq 0.$$

Then, by definition of the dual cone, $h \in U^*_+$ and $A^\top \mathbb{R}^m_+ \subset U^*_+$.

To prove that $U^*_+ \subset A^\top \mathbb{R}^m_+$, we first show thate $(A^\top \mathbb{R}^m_+)^* \subset U_+$ and since $A^\top \mathbb{R}^m_+$ is a closed convex cone in 0 by Theorem 5.2, we get, by Theorems 5.4(i) and 5.5(i),

$$(A^\top \mathbb{R}^m_+)^* \subset U_+ \quad \Rightarrow U^*_+ \subset ((A^\top \mathbb{R}^m_+)^*)^* = A^\top \mathbb{R}^m_+.$$

This, combined with $U^*_+ \supset A^\top \mathbb{R}^m_+$, yields $U^*_+ = A^\top \mathbb{R}^m_+$. So we now prove that $(A^\top \mathbb{R}^m_+)^* \subset U_+$. Let $h \in (A^\top \mathbb{R}^m_+)^*$. For each j, $1 \leq j \leq m$, $e^m_j \in \mathbb{R}^m_+$ and $A^\top e^m_j \in A^\top \mathbb{R}^m_+$, and hence

$$\forall j, \quad Ah \cdot e^m_j = h \cdot A^\top e^m_j \geq 0 \quad \Rightarrow \forall j, \quad Ah \cdot e^m_j \geq 0 \quad \Rightarrow h \in U_+,$$

and $(A^\top \mathbb{R}^m_+)^* \subset U_+$. □

Now consider U specified by a finite number of linear equalities and inequalities:

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : a_i \cdot x \leq 0,\, 1 \leq i \leq m \text{ and } b_j \cdot x = 0,\, 1 \leq j \leq k \right\},$$

where $\{a_i : 1 \leq i \leq m\}$ and $\{b_j : 1 \leq j \leq k\}$ are sequences of vectors in $\mathbb{R}^n$. As in the context of Theorem 5.5, introduce the matrices A $(m \times n)$ and B $(k \times n)$:

$$A_{ij} = (a_i)_j,\, 1 \leq i \leq m,\, 1 \leq j \leq n, \text{ and } B_{ij} = (b_i)_j,\, 1 \leq i \leq k,\, 1 \leq j \leq n,$$

and the linear maps $A : \mathbb{R}^n \to \mathbb{R}^m$ and $B : \mathbb{R}^n \to \mathbb{R}^k$ using the canonical orthonormal bases $\{e^m_i\}$ of $\mathbb{R}^m$ and $\{e^k_j\}$ of $\mathbb{R}^k$ (see (7.6) before Theorem 7.2). The set U can then be rewritten as

$$U = \left\{ x \in \mathbb{R}^n : -Ax \in \mathbb{R}^m_+ \text{ and } Bx = 0 \right\}.$$

Theorem 7.3. *Let $m \geq 1$, $n \geq 1$, and $k \geq 1$ be three integers and $\{a_1, \dots, a_m\}$ and $\{b_1, \dots, b_k\}$ two sequences of vectors in $\mathbb{R}^n$.*

(i) *The set*

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : a_i \cdot x \leq 0,\, 1 \leq i \leq m, \text{ and } b_j \cdot x = 0,\, 1 \leq j \leq k \right\} \tag{7.16}$$

is a closed convex cone in 0 *and its dual* U^* *is given by*

$$U^* = \left\{ -\sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j b_j \,:\, \lambda_i \geq 0,\, 1 \leq i \leq m,\; \mu_j \in \mathbb{R},\, 1 \leq j \leq k \right\}. \tag{7.17}$$

(ii) *Given the positive cone* $\mathbb{R}^m_+$ *in* $\mathbb{R}^m$ *defined by* (7.8) *and two linear maps* $A : \mathbb{R}^n \to \mathbb{R}^m$ *and* $B : \mathbb{R}^n \to \mathbb{R}^k$, *the dual* U^* *of*

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n \,:\, -Ax \in \mathbb{R}^m_+ \text{ and } Bx = 0 \right\} \tag{7.18}$$

is given by

$$U^* = \left\{ -A^\top \lambda + B^\top \mu \,:\, \lambda \in \mathbb{R}^m_+,\; \mu \in \mathbb{R}^k \right\}. \tag{7.19}$$

Equivalently, U *and* U^* *are, respectively, given by* (7.16) *and* (7.17) *by choosing* $a_i = A^\top e^m_i$, $1 \leq i \leq m$, *and* $b_j = B^\top e^k_j$, $1 \leq j \leq k$.

Proof. (i) U can be rewritten in the form

$$U = \left\{ x \in \mathbb{R}^n \,:\, \begin{array}{l} a_i \cdot x \leq 0,\; 1 \leq i \leq m, \text{ and} \\ b_j \cdot x \leq 0 \text{ and } -b_j \cdot x \leq 0,\; 1 \leq j \leq k \end{array} \right\}.$$

By Theorem 7.2(ii),

$$\begin{aligned} U^* &= \left\{ -\sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} (\lambda^+_j - \lambda^-_j) b_j \,:\, \begin{array}{l} \lambda_i \geq 0,\, 1 \leq i \leq m \\ \lambda^+_j \geq 0,\, 1 \leq j \leq k \\ \lambda^-_j \geq 0,\, 1 \leq j \leq k \end{array} \right\} \\ &= \left\{ -\sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j b_j \,:\, \begin{array}{l} \lambda_i \geq 0,\, 1 \leq i \leq m \\ \mu_j \in \mathbb{R},\, 1 \leq j \leq k \end{array} \right\}. \end{aligned}$$

(ii) Since we already have $-Ax \in \mathbb{R}^m_+$, rewrite condition $Bx = 0$ as two inequality conditions:

$$Bx \in \mathbb{R}^k_+ \text{ and } -Bx \in \mathbb{R}^k_+.$$

Introduce the new linear map

$$x \mapsto Cx \stackrel{\text{def}}{=} (-Ax, Bx, -Bx) : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^k$$

whose transpose is

$$(\lambda, \xi, \zeta) \mapsto C^\top(\lambda, \xi, \zeta) = -A^\top \lambda + B^\top \xi - B^\top \zeta : \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^n. \tag{7.20}$$

It is now sufficient to invoke Theorem 7.2(iii) for the set

$$U = \left\{ x \in \mathbb{R}^n \,:\, Cx \in \mathbb{R}^m_+ \times \mathbb{R}^k_+ \times \mathbb{R}^k_+ \right\}$$

to get

$$U^* = \left\{-A^\top\lambda + B^\top\xi - B^\top\zeta \,:\, \lambda \in \mathbb{R}^m_+ \text{ and } \xi,\zeta \in \mathbb{R}^k_+\right\}.$$

But

$$\left\{\xi - \zeta \,:\, \xi,\zeta \in \mathbb{R}^k_+\right\} = \mathbb{R}^k$$

and finally

$$U^* = \left\{-A^\top\lambda + B^\top\mu \,:\, \lambda \in \mathbb{R}^m_+, \text{ and } \mu \in \mathbb{R}^k\right\}. \qquad \square$$

7.3 Linear Programming Problem

Let $n \geq 1$, $m \geq 1$, and $k \geq 1$ be three integers, let q, $a_1,a_2,\dots,a_m$, $b_1,b_2,\dots,b_k$ be vectors in $\mathbb{R}^n$, and let $\alpha_1,\alpha_2,\dots,\alpha_m$, $\beta_1,\beta_2,\dots,\beta_k$ be scalars. As a direct application of the material developed in the last two sections, consider the following *linear programming* (LP) problem:[9]

$$\inf_{x\in U} f(x), \quad f(x) \overset{\text{def}}{=} q\cdot x, \quad U \overset{\text{def}}{=} \left\{x \in \mathbb{R}^n : \begin{array}{l} a_i \cdot x \leq \alpha_i,\ 1 \leq i \leq m \\ b_j \cdot x = \beta_j,\ 1 \leq j \leq k \end{array}\right\}, \tag{7.21}$$

where U is specified by a finite number of equality or inequality constraints on affine functions. The problem is *feasible*[10] when $U \neq \varnothing$ and *infeasible* when $U = \varnothing$.

The objective function f is convex on $\mathbb{R}^n$. The set U is closed and convex. So the necessary and sufficient condition (3.4) of Theorem 3.2 yields

$$\exists x \in U \text{ such that } q\cdot(y-x) \geq 0 \text{ for all } y \in U. \tag{7.22}$$

In this case, it is equivalent to

$$\exists x \in U \text{ such that } q\cdot h \geq 0 \text{ for all } h \in T_U(x) \quad \text{or} \quad q \in T_U(x)^*.$$

For affine constraints, the cone $T_U(x)$ has been characterized by (7.3) of Theorem 7.1: for each $x \in U$,

$$T_U(x) = \left\{h \in \mathbb{R}^n : \begin{array}{ll} a_i \cdot h \leq 0, & i \in I(x) \\ b_j \cdot h = 0, & 1 \leq j \leq k \end{array}\right\}, \tag{7.23}$$

where

$$I(x) \overset{\text{def}}{=} \{i : 1 \leq i \leq m \text{ and } a_i \cdot x = \alpha_i\} \tag{7.24}$$

is the set of indices of *active constraints*.

[9] From L. D. BERKOVITZ [**1**]: "For a brief account of the origin and early days of linear programming by a pioneer in the field, see G. B. DANTZIG [**4**]."

[10] Some authors use the terminology *consistent* (see, for instance, L. D. BERKOVITZ [**1**, p. 180]).

Theorem 7.4. *Let U and q be as specified by* (7.21).

(i) *The necessary and sufficient condition for the existence of a minimizer on U to problem* (7.21) *is*

$$\exists x \in U \text{ such that } q \in T_U(x)^*, \tag{7.25}$$

$$\begin{aligned} T_U(x)^* &= \left\{ -\sum_{i\in I(x)} \lambda_i a_i - \sum_{j=1}^{k} \mu_j b_j : \begin{matrix} \lambda_i \geq 0,\ i \in I(x) \\ \mu_j \in \mathbb{R},\ 1 \leq j \leq k \end{matrix} \right\} \\ &= \left\{ -\sum_{i=1}^{m} \lambda_i a_i - \sum_{j=1}^{k} \mu_j b_j : \begin{matrix} \lambda_i \geq 0,\ \lambda_i (a_i \cdot x - \alpha_i) = 0 \\ a_i \cdot x - \alpha_i \leq 0,\ 1 \leq i \leq m \\ \mu_j \in \mathbb{R},\ 1 \leq j \leq k \end{matrix} \right\}. \end{aligned} \tag{7.26}$$

(ii) *The necessary and sufficient condition* (7.25) *is equivalent to the existence of $x \in \mathbb{R}^n$, $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, and $\mu = (\mu_1, \dots, \mu_k) \in \mathbb{R}^k$ such that*

$$\begin{aligned} &\lambda_i \geq 0, \quad a_i \cdot x - \alpha_i \leq 0, \quad \lambda_i (a_i \cdot x - \alpha_i) = 0, \quad 1 \leq i \leq m, \\ &b_j \cdot x = \beta_j, \quad 1 \leq j \leq k, \\ &q + \sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j b_j = 0. \end{aligned} \tag{7.27}$$

Proof. (i) From Theorem 3.4 of section 3.3, from the dual condition of Theorem 6.3(ii), from the characterization of $T_U(x)$ in Theorem 7.1, and from formula (7.17) of Theorem 7.3. The second expression of $T_U(x)^*$ is obtained by choosing $\lambda_i = 0$ for $i \notin I(x)$, that is, when $a_i \cdot x - \alpha_i < 0$. This way the condition $i \in I(x)$ can be replaced by $\lambda_i (a_i \cdot x - \alpha_i) = 0$ for all i, $1 \leq i \leq m$.

(ii) From part (i). □

It will be convenient to associate with the families of vectors $\{a_i\}$ and $\{b_j\}$, the linear maps

$$x \mapsto Ax \overset{\text{def}}{=} \sum_{i=1}^{m} (a_i \cdot x) e_i^m : \mathbb{R}^n \to \mathbb{R}^m, \ x \mapsto Bx \overset{\text{def}}{=} \sum_{i=1}^{k} (b_j \cdot x) e_j^k : \mathbb{R}^n \to \mathbb{R}^k \tag{7.28}$$

$$\Rightarrow A^\top \lambda = \sum_{i=1}^{m} \lambda_i a_i \text{ and } B^\top \mu = \sum_{j=1}^{k} \mu_j b_j \tag{7.29}$$

and, with $\{\alpha_i\}$ and $\{\beta_j\}$, the vectors

$$\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m \quad \text{and} \quad \beta = (\beta_1, \dots, \beta_k) \in \mathbb{R}^k. \tag{7.30}$$

The vectors of parameters (λ, μ) naturally arise from the characterization of the dual cone $T_U(x)$. They turn out to be the *Lagrange multipliers* associated with the constraint set U. So, the abstract necessary and sufficient condition in terms of $T_U(x)$ becomes an explicit condition in terms of the constraint functions.

In general, for an arbitrary function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, it is natural to introduce the function $L : \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k \to \mathbb{R}$ defined as

$$L(x,\lambda,\mu) \stackrel{\text{def}}{=} f(x) + (Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu. \tag{7.31}$$

The function L is called the *Lagrangian* of the problem. It is readily checked that

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \begin{cases} f(x), & \text{if } x \in U \\ +\infty, & \text{if } x \notin U \end{cases} \tag{7.32}$$

$$\Rightarrow v_+ \stackrel{\text{def}}{=} \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \inf_{x\in U} f(x). \tag{7.33}$$

By augmenting the number of variables, we now have a new function (7.32) from $\mathbb{R}^n$ into $\mathbb{R}\cup\{+\infty\}$ with domain U to minimize over all $\mathbb{R}^n$ instead of over U. This function is exactly the function $f_U = f + I_U$ introduced in Chapter 2, where I_U is the indicator function of U:

$$\begin{aligned} f(x) + I_U(x) &= \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) \\ &= f(x) + \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} [(Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu]. \end{aligned}$$

It is easy to verify that

$$I_U(x) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} [(Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu]. \tag{7.34}$$

In the terminology of *game theory* the *inf sup* is the *upper value* of the game and the associated problem is the *upper* problem, while

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \tag{7.35}$$

is the *lower value* of the game and the associated problem is the *lower* problem. This terminology arises from the fact that we always have the inequality

$$v_- \stackrel{\text{def}}{=} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \le v_+ \stackrel{\text{def}}{=} \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu). \tag{7.36}$$

The existence of a minimizer is directly related to the notion of *saddle point* for the Lagrangian: there exists $(\hat{x},\hat{\lambda},\hat{\mu}) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$ such that

$$\forall x \in \mathbb{R}^n, \forall(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k, \quad L(\hat{x},\lambda,\mu) \le L(\hat{x},\hat{\lambda},\hat{\mu}) \le L(x,\hat{\lambda},\hat{\mu}). \tag{7.37}$$

The general framework of game theory will be presented in section 7.4 and its connection with the Fenchel primal and dual problems in section 7.5. We shall see that the upper problem coincides with the Fenchel primal problem and that the lower problem coincides with the Fenchel dual problem.

In *Linear Programming*, the lower problem is called the *Lagrange dual* problem, inequality (7.36) is called the *weak duality* property. This terminology is misleading since this inequality is always verified. It is not a “mathematical property” enjoyed by some specific games. The difference between the inf sup and the sup inf is called the *duality gap*. When we have equality in (7.36), we say that *strong duality* holds. But equality does not necessarily mean existence of a saddle point!

The function

$$(\lambda,\mu) \mapsto g(\lambda,\mu) \overset{\text{def}}{=} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) : \mathbb{R}^m_+ \times \mathbb{R}^k \to \mathbb{R}\cup\{\pm\infty\} \tag{7.38}$$

is called the *Lagrange dual function*. If $\operatorname{dom} f \neq \varnothing$, $g : \mathbb{R}^m_+ \times \mathbb{R}^k \to \mathbb{R}\cup\{-\infty\}$. The *Lagrange dual problem* consists in maximizing this function. Since the function

$$L(x,\lambda,\mu) = q\cdot x + (Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu$$

is linear in x without constraints, its infimum is given by

$$\begin{aligned} g(\lambda,\mu) &= \inf_{x\in\mathbb{R}^n} q\cdot x + (Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu \\ &= \begin{cases} -\alpha\cdot\lambda - \beta\cdot\mu, & \text{if } q + A^\top\lambda + B^\top\mu = 0 \\ -\infty, & \text{if } q + A^\top\lambda + B^\top\mu \neq 0, \end{cases} \end{aligned} \tag{7.39}$$

$$\operatorname{dom} g = \left\{(\lambda,\mu)\in\mathbb{R}^m\times\mathbb{R}^k : \begin{bmatrix} A^\top & B^\top \end{bmatrix}\begin{bmatrix}\lambda\\ \mu\end{bmatrix} = -q,\ -\lambda_i \le 0,\ 1\le i\le m\right\}. \tag{7.40}$$

This yields the *dual* maximization problem

$$\begin{aligned} &\sup_{(\lambda,\mu)\in V} -\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\cdot\begin{bmatrix}\lambda\\ \mu\end{bmatrix}, \\ &V \overset{\text{def}}{=} \left\{(\lambda,\mu)\in\mathbb{R}^m\times\mathbb{R}^k : \begin{bmatrix} A^\top & B^\top \end{bmatrix}\begin{bmatrix}\lambda\\ \mu\end{bmatrix} = -q,\ -\lambda_i \le 0,\ 1\le i\le m\right\}, \end{aligned} \tag{7.41}$$

where, by definition of V,

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k}\ \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \sup_{(\lambda,\mu)\in V} -\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\cdot\begin{bmatrix}\lambda\\ \mu\end{bmatrix}. \tag{7.42}$$

When $V \neq \varnothing$, the dual problem is *feasible*; otherwise it is *infeasible.*

It has the same form as the initial minimization problem, but it is the maximization of a linear objective function subject to affine inequality and equality constraints. In order to make the primal problem look more like the dual problem, we can introduce a vector of *slack variables* $y \in \mathbb{R}^m_+$:

$$\begin{matrix} -(Ax-\alpha)\in\mathbb{R}^m_+ \\ Bx-\beta = 0 \end{matrix} \iff \begin{matrix} y\in\mathbb{R}^m_+, \quad Ax-\alpha+y=0 \\ Bx-\beta=0. \end{matrix} \tag{7.43}$$

So we can rewrite the constraints and the linear objective function as follows:

$$\begin{aligned} &\inf_{(x,y)\in\hat{U}} \begin{bmatrix} q \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \\ &\hat{U} \stackrel{\text{def}}{=} \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m : \begin{bmatrix} A & I \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, -y_i \le 0, 1 \le i \le m \right\}. \end{aligned} \tag{7.44}$$

This means that there are several ways to write the LP problem by introducing extra variables and transforming equalities into inequalities and vice versa (see Exercise 9.6). Some formulations are more interesting than others. We give the terminology of L. D. BERKOVITZ [**1**, pp. 139–145]. The following notation will be useful to write componentwise inequalities between two vectors $x, y \in \mathbb{R}^n$:

$$x \ge y \iff x_i \ge y_i,\ 1 \le i \le n \iff x - y \in \mathbb{R}^n_+ \tag{7.45}$$

$$x \le y \iff x_i \le y_i,\ 1 \le i \le n \iff y - x \in \mathbb{R}^n_+ \tag{7.46}$$

$$x > y \iff x_i > y_i,\ 1 \le i \le n \quad \text{and} \quad x < y \iff x_i < y_i,\ 1 \le i \le n. \tag{7.47}$$

Definition 7.1.
Let $x, b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$, and A be an $m \times n$ matrix.

(i) The LP problem in *standard form* (SLP) is given as

$$\begin{aligned} &\text{minimize } -b \cdot x \\ &\text{subject to } Ax \le c \quad \text{and} \quad x \ge 0. \end{aligned} \tag{7.48}$$

(ii) The LP problem in *canonical form* (CLP) is given as

$$\begin{aligned} &\text{minimize } -b \cdot x \\ &\text{subject to } Ax = c \quad \text{and} \quad x \ge 0. \end{aligned} \tag{7.49}$$

(iii) The LP problem in *inequality form* (LPI) is given as

$$\begin{aligned} &\text{minimize } -b \cdot x \\ &\text{subject to } Ax \le c. \end{aligned} \tag{7.50}$$

□

For instance, the *simplex method* is applied to the CLP. As for the SLP, it provides the following more symmetrical expression of the dual problem:

$$\begin{aligned} &\text{maximize } -c \cdot \lambda \\ &\text{subject to } A^\top \lambda \ge b \quad \text{and} \quad \lambda \ge 0. \end{aligned} \tag{7.51}$$

We associate with the primal-dual problem the *payoff matrix* of game theory

$$P \stackrel{\text{def}}{=} \begin{bmatrix} 0 & A^\top & -b \\ -A & 0 & c \\ b^\top & -c^\top & 0 \end{bmatrix} \tag{7.52}$$

which is skew-symetric: $P + P^\top = 0$. The saddle point equations can be written as

$$\begin{bmatrix} 0 & A^\top & -b \\ -A & 0 & c \\ b^\top & -c^\top & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \\ 1 \end{bmatrix} \begin{matrix} \geq 0 \\ \geq 0 \\ = 0 \end{matrix} \quad \text{and} \quad \begin{matrix} x \geq 0 \\ \lambda \geq 0. \end{matrix}$$

So there is a connection between LP and symmetric *matrix games*.[11] The underlying theorem for an $(n+m+1) \times (n+m+1)$ skew-symmetric matrix G is that there exists $(\hat{x}, \hat{\lambda}, \hat{t}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ such that

$$G \begin{bmatrix} \hat{x} \\ \hat{\lambda} \\ \hat{t} \end{bmatrix} \geq 0, \quad \begin{bmatrix} \hat{x} \\ \hat{\lambda} \\ \hat{t} \end{bmatrix} \geq 0, \quad \text{and} \sum_{i=1}^{m} \hat{\lambda}_i + \sum_{j=1}^{m} \hat{x}_j + \hat{t} = 1.$$

When applied to the matrix associated with the LP problem we get

$$\begin{bmatrix} 0 & A^\top & -b \\ -A & 0 & c \\ b^\top & -c^\top & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\lambda} \\ \hat{t} \end{bmatrix} \geq 0, \quad \begin{bmatrix} \hat{x} \\ \hat{\lambda} \\ \hat{t} \end{bmatrix} \geq 0, \quad \text{and} \sum_{i=1}^{m} \hat{\lambda}_i + \sum_{j=1}^{m} \hat{x}_j + \hat{t} = 1.$$

By taking the inner product of the first row by $\hat{t}\,\hat{x}$ and of the second row by $\hat{t}\,\hat{\lambda}$ and adding them up we get $\hat{t}\,(b \cdot x - c \cdot \lambda) = 0$. It is shown by G. B. DANTZIG [**4**] that

(a) $\hat{t}\,(b \cdot x - c \cdot \lambda) = 0$;

(b) if $\hat{t} > 0$, then $(\hat{x}\,\hat{t}^{-1}, \hat{\lambda}\,\hat{t}^{-1})$ are optimal solutions to the primal-dual LP problems, respectively;

(c) if $b \cdot x - c \cdot \lambda > 0$, then either the primal problem or the dual problem is infeasible (so neither has an optimal solution).

I. ADLER [**1**] points out that the case $\hat{t} = 0$ and $b \cdot x - c \cdot \lambda = 0$ is missing. From his abstract: "In 1951, G. B. DANTZIG [**2**] showed the equivalence of linear programming and two-person zero-sum games. However, in the description of his reduction from linear programming to zero-sum games, he noted that there was one case in which his reduction does not work. This also led to incomplete proofs of the relationship between the Minimax Theorem of game theory and the Strong Duality Theorem of linear programming. In this note, we fill these gaps."

The next theorem relates the linear programming problem with the notions we have just introduced for the Lagrangian. Its proof will be skipped since it is essentially the same as the proof for the *quadratic programming* problem that will be given in Theorem 7.12 of section 7.6.

[11] What makes this connection really interesting is that "the Minimax Theorem can be proved by rational algebraic means and holds for any ordered field. This was first specifically pointed out and carried through by Hermann Weyl ..." (see the survey paper by H. W. KUHN and A. W. TUCKER [**2**]).

Theorem 7.5. *Consider the primal and dual problems where f, U, g, and V are specified by* (7.21) *and* (7.41)*, and the Lagrangian $L(x,\lambda,\mu)$ by* (7.31)*. Then*

$$\begin{aligned} &\textit{primal problem} && \inf_{x\in U} q\cdot x = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu),\\ &\textit{dual problem} && \sup_{(\lambda,\mu)\in V} -\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\cdot\begin{bmatrix}\lambda\\ \mu\end{bmatrix} = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu),\\ & && \inf_{x\in U} q\cdot x \geq \sup_{(\lambda,\mu)\in V} -\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\cdot\begin{bmatrix}\lambda\\ \mu\end{bmatrix}. \end{aligned} \tag{7.53}$$

(i) *If $U \neq \varnothing$ (primal feasible), then* $\inf f(U) < +\infty$*, and we have strong duality*

$$\inf_{x\in U} q\cdot x = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu), \tag{7.54}$$

and there are two cases:

(a) *if $-q \notin A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k$ (dual infeasible, $V=\varnothing$),*

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = -\infty;$$

(b) *if $-q \in A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k$ (dual feasible, $V\neq\varnothing$), there exist $\hat{x}\in U$ and $(\hat\lambda,\hat\mu)\in \mathbb{R}^m_+\times\mathbb{R}^k$ such that*

$$\forall x\in\mathbb{R}^n,\ \forall(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k,\quad L(\hat{x},\lambda,\mu)\leq L(\hat{x},\hat\lambda,\hat\mu)\leq L(x,\hat\lambda,\hat\mu), \tag{7.55}$$

and $(\hat{x},\hat\lambda,\hat\mu)$ are a solution of the system (7.27) *of Theorem* 7.4(ii).

In case (a)*, U is inbounded.*

(ii) *If $U=\varnothing$ (primal infeasible),* $\inf_{x\in U} q\cdot x = +\infty$*, and we have two cases:*

(c) *if $-q \notin A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k$ (dual infeasible, $V=\varnothing$),*

$$-\infty = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) < \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = +\infty;$$

(d) *if $-q \in A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k$ (dual feasible, $V\neq\varnothing$), we have strong duality*

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = +\infty.$$

In case (d)*, V is inbounded.*

The following simple examples illustrate cases (a) to (d) with the following notation: $X = \mathbb{R}^n$, $Y = \mathbb{R}^m_+ \times \mathbb{R}^k$, S the set of saddle points,

$$X_0 = \left\{ x_0 \in \mathbb{R}^n : \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x_0,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) \right\},$$

$$Y_0 = \left\{ (\lambda_0,\mu_0) \in \mathbb{R}^m_+\times\mathbb{R}^k : \inf_{x\in\mathbb{R}^n} L(x,\lambda_0,\mu_0) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \right\},$$

$$v^+ = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu), \quad \text{and} \quad v^- = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu).$$

Example 7.1 (Case (a) : $v^- = v^+ = -\infty$)**.**
Consider the following one-dimensional problem:

$$\begin{cases} \text{minimize } x \\ \text{subject to } x \le 0. \end{cases}$$

$X = \mathbb{R}$, $Y = \mathbb{R}_+$, $X_0 = \varnothing$, $Y_0 = \mathbb{R}_+$, and $S = \varnothing$. □

Example 7.2 (Case (b) : $v^- = v^+ = 0$)**.**
Consider the following one-dimensional problem:

$$\begin{cases} \text{minimize } x \\ \text{subject to } x = 0. \end{cases}$$

$X = \mathbb{R}$, $Y = \mathbb{R}$, $X_0 = \varnothing$, $X_0 = \{0\}$, $Y_0 = \{-1\}$, and $S = \{(0,-1)\}$. □

Example 7.3 (Case (c): $-\infty = v^- < v^+ = +\infty$)**.**
Consider the following two-dimensional problem:

$$\begin{cases} \text{minimize } x_1 + x_2 \\ \text{subject to } \begin{cases} x_1 - x_2 = 1 \\ -x_1 + x_2 = 2. \end{cases} \end{cases}$$

$X = X_0 = \mathbb{R}^2$, and $Y = Y_0 = \mathbb{R}^2$, $S = \varnothing$. □

Example 7.4 (Case (d): $v^- = v^+ = +\infty$)**.**
Consider the following two-dimensional problem:

$$\begin{cases} \text{minimize } x_1 - x_2 \\ \text{subject to } \begin{cases} x_1 - x_2 = 1 \\ -x_1 + x_2 = 2. \end{cases} \end{cases}$$

$X = \mathbb{R}^2$, $Y = \mathbb{R}^2$, $X_0 = \mathbb{R}^2$, $Y_0 = \varnothing$, and $S = \varnothing$. □

For the primal SLP minimization problem

$$\begin{cases} \text{minimize } & -b\cdot x \\ \text{subject to } & Ax \le c \quad \text{and} \quad x \ge 0, \end{cases} \tag{7.56}$$

$$U \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : Ax \le c \quad \text{and} \quad x \ge 0\right\}, \tag{7.57}$$

its dual maximization problem

$$\begin{cases} \text{maximize } & -c\cdot \lambda \\ \text{subject to } & A^\top \lambda \ge b \quad \text{and} \quad \lambda \ge 0, \end{cases} \tag{7.58}$$

$$V \stackrel{\text{def}}{=} \left\{\lambda \in \mathbb{R}^m : A^\top \lambda \ge b \quad \text{and} \quad \lambda \ge 0\right\}, \tag{7.59}$$

and the Lagrangian

$$(x,\lambda,\lambda_+) \mapsto L(x,\lambda,\lambda_+) = -b\cdot x + \lambda\cdot(Ax-c) - \lambda_+\cdot x : \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^n_+ \to \mathbb{R}$$

or, in its more compact form,

$$(x,\lambda) \mapsto L(x,\lambda) = -b\cdot x + \lambda\cdot(Ax-c) : \mathbb{R}^n_+ \times \mathbb{R}^m_+ \to \mathbb{R}, \tag{7.60}$$

we recover and sharpen the four cases.

(i) If $U \ne \varnothing$, that is, $c \in \mathbb{R}^m_+ + A\,\mathbb{R}^n_+$ (primal feasible), then

(a) if $-b \notin -A^\top\mathbb{R}^m_+ + \mathbb{R}^n_+$ (dual infeasible, $V = \varnothing$),

$$v^- = v^+ = -\infty;$$

(b) if $-b \in -A^\top\mathbb{R}^m_+ + \mathbb{R}^n_+$ (dual feasible, $V \ne \varnothing$), there exists a saddle point $(\hat{x},\hat{\lambda},\hat{\lambda}_+) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^n_+$ and

$$v^- = 0 = v^+. \tag{7.61}$$

(ii) If $U = \varnothing$, that is, $c \notin \mathbb{R}^m_+ + A\,\mathbb{R}^n_+$ (primal infeasible), then

(c) if $-b \notin -A^\top\mathbb{R}^m_+ + \mathbb{R}^n_+$ (dual infeasible, $V = \varnothing$),

$$-\infty = v^- < v^+ = +\infty;$$

(d) if $-b \in -A^\top\mathbb{R}^m_+ + \mathbb{R}^n_+$ (dual feasible, $V \ne \varnothing$),

$$v^- = v^+ = +\infty.$$

In case (a) U is unbounded and in case (d) V is unbounded.

7.4 Some Elements of Two-Person Zero-Sum Games

The introduction of the Lagrange multipliers and the Lagrangian in Theorem 7.5 of section 7.3 for the LP problem provides an alternate approach to the optimization problem in the presence of affine constraints.

Associate with the function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, dom $f \neq \varnothing$, and the set U defined in (7.21), the *Lagrangian*

$$\begin{aligned} (x,\lambda,\mu) \to L(x,\lambda,\mu) &: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}\cup\{+\infty\} \\ L(x,\lambda,\mu) &\overset{\text{def}}{=} f(x) + \sum_{i=1}^{m} \lambda_i \, [a_i \cdot x - \alpha_i] + \sum_{j=1}^{k} \mu_j \left[b_j \cdot x - \beta_j\right], \end{aligned} \tag{7.62}$$

where $\lambda = (\lambda_1, \dots, \lambda_m)$ and $\mu = (\mu_1, \dots, \mu_k)$. The set U of affine constraints is closed and convex (possibly empty). It is readily seen that the primal problem can be written in the form

$$\inf f(U) = \inf_{x\in U} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu), \tag{7.63}$$

since any $x \notin U$ will make the sup equal to $+\infty$. If $U \cap \text{dom}\, f = \varnothing$, then the sup is always $+\infty$ and $\inf f(U) = +\infty$ in good agreement with our convention of Definition 3.1(ii) in Chapter 2; if $U \cap \text{dom}\, f \neq \varnothing$, then $\inf f(U) < +\infty$.

We can look at this problem as a *two-player zero-sum game*. The *first player* wants to minimize the *utility function* $L(x,\lambda,\mu)$ using *strategies* $x \in U$ and the *second player* wants to minimize the *utility function* $-L(x,\lambda,\mu)$ using strategies $(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k$. This explains this terminology.

Such games have been studied by Émile Borel.[12] The modern epoch of game theory began with the statement and proof of the minimax theorem for zero-sum games in 1928 by John von Neumann[13] followed by the influential book of J. von Neumann and O. Morgenstern [**1**] in 1944. In game theory the extrema are replaced by several notions of equilibria. The one that came up in Theorem 7.5(i)(b) for LP is the one of J. F. Nash [**1**][14] introduced in 1951. The equivalence of LP and two-person zero-sum games was shown in 1951 by G. B. Dantzig [**2**].[15]

The following standard notions are associated with such a game.

Definition 7.2.
Associate with a function $G : X \times Y \to \mathbb{R}$ the following notation:

$$v^+ \overset{\text{def}}{=} \inf_{x\in X} \sup_{y\in Y} G(x,y), \quad X_0 \overset{\text{def}}{=} \left\{ x \in X : \sup_{y\in Y} G(x,y) = v^+ \right\}, \tag{7.64}$$

$$v^- \overset{\text{def}}{=} \sup_{y\in Y} \inf_{x\in X} G(x,y), \quad Y_0 \overset{\text{def}}{=} \left\{ y \in Y : \inf_{x\in X} G(x,y) = v^- \right\}. \tag{7.65}$$

[12] See, for instance, the book of É. Borel [**1**] in 1938.

[13] John von Neumann (1903–1957), J. von Neumann [**1**].

[14] John Forbes Nash, Jr. (1928–). He shared the 1994 Nobel Memorial Prize in Economic Sciences with game theorists Reinhard Selten and John Harsanyi.

[15] George Bernard Dantzig (1914–2005).

(i) The game has an *upper value* if v^+ is finite; it has a *lower value* if v^- is finite; and it has a *value* if it has a lower and an upper value and $v^+ = v^-$.

(ii) The game has a *saddle point* if there exists $(\hat{x}, \hat{y}) \in X \times Y$ such that

$$\forall x \in X, \forall y \in Y, \quad G(\hat{x}, y) \leq G(\hat{x}, \hat{y}) \leq G(x, \hat{y}). \tag{7.66}$$

The set of all saddle points of G will be denoted by S. □

Theorem 7.6. *Let $G : X \times Y \to \mathbb{R}$.*

(i) *The inequality $v^- \leq v^+$ is always verified and*

$$\forall (x_0, y_0) \in X_0 \times Y_0, \quad v^- \leq G(x_0, y_0) \leq v^+. \tag{7.67}$$

(ii) *We always have $S \subset X_0 \times Y_0$ and, if $S \neq \varnothing$, $v^- = v^+$.*

(iii) *If $v^- = v^+$, then $X_0 \times Y_0 \subset S$.*

(iv) *If $X_0 \times Y_0 \neq \varnothing$, then*

$$v^- = v^+ \quad \Longleftrightarrow \quad S = X_0 \times Y_0 \tag{7.68}$$

$$v^- < v^+ \quad \Longleftrightarrow \quad S = \varnothing. \tag{7.69}$$

If $X_0 \times Y_0 = \varnothing$,[16] *then $S = \varnothing$.*

Remark 7.1.
The case $X_0 \times Y_0 \neq \varnothing$, $v^- < v^+$, and $S = \varnothing$ can occur. Consider the sets $X = Y = \mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ and the function $G(x, y) = y - x$:

$$\inf_{x \in X} y - x = -\infty \text{ and } v^- = \sup_{y \in Y} \inf_{x \in X} y - x = -\infty \quad \Rightarrow Y_0 = Y = \mathbb{R}_+$$

$$\sup_{y \in Y} y - x = +\infty \text{ and } v^+ = \inf_{x \in X} \sup_{y \in Y} y - x = +\infty \quad \Rightarrow X_0 = X = \mathbb{R}_+.$$

□

Proof. (i) By definition of the infimum $\inf_{x \in X} G(x, y) \leq G(x, y)$,

$$\sup_{y \in Y} \inf_{x \in X} G(x, y) \leq \sup_{y \in Y} G(x, y) \quad \Rightarrow v^- = \sup_{y \in Y} \inf_{x \in X} G(x, y) \leq \inf_{x \in X} \sup_{y \in Y} G(x, y) = v^+.$$

If $X_0 \times Y_0 = \varnothing$, there is nothing to prove. If there exist $x_0 \in X_0$ and $y_0 \in Y_0$, then by definition of X_0 and Y_0,

$$\inf_{x \in X} G(x, y_0) \leq G(x_0, y_0) \leq \sup_{y \in Y} G(x_0, y)$$

$$\Rightarrow v^- = \sup_{y \in Y} \inf_{x \in X} G(x, y) = \inf_{x \in X} G(x, y_0) \leq G(x_0, y_0) \leq \sup_{y \in Y} G(x_0, y)$$

$$\sup_{y \in Y} G(x_0, y) = \inf_{x \in X} \sup_{y \in Y} G(x, y) = v^+$$

$$\Rightarrow v^- = \inf_{x \in X} G(x, y_0) \leq G(x_0, y_0) \leq \sup_{y \in Y} G(x_0, y) = v^+. \tag{7.70}$$

[16] $X_0 \times Y_0 = \varnothing$ means that either $X_0 = \varnothing$, or $Y_0 = \varnothing$, or both are empty.

(ii) If $S = \varnothing$, then $S = \varnothing \subset X_0 \times Y_0$. If there exists $(x_0, y_0) \in S$, then

$$v^+ = \inf_{x \in X} \sup_{y \in Y} G(x,y) \le \sup_{y \in Y} G(x_0, y) \le G(x_0, y_0)$$
$$G(x_0, y_0) \le \inf_{x \in X} G(x, y_0) \le \sup_{y \in Y} \inf_{x \in X} G(x,y) = v^- \le v^+$$
$$\Rightarrow v^+ = \sup_{y \in Y} G(x_0, y) = G(x_0, y_0) = \inf_{x \in X} G(x, y_0) = v^-,$$

and, by the definitions of X_0 and Y_0, $(x_0, y_0) \in X_0 \times Y_0$. Hence $S \subset X_0 \times Y_0$.

(iii) If $X_0 \times Y_0 = \varnothing$, from part (ii) $S = \varnothing$. If $X_0 \times Y_0 \ne \varnothing$ and if $v^- = v^+$, then for all $(x_0, y_0) \in X_0 \times Y_0$,

$$G(x_0, y_0) \le \sup_{y \in Y} G(x_0, y) = v^+ = v^- = \inf_{x \in X} G(x, y_0) \le G(x_0, y_0)$$
$$\Rightarrow \sup_{y \in Y} G(x_0, y) = G(x_0, y_0) = \inf_{x \in X} G(x, y_0)$$
$$\Rightarrow \forall x \in X, \forall y \in Y, \quad G(x_0, y) \le G(x_0, y_0) \le G(x, y_0)$$

and (x_0, y_0) is a saddle point of G and $X_0 \times Y_0 \subset S$.

(iv) From (i), (ii), and (iii). □

Theorem 7.7. *Let $G : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$, $N \ge 1$ an integer, and $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^N$ two nonempty convex sets with the following assumptions:*

(a) *for each $y \in Y$, $x \mapsto G(x,y)$ is semidifferentiable on $\mathbb{R}^n$ and convex on X;*

(b) *for each $x \in X$, $y \mapsto G(x,y)$ is semidifferentiable on $\mathbb{R}^N$ and concave on Y.*

Then the following statements are equivalent:

(i) *G has a saddle point in $X \times Y$: there exists $(\hat{x}, \hat{y}) \in X \times Y$ such that*

$$\forall x \in X, \forall y \in Y, \quad G(\hat{x}, y) \le G(\hat{x}, \hat{y}) \le G(x, \hat{y}); \tag{7.71}$$

(ii) *there exists $(\hat{x}, \hat{y}) \in X \times Y$ such that*

$$\forall x \in X,\ dG(\hat{x}, \hat{y}; x - \hat{x}, 0) \ge 0 \ \text{ and } \ \forall y \in Y,\ dG(\hat{x}, \hat{y}; 0, y - \hat{y}) \le 0. \tag{7.72}$$

Proof. (i) $\Rightarrow$ (ii) . If $(\hat{x}, \hat{y}) \in X \times Y$ is a saddle point of G in $X \times Y$, then

$$-\inf_{y \in Y} -G(\hat{x}, y) = \sup_{y \in Y} G(\hat{x}, y) = G(\hat{x}, \hat{y}) = \inf_{x \in X} G(x, \hat{y}).$$

Since the map $x \mapsto G(x, \hat{y})$ is semidifferentiable on $\mathbb{R}^n$ and convex on X,

$$\forall x \in X, \quad dG(\hat{x}, \hat{y}; x - \hat{x}, 0) \ge 0$$

by Theorem 3.4 or its general version Theorem 3.5. Similarly, since the map $y \mapsto -G(\hat{x}, y)$ is semidifferentiable on $\mathbb{R}^N$ and convex on Y,

$$\forall y \in Y, \quad -dG(\hat{x}, \hat{y}; 0, y - \hat{y}) \le 0$$

by the same theorem.

(ii) $\Rightarrow$ (i) Given a semidifferentiable function convex on a convex U,

$$\begin{aligned}
&\forall \varepsilon, 0 < \varepsilon \le 1, \forall x, y \in U, \quad f(x+\varepsilon(y-x)) \le f(x) + \varepsilon[f(y) - f(x)] \\
\Rightarrow\ & \frac{f(x+\varepsilon(y-x)) - f(x)}{\varepsilon} \le [f(y) - f(x)] \quad \Rightarrow df(x; y-x) \le f(y) - f(x) \\
&\Rightarrow \forall x, y \in U, \quad f(y) \ge f(x) + df(x; y-x).
\end{aligned}$$

As a result, for all $x \in X$,

$$G(x, \hat{y}) \ge G(\hat{x}, \hat{y}) + dG(\hat{x}, \hat{y}; x - \hat{x}, 0) \ge G(\hat{x}, \hat{y})$$

and for all $y \in Y$,

$$-G(\hat{x}, y) \ge -G(\hat{x}, \hat{y}) - dG(\hat{x}, \hat{y}; 0, y - \hat{y}) \ge -G(\hat{x}, \hat{y}).$$

From the above two inequalities,

$$\forall (x, y) \in X \times Y, \quad G(\hat{x}, y) \le G(\hat{x}, \hat{y}) \le G(x, \hat{y})$$

and, by definition, $(\hat{x}, \hat{y})$ is a saddle point of G in $X \times Y$. □

7.5 Fenchel Primal and Dual Problems and the Lagrangian

As we have seen in (7.63), the primal problem which coincides with the inf sup problem is related to the notion of upper value of the game; the sup inf problem

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+ \times \mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x, \lambda, \mu) \tag{7.73}$$

is related to the lower value of the game. In general,

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+ \times \mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x, \lambda, \mu) \le \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+ \times \mathbb{R}^k} L(x, \lambda, \mu). \tag{7.74}$$

For the LP problem, we have seen in Theorem 7.5 that when $U \neq \varnothing$, the two values are equal. When they are finite, we have a saddle point; when they are equal to $-\infty$, it corresponds to an unbounded U without minimizer; when they are equal to $+\infty$, U is empty and the problem is infeasible.

Recall that we have shown in (7.34) that

$$I_U(x) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+ \times \mathbb{R}^k} [(Ax - \alpha) \cdot \lambda + (Bx - \beta) \cdot \mu] \tag{7.75}$$

and that

$$\begin{aligned}
&\sup_{(\lambda,\mu)\in\mathbb{R}^m_+ \times \mathbb{R}^k} L(x, \lambda, \mu) = q \cdot x + I_U(x) \\
\Rightarrow\ & \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+ \times \mathbb{R}^k} L(x, \lambda, \mu) = \inf_{x\in\mathbb{R}^n} q \cdot x + I_U(x) \\
&\qquad = -\sup_{x\in\mathbb{R}^n} -q \cdot x - I_U(x) = -I_U^*(-q).
\end{aligned}$$

From Example 9.2 of Chapter 2, the Fenchel–Legendre transform of $f(x) = q \cdot x$ is

$$f^*(x^*) = \begin{cases} 0, & \text{if } x^* = q \\ +\infty, & \text{if } x^* \neq q. \end{cases}$$

As a result, for $U \neq \varnothing$,

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \sup_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} L(x,\lambda,\mu) &= -I^*_U(-q) \\ &= -\inf_{x^* \in \mathbb{R}^n} \left[f^*(-x^*) + I^*_U(x^*) \right] \\ \Rightarrow \inf_{x \in \mathbb{R}^n} [f(x) + I_U(x)] &= -\inf_{x^* \in \mathbb{R}^n} \left[f^*(-x^*) + I^*_U(x^*) \right] \end{aligned}$$

and we recognize the Fenchel primal and dual problems of Chapter 2. The assumption $U \neq \varnothing$ is necessary. Otherwise, we would have $I^*_U(x^*) = -\infty$ and, for $x^* \neq -q$, the sum $f^*(-x^*) + I^*_U(x^*) = +\infty + (-\infty)$ would be indefinite.

In order to relate the results of Theorem 7.5 with the Fenchel primal and dual problems[17] of Theorem 9.7 of Chapter 2, we now express the indicator function I_U and its Fenchel–Legendre transform in terms of the Lagrange multipliers. Fortunately, this comes as a corollary to Theorem 7.5(i).

Theorem 7.8. *Let U be as specified by* (7.21). *If $U \neq \varnothing$, the support function $\sigma_U = I^*_U : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by*

$$\begin{aligned} \sigma_U(x^*) &= \inf_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} \sup_{x \in \mathbb{R}^n} \bar{\ell}(x,\lambda,\mu) = \sup_{x \in \mathbb{R}^n} \inf_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} \bar{\ell}(x,\lambda,\mu) \\ &= \begin{cases} \displaystyle\inf_{\substack{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k \\ x^* = A^\top \lambda + B^\top \mu}} (\alpha \cdot \lambda + \beta \cdot \mu), & \text{if } x^* \in A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k \\ +\infty, & \text{if } x^* \notin A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k \end{cases} \end{aligned} \tag{7.76}$$

$$\bar{\ell}(x,\lambda,\mu) \overset{\text{def}}{=} x^* \cdot x - (Ax - \alpha) \cdot \lambda - (Bx - \beta) \cdot \mu \tag{7.77}$$

$$\operatorname{dom} \sigma_U = A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k. \tag{7.78}$$

If $x^* \in A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k$, there exist $\hat{x} \in U$ and $(\hat{\lambda}, \hat{\mu}) \in \mathbb{R}^m_+ \times \mathbb{R}^k$ such that

$$\forall x \in \mathbb{R}^n, \forall (\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k, \quad \bar{\ell}(x,\hat{\lambda},\hat{\mu}) \le \bar{\ell}(\hat{x},\hat{\lambda},\hat{\mu}) \le \bar{\ell}(\hat{x},\lambda,\mu), \tag{7.79}$$

$$\sigma_U(x^*) = \bar{\ell}(\hat{x},\hat{\lambda},\hat{\mu}). \tag{7.80}$$

Proof. This is the LP problem for a supremum instead of an infimum. In the case of the supremum,

$$\sigma_U(x^*) = \sup_{x \in U} x^* \cdot x = -\inf_{x \in U} (-x^*) \cdot x$$

and the results follow from Theorem 7.5(i) with $q = -x^*$. □

[17] It will be convenient to use the form (9.45) of the dual problem following Theorem 9.7 of Chapter 2 along with the version of the Fenchel duality theorem given in Theorem 9.8 of Chapter 2 in order to avoid extra notation.

We can now connect the Lagrange and Fenchel formulations for a much larger class of functions.

Theorem 7.9. *Let* $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, dom $f \neq \varnothing$, $U \neq \varnothing$ *the set defined in* (7.21), *and* L *the Lagrangian* (7.62).

(i) *The upper problem for* L *coincides with the Fenchel primal problem*

$$\inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \inf f(U) = \inf_{x\in\mathbb{R}^n} [f(x)+I_U(x)], \tag{7.81}$$

$$I_U(x) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} (Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu. \tag{7.82}$$

(ii) *The lower problem for* L *coincides with the Fenchel dual problem*

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = -\inf_{x^*\in\mathbb{R}^n} \left[f^*\left(-x^*\right) + I_U^*(x^*)\right], \tag{7.83}$$

$$I_U^*(x^*) = \begin{cases} \displaystyle\inf_{\substack{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k \\ x^*=A^\top\lambda+B^\top\mu}} (\alpha\cdot\lambda+\beta\cdot\mu), & \text{if } x^*\in A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k \\ +\infty, & \text{if } x^*\notin A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k \end{cases} \tag{7.84}$$
$$\operatorname{dom} I_U^* = A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k,$$

where f^* *is the Fenchel–Legendre transform of* f.

Remark 7.2.
In Theorems 9.7 and 9.8 of Chapter 2 pick $g = I_U$ and $f = f$. □

Proof. (i) For $x \notin U \cap \operatorname{dom} f$, it is equal to $+\infty$ and for $x \in U \cap \operatorname{dom} f$, it is bounded above by zero. It could be negative only if $a_i\cdot x - \alpha_i < 0$ for some i in which case we pick $\lambda_i = 0$. So $g = I_U$.

(ii) By direct computation,

$$f(x) + (Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu = -[\alpha\cdot\lambda+\beta\cdot\mu] + \left[A^\top\lambda + B^\top\mu\cdot x - f(x)\right].$$

As a result,

$$\begin{aligned}\inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) &= -[\alpha\cdot\lambda+\beta\cdot\mu] - \sup_{x\in\mathbb{R}^n}\left\{-\left[A^\top\lambda+B^\top\mu\right]\cdot x - f(x)\right\} \\ &= -[\alpha\cdot\lambda+\beta\cdot\mu] - f^*\left(-A^\top\lambda - B^\top\mu\right).\end{aligned}$$

From Theorem 7.8, we recognize I_U^*:

$$\begin{aligned}&\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \\ &= \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \left\{-[\alpha\cdot\lambda+\beta\cdot\mu] - f^*\left(-A^\top\lambda - B^\top\mu\right)\right\}\end{aligned}$$

$$
\begin{aligned}
&= -\inf_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \left[\alpha\cdot\lambda+\beta\cdot\mu+f^*\left(-A^\top\lambda-B^\top\mu\right)\right] \\
&= -\inf_{x^*\in A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k} \left[\inf_{\substack{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k \\ x^*=A^\top\lambda+B^\top\mu}} (\alpha\cdot\lambda+\beta\cdot\mu)+f^*\left(-x^*\right)\right] \\
&= -\inf_{x^*\in \operatorname{dom} I_U^*} \left[I_U^*(x^*)+f^*\left(-x^*\right)\right] = -\inf_{x^*\in\mathbb{R}^n} \left[I_U^*(x^*)+f^*\left(-x^*\right)\right],
\end{aligned}
$$

since $I_U^*(x^*)=+\infty$ if $x^*\notin \operatorname{dom} I_U^*$ and $\operatorname{dom} I_U^* = A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k$. □

7.6 Quadratic Programming Problem

The results of the LP problem readily extend to a quadratic objective function with U specified by a finite number of equality or inequality constraints on affine functions.

Let $n\geq 1$, $m\geq 1$, and $k\geq 1$ be three integers, $a_1,a_2,\dots,a_m$, $b_1,b_2,\dots,b_k$ be vectors in $\mathbb{R}^n$, and $\alpha_1,\alpha_2,\dots,\alpha_m$, $\beta_1,\beta_2,\dots,\beta_k$, be scalars. Let A and B be the matrices and α and β be the vectors introduced in (7.28). Consider the following *quadratic programming* (QP) problem:

$$
\inf_{x\in U} f(x),\ f(x)\stackrel{\text{def}}{=}\frac{1}{2}Qx\cdot x+q\cdot x,\ U\stackrel{\text{def}}{=}\left\{x\in\mathbb{R}^n : Ax\leq\alpha \quad\text{and}\quad Bx=\beta\right\} \tag{7.85}
$$

for a symmetric $n\times n$ matrix Q and $q\in\mathbb{R}^n$.

7.6.1 Theorem of Frank–Wolfe

The first fundamental result[18] by M. FRANK and P. WOLFE [**1**] in 1956 extends Theorem 3.3 to polyhedral convex domains. In that paper the authors introduce the *reduced gradient method*, an iterative algorithm for QP. They prove that when minimizing a quadratic function over a nonempty convex polyhedral domain, only two cases are possible: either there exists a minimizer in U or the infimum is $-\infty$. In general, this is not true for a convex objective function. Indeed, the infimum $\inf f(U)$ of $f(x)=e^{-x}$ with respect to $U=\{x\in\mathbb{R} : x\geq 0\}$ is equal to 0, but the set of minimizers $\operatorname{argmin} f(U)$ is empty.

Theorem 7.10. *Let U and f be as specified in* (7.85) *with $q\in\mathbb{R}^n$ and Q a symmetric matrix. Then*

$$
U\neq\varnothing \quad\textit{and}\quad \inf f(U)>-\infty \quad\Longleftrightarrow\quad \exists\hat{x}\in U \textit{ such that } f(\hat{x})=\inf f(U).
$$

As a consequence, if $U\neq\varnothing$ and $\operatorname{argmin} f(U)=\varnothing$*, we have* $\inf f(U)=-\infty$.

Proof. Since each equality constraint $b_j\cdot x-\beta_j=0$ is equivalent to two inequality constraints $b_j\cdot x-\beta_j\leq 0$ and $-b_j\cdot x+\beta_j\leq 0$, we can, without loss of generality, prove

[18] We give the longer but more elementary proof of E. BLUM and W. OETTLI [**1**] for the general case. For the convex case, see the proof of L. COLLATZ and W. WETTERLING [**1**] in 1966 via Lagrange multipliers. Lothar Collatz (1910–1990). Marguerite Josephine Straus Frank. Philip Wolfe.

the theorem only for inequality constraints: $Ax - \alpha \le 0$. We also introduce some notation: $I = \{1,\dots,m\}$, $g_i(x) = a_i \cdot x - \alpha_i$. As $U \neq \varnothing$, choose a point $x_0 \in U$ and define for each $\rho \ge 0$:

$$U_\rho \stackrel{\text{def}}{=} \{x \in U : \|x - x_0\| \le \rho\}, \quad m(\rho) \stackrel{\text{def}}{=} \inf f(U_\rho), \quad m \stackrel{\text{def}}{=} \inf f(U),$$
$$S_\rho \stackrel{\text{def}}{=} \left\{x \in U_\rho : f(x) = m(\rho)\right\}.$$

As U_ρ is nonempty and compact, the minimum of f is achieved in U_ρ and the section S_ρ is also nonempty and compact. Therefore, there exists $x_\rho \in S_\rho$ of minimal norm:

$$x_\rho \in U_\rho, \quad \|x_\rho\| = \inf_{x \in S_\rho} \|x\| \text{ and } f(x_\rho) = m(\rho).$$

The proof proceeds in two steps.

(i) We first prove the following property:

$$\text{(P)} \qquad \exists \bar\rho > 0 \text{ such that } \forall \rho > \bar\rho, \quad \|x_\rho\| < \rho. \tag{7.86}$$

If (P) were not true, then

$$\forall n \ge 1, \exists \rho_n > n \text{ such that } \|x_{\rho_n}\| = n.$$

To keep the notation simple, we shall write x_n in place of x_{ρ_n}. By construction $x_n \in U$ and $g_i(x_n) \le 0$ for each i. Let

$$I_0 \stackrel{\text{def}}{=} \left\{ i \in I : \limsup_{n\to\infty} g_i(x_n) = 0 \right\}.$$

For the other indices i, we have $\limsup_{n\to\infty} g_i(x_n) < 0$. Therefore, there exists $\varepsilon > 0$ such that

$$\forall i \in I \backslash I_0, \quad \limsup_{n\to\infty} g_i(x_n) \le -\varepsilon.$$

Since $\|x_n\|/n = 1$, there exists s, $\|s\| = 1$, and a subsequence, still labeled $\{x_n\}$, such that

$$s_n \stackrel{\text{def}}{=} \frac{x_n}{n} \to s \text{ and for all } i,\ g_i(x_n) \to \limsup_{n\to\infty} g_i(x_n) \text{ as } n \to \infty.$$

We have

$$\begin{aligned} \forall i \in I_0, &\quad g_i(x_n) = n\, a_i \cdot s_n - \alpha_i \to 0 &&\Rightarrow \boxed{a_i \cdot s = 0} \\ \forall i \in I\backslash I_0, &\quad \lim_{n\to\infty} (n\, a_i \cdot s_n - \alpha_i) \le -\varepsilon &&\Rightarrow \boxed{a_i \cdot s \le 0.} \end{aligned}$$

On the other hand,

$$m(\rho_n) = f(x_n) = f(n\, s_n) = \frac{n^2}{2} Q s_n \cdot s_n + n\, q \cdot s_n \to m \quad \Rightarrow \boxed{Qs \cdot s = 0}$$

since m is finite.

Now consider the points of the form $x_n + \lambda s$, $\lambda \geq 0$. We have

$$\forall i \in I, \quad g_i(x_n + \lambda s) = g_i(x_n) + \lambda\, a_i \cdot s \leq 0 \quad \Rightarrow \boxed{\forall \lambda \geq 0, x_n + \lambda s \in U}$$

since $a_i \cdot s \leq 0$ for all i. As f is quadratic and $Qs \cdot s = 0$,

$$f(x_n + \lambda s) = f(x_n) + (Qx_n + q) \cdot (\lambda s) + \frac{\lambda^2}{2} Qs \cdot s = f(x_n) + \lambda (Qx_n + q) \cdot s.$$

If we had $(Qx_n + q) \cdot s < 0$, we could construct the sequence $\{x_n + ks : k \geq 1\}$ in U such that $\lim_{k\to\infty} f(x_n + ks) = -\infty$, which would contradict the assumption on the finiteness of the infimum m. Therefore

$$\boxed{\forall \lambda \geq 0, \forall n, \quad (Qx_n + q) \cdot s \geq 0.}$$

Now consider points of the form $x_n - \lambda s$, $\lambda \geq 0$, that do not necessarily belong to U_{ρ_n}. We have

$$\begin{aligned}
\forall i \in I_0, \quad & g_i(x_n - \lambda s) = g_i(x_n) - \lambda\, a_i \cdot s = g_i(x_n) \leq 0, \\
\forall i \in I \backslash I_0, \quad & g_i(x_n - \lambda s) = g_i(x_n) - \lambda\, a_i \cdot s \leq -\varepsilon + \lambda \|a_i\|.
\end{aligned}$$

Hence $x_n - \lambda s \in U$ if

$$0 \leq \lambda < \delta_1 \overset{\text{def}}{=} \varepsilon \min_{1 \leq i \leq m} \{(1 + \|a_i\|)^{-1}\}.$$

At last,

$$\|x_n - \lambda s\|^2 = \|x_n\|^2 - 2\lambda\, x_n \cdot s + \lambda^2 \|s\|^2 = \|x_n\|^2 - \lambda (2x_n \cdot s - \lambda).$$

As $x_n/n \to s$, $\|s\| = 1$, there exists $\bar{n}$ such that

$$\begin{aligned}
\forall n > \bar{n}, \quad & \left\| s - \frac{x_n}{n} \right\|^2 < 1 \quad \Rightarrow 2\, x_n \cdot s > n \\
\Rightarrow \forall n > \bar{n}, \forall 0 < \lambda < n, \quad & \|x_n - \lambda s\| < \|x_n\| = n \\
\Rightarrow & \boxed{\forall 0 < \lambda < \bar{n}, \forall n > \bar{n}, \quad \|x_n - \lambda s\| < \|x_n\| = n.}
\end{aligned}$$

Finally

$$\forall 0 < \lambda < \delta \overset{\text{def}}{=} \min\{\delta_1, \bar{n}\}, \forall n > \bar{n}, \quad x_n - \lambda s \in U_{\rho_n}.$$

Since $Qs \cdot s = 0$ and $(Qx_n + q) \cdot s \geq 0$ for all $n \geq 1$ and $\lambda \geq 0$, we have

$$f(x_n - \lambda s) = f(x_n) + (Qx_n + q) \cdot (-\lambda s) = f(x_n) - \lambda (Qx_n + q) \cdot s \leq f(x_n)$$

and for $0 < \lambda < \delta$ and $n > \bar{n}$,

$$x_n - \lambda s \in U_{\rho_n} \quad \text{and} \quad \|x_n - \lambda s\| < \|x_n\| = n.$$

As $x_n \in U_{\rho_n}$ is a minimizer of f with respect to U_{ρ_n}, we have

$$f(x_n - \lambda s) = f(x_n) = m(\rho_n).$$

On the other hand, we know that x_n is a point of $S_{\rho_n} = \{x \in U_{\rho_n} : f(x) = m(\rho_n)\}$ of minimum norm. We then get a contradiction since we have constructed points $x_n - \lambda s \in S_{\rho_n}$ such that $\|x_n - \lambda s\| < \|x_n\|$. This contradiction completes the first part of the proof.

(ii) We now prove that property (P) implies the existence of a ρ such that $m(\rho) = m$, which yields the existence since

$$x_\rho \in U \text{ and } f(x_\rho) = \inf f(U_\rho) = m(\rho) = m = \inf f(U).$$

As the function $\rho \mapsto m(\rho)$ is monotone nonincreasing and goes to the finite limit m, we can find $\bar{\rho} < \rho_1 < \rho_2$ such that $m(\rho_1) > m(\rho_2)$. From property (P), $\|x_{\rho_2}\| < \rho_2$ and, since $m(\rho_1) > m(\rho_2)$, we get $\rho_1 < \|x_{\rho_2}\|$. Choose $\rho_3 = \|x_{\rho_2}\|$, which implies that $\rho_1 < \rho_3 = \|x_{\rho_2}\| < \rho_2$. This yields the following two properties: (a) $\|x_{\rho_3}\| < \rho_3 = \|x_{\rho_2}\|$ since $\bar{\rho} < \rho_3$, and (b) $f(x_{\rho_2}) \le f(x_{\rho_3})$, since $\rho_2 > \rho_3$.

If $f(x_{\rho_2}) = f(x_{\rho_3})$, then $x_{\rho_3} \in U_{\rho_2}$ since $\|x_{\rho_3}\| < \rho_3 = \|x_{\rho_2}\| < \rho_2$ and inequality $\|x_{\rho_3}\| < \|x_{\rho_2}\|$ contradicts the fact that, by definition, x_{ρ_2} is an element of minimum norm in S_{ρ_2}. Hence, $f(x_{\rho_2}) < f(x_{\rho_3})$. As $\|x_{\rho_3}\| < \rho_3 = \|x_{\rho_2}\|$, we have $x_{\rho_2} \in U_{\rho_3}$ and we obtain the contradiction

$$x_{\rho_2} \in U_{\rho_3}, \quad f(x_{\rho_2}) < f(x_{\rho_3}) = \inf f(U_{\rho_3}).$$

This final contradiction completes the proof. □

7.6.2 Nonconvex Objective Function

If the matrix Q is positive semidefinite on $\mathbb{R}^n$, f is convex as in the case of LP for which $Q = 0$, but, in general, f is not convex. When the matrix Q is not positive semidefinite over all points of $\mathbb{R}^n$, there exists $y \in \mathbb{R}^n$ such that $Qy \cdot y < 0$ and necessarily

$$\forall \lambda \in \mathbb{R}^m, \quad \inf_{x \in \mathbb{R}^n} L(x, \lambda) = -\infty \quad \Rightarrow \quad \sup_{(\lambda \in, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) = -\infty,$$

$$L(x, \lambda, \mu) \overset{\text{def}}{=} \frac{1}{2} Qx \cdot x + q \cdot x + \lambda \cdot [Ax - \alpha] + \mu \cdot [Bx - \beta].$$

So we cannot use duality or the theory of saddle points associated with the Lagrangian. As for the primal problem, from Theorem 7.10,

$$\inf_{x \in \mathbb{R}^n} \sup_{(\lambda \in, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} L(x, \lambda, \mu) = \inf f(U) = \begin{cases} \inf f(U) < +\infty, & \text{if } U \neq \varnothing \\ +\infty, & \text{if } U = \varnothing. \end{cases}$$

The difficulty is that the necessary and sufficient condition in the convex case (Q positive semidefinite) is now only necessary and that a second condition is required to get the existence of a minimizing solution in U.

In this section, we consider the quadratic function (7.85) for an $n \times n$ symmetric matrix Q but with only equality constraints. As in the case of LP, we obtain sharper conditions than those of the classical Lagrange multipliers theorem. This problem arises in discretizing the Stokes equations, where the matrix B is associated with the discretization of the divergence operator (see the papers of M. FORTIN and Z. MGHAZLI [**1**, **2**]). The strict positivity of Q on Ker B is essential to construct mixed finite element schemes (see Exercise 5.7). The inequality case would correspond to an expanding or a contracting fluid through a positivity or a negativity constraint on the divergence of the fluid velocity. The following theorem generalizes Theorem 3.3.

Theorem 7.11 (cf. the marginal function of Example 9.7 in Chapter 2)**.** *Let $q \in \mathbb{R}^n$, Q be an $n \times n$ symmetric matrix, f be as specified by* (7.85) *and*

$$U \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : Bx = \beta\right\} \tag{7.87}$$

for some vector $\beta \in \mathbb{R}^m$ and some $m \times n$ matrix B.

(i) *There exists a minimizer of f on U if and only if there exists $(\hat{x}, \hat{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$ such that*

$$\text{(a)} \begin{bmatrix} Q & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} -q \\ \beta \end{bmatrix}, \quad \text{(b) } \forall h \in \operatorname{Ker} B,\ Qh \cdot h \geq 0.$$

(ii) *There exists a unique minimizer of f on U if and only if there exists $(\hat{x}, \hat{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$ such that*

$$\text{(a)} \begin{bmatrix} Q & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} -q \\ \beta \end{bmatrix}, \quad \text{(b) } \forall h \in \operatorname{Ker} B,\ h \neq 0,\ Qh \cdot h > 0.$$

Proof. (i) If $\hat{x} \in U$ is a minimizer, then $f(x) - f(\hat{x}) \geq 0$ for all $x \in U$. Since f is quadratic,

$$f(x) - f(\hat{x}) = (Q\hat{x} + q) \cdot (x - \hat{x}) + \frac{1}{2} Q(x - \hat{x}) \cdot (x - \hat{x}). \tag{7.88}$$

For $t > 0$ and $h \in \operatorname{Ker} B$, $\hat{x} \pm th \in U$ since $B(\hat{x} \pm th) = B(\hat{x}) \pm t\, Bh = \beta$. Substituting above and dividing by $t > 0$,

$$0 \leq \pm(Q\hat{x} + q) \cdot h + \frac{t}{2} Qh \cdot h.$$

By letting t go to zero,

$$\begin{aligned} &\forall h \in \operatorname{Ker} B, \quad 0 \leq \pm(Q\hat{x} + q) \cdot h &&\Rightarrow \forall h \in \operatorname{Ker} B, \quad (Q\hat{x} + q) \cdot h = 0 \\ &\Rightarrow\ Q\hat{x} + q \in (\operatorname{Ker} B)^\perp = \operatorname{Im} B^\top &&\Rightarrow \exists \hat{\lambda} \in \mathbb{R}^m \text{ such that } Q\hat{x} + B^\top \hat{\lambda} = q. \end{aligned}$$

The second equation $B\hat{x} = \beta$ results from the fact that $\hat{x} \in U$. As for condition (b), for each $h \in \operatorname{Ker} B$, $x = \hat{x} + h \in U$. Substitute in (7.88),

$$\forall h \in \operatorname{Ker} B, \quad 0 \leq (Q\hat{x} + q) \cdot h + \frac{1}{2} Qh \cdot h = \frac{1}{2} Qh \cdot h. \tag{7.89}$$

Conversely, from the fact that for $x \in U$, $x - \hat{x} \in \operatorname{Ker} B$ and the fact that

$$Q\hat{x} + B^\top \hat{\lambda} = q \quad \Rightarrow Q\hat{x} + q = -B^\top \hat{\lambda} \in \operatorname{Im} B^\top = (\operatorname{Ker} B)^\perp,$$

we get $(Q\hat{x} + q) \cdot (x - \hat{x}) = 0$. Going back to (7.88), for all $x \in U$,

$$f(x) - f(\hat{x}) = (Q\hat{x} + q) \cdot (x - \hat{x}) + \frac{1}{2} Q(x - \hat{x}) \cdot (x - \hat{x}) = \frac{1}{2} Q(x - \hat{x}) \cdot (x - \hat{x}) \geq 0$$

from condition (b).

(ii) Same proof, but with strict inequalities (see Exercise 9.7). □

7.6.3 Convex Objective Function

When the $n \times n$ symmetric matrix Q is positive semidefinite, the objective function f is convex and we can use the necessary and sufficient condition (3.4) of Theorem 3.2, which yields

$$\exists x \in U \text{ such that } (Qx + q) \cdot (y - x) \geq 0 \text{ for all } y \in U, \tag{7.90}$$

which is equivalent to

$$\exists x \in U \text{ such that } (Qx + q) \cdot h \geq 0 \text{ for all } h \in T_U(x) \quad \text{or} \quad Qx + q \in T_U(x)^*.$$

The *Lagrangian* is given by

$$L(x, \lambda, \mu) \stackrel{\text{def}}{=} \frac{1}{2} Qx \cdot x + q \cdot x + (Ax - \alpha) \cdot \lambda + (Bx - \beta) \cdot \mu. \tag{7.91}$$

The Fenchel–Legendre transform of f (Example 9.3 of Chapter 2) is

$$f^*(x^*) = \begin{cases} \displaystyle\sup_{y \in \mathbb{R}^n,\, Qy + q = x^*} \frac{1}{2} Qy \cdot y, & \text{if } x^* \in q + Q\,\mathbb{R}^n \\ +\infty, & \text{if } x^* \notin q + Q\,\mathbb{R}^n \end{cases}$$
$$\operatorname{dom} f^* = q + Q\,\mathbb{R}^n.$$

By Theorem 7.9(ii), the Fenchel–Legendre transform of I_U is

$$I_U^*(x^*) = \begin{cases} \displaystyle\inf_{\substack{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k \\ x^* = A^\top \lambda + B^\top \mu}} (\alpha \cdot \lambda + \beta \cdot \mu), & \text{if } x^* \in A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k \\ +\infty, & \text{if } x^* \notin A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k \end{cases} \tag{7.92}$$
$$\operatorname{dom} I_U^* = A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k.$$

The *Fenchel dual problem*, which coincides with the *dual Lagrange problem*, is given by

$$\begin{aligned} &\text{maximize } -\alpha \cdot \lambda - \beta \cdot \mu - \frac{1}{2} Qx \cdot x \\ &\text{subject to } A^\top \lambda + B^\top \mu + Qx = -q \quad \text{and} \quad \lambda \geq 0, \end{aligned} \tag{7.93}$$

$$V \stackrel{\text{def}}{=} \left\{ (x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k : A^\top \lambda + B^\top \mu + Qx = -q \quad \text{and} \quad \lambda \geq 0 \right\}. \tag{7.94}$$

The dual problem is *feasible* if there exists $(x,\lambda,\mu) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$ such that $A^\top\lambda + B^\top\mu + Qx = -q$; otherwise, it is *infeasible*. Finally, as U is closed and convex and f is continuous of domain $\mathrm{dom}\, f = \mathbb{R}^n$, I_U is convex lsc and f and I_U are convex lsc. From the Fenchel duality theorem (Theorem 9.8 of Chapter 2) we have strong duality if either $U \neq \varnothing$ or $V \neq \varnothing$.

As in LP, the QP problem can be put in several equivalent forms.

Definition 7.3.
Let $x, b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$, $Q \geq 0$ be an $n \times n$ symmetrical matrix, and A be an $m \times n$ matrix.

(i) The QP problem in *standard form* (SQP) is given as

$$\begin{aligned} &\text{minimize } -b\cdot x + \frac{1}{2} Qx\cdot x \\ &\text{subject to } Ax \leq c \quad \text{and} \quad x \geq 0. \end{aligned} \tag{7.95}$$

(ii) The QP problem in *canonical form* (CQP) is given as

$$\begin{aligned} &\text{minimize } -b\cdot x + \frac{1}{2} Qx\cdot x \\ &\text{subject to } Ax = c \quad \text{and} \quad x \geq 0. \end{aligned} \tag{7.96}$$

(iii) The QP problem in *inequality form* (IQP) is given as

$$\begin{aligned} &\text{minimize } -b\cdot x + \frac{1}{2} Qx\cdot x \\ &\text{subject to } Ax \leq c. \end{aligned} \tag{7.97}$$

□

Theorem 7.12. *Let U and f be as specified by* (7.85) *with $q \in \mathbb{R}^n$ and $Q \geq 0$.*

(i) *The necessary and sufficient condition for the existence of a minimizer on U to problem* (7.85) *is*

$$\exists x \in U \text{ such that } Qx + q \in T_U(x)^*, \tag{7.98}$$

$$T_U(x)^* = \left\{ -\sum_{i\in I(x)} \lambda_i a_i - \sum_{j=1}^{k} \mu_j b_j : \begin{array}{l} \forall \lambda_i \geq 0,\ i \in I(x) \\ \forall \mu_j \in \mathbb{R},\ 1 \leq j \leq k \end{array} \right\}. \tag{7.99}$$

(ii) *The necessary and sufficient condition* (7.98) *is equivalent to the existence of $\hat{x} \in \mathbb{R}^n$, $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_m) \in \mathbb{R}^m$, and $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_k) \in \mathbb{R}^k$ such that*

$$\begin{aligned} &\hat{\lambda}_i \geq 0, \quad a_i\cdot\hat{x} - \alpha_i \leq 0, \quad \hat{\lambda}_i\,(a_i\cdot\hat{x} - \alpha_i) = 0, \quad 1 \leq i \leq m, \\ &b_j\cdot\hat{x} = \beta_j, \quad 1 \leq j \leq k, \\ &Q\hat{x} + q + \sum_{i=1}^{m} \hat{\lambda}_i a_i + \sum_{j=1}^{k} \hat{\mu}_j b_j = 0. \end{aligned} \tag{7.100}$$

Moreover, $(\hat{x},\hat{\lambda},\hat{\mu}) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$ *is a saddle point of the Lagrangian L:*

$$\forall x \in \mathbb{R}^n, \forall(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k, \quad L(\hat{x},\lambda,\mu) \leq L(\hat{x},\hat{\lambda},\hat{\mu}) \leq L(x,\hat{\lambda},\hat{\mu}). \tag{7.101}$$

(iii) *If* $U \neq \varnothing$ *(primal feasible),* $\inf_{x\in U} f(x) < +\infty$*, and we have strong duality*

$$\inf_{x\in U} f(x) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu). \tag{7.102}$$

There are two cases:

(a) *if* $-q \notin A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k + Q\,\mathbb{R}^n$ *(dual infeasible,* $V = \varnothing$*),*

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = -\infty$$

and U is unbounded;

(b) *if* $-q \in A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k + Q\,\mathbb{R}^n$ *(dual feasible,* $V = \varnothing$*), then*

$$-\infty < \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) < +\infty$$

and the Lagrangian has a saddle point $(\hat{x},\hat{\lambda},\hat{\mu}) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$ *solution of system* (7.100) *of part* (ii).

If $U = \varnothing$ *(primal infeasible),* $\inf_{x\in U} f(x) = +\infty$ *and we have two cases:*

(c) *for* $-q \notin A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k + Q\,\mathbb{R}^n$ *(dual infeasible,* $V = \varnothing$*),*

$$-\infty = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) < \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = +\infty;$$

(d) *for* $-q \in A^\top\mathbb{R}^m_+ + B^\top\mathbb{R}^k + Q\,\mathbb{R}^n$ *(dual feasible,* $V \neq \varnothing$*), we have strong duality*

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = +\infty$$

and V is unbounded.

Proof. (i) Since f and U are convex, the dual necessary optimality condition $\nabla f(\hat{x}) \in T_U(\hat{x})^*$ of Theorem 6.3(iii) becomes necessary and sufficient. Moreover, for affine constraints, the dual cone $T_U(\hat{x})^*$ is given by (7.26) Theorem 7.4.

(ii) From part (i), by using the second characterization (7.26) in Theorem 7.4. Since the Lagrangian is quadratic, for all $(x,\lambda,\mu) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$

$$\begin{aligned}
L(x,\lambda,\mu) &= L(\hat{x},\hat{\lambda},\hat{\mu}) + (\underbrace{Q\hat{x} + q + A^\top\hat{\lambda} + B^\top\hat{\mu}}_{=0}) \cdot (x - \hat{x}) \\
&\quad + (\underbrace{A\hat{x} - \alpha}_{\leq 0}) \cdot (\lambda - \hat{\lambda}) + (\underbrace{B\hat{x} - \beta}_{=0}) \cdot (\mu - \hat{\mu}) + \frac{1}{2}\underbrace{Q(x - \hat{x}) \cdot (x - \hat{x})}_{\geq 0} \\
&= L(\hat{x},\hat{\lambda},\hat{\mu}) + (\underbrace{A\hat{x} - \alpha}_{\leq 0}) \cdot \underbrace{\lambda}_{\geq 0} + \frac{1}{2}\underbrace{Q(x - \hat{x}) \cdot (x - \hat{x})}_{\geq 0}.
\end{aligned}$$

Therefore

$$\forall x \in \mathbb{R}^n, \quad L(x,\hat\lambda,\hat\mu) = L(\hat x,\hat\lambda,\hat\mu) + \frac{1}{2}\underbrace{Q(x-\hat x)\cdot(x-\hat x)}_{\geq 0} \geq L(\hat x,\hat\lambda,\hat\mu)$$
$$\forall(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k, \quad L(\hat x,\lambda,\mu) = L(\hat x,\hat\lambda,\hat\mu) + \underbrace{(A\hat x - \alpha)}_{\leq 0}\cdot \underbrace{\lambda}_{\geq 0} \leq L(\hat x,\hat\lambda,\hat\mu)$$
$$\Rightarrow \forall (x,\lambda,\mu) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k, \quad L(\hat x,\lambda,\mu) \leq L(\hat x,\hat\lambda,\hat\mu) \leq L(x,\hat\lambda,\hat\mu)$$

and $(\hat x,\hat\lambda,\hat\mu)$ is a saddle point of L. Conversely, if $U \neq \varnothing$ and $(\hat x,\hat\lambda,\hat\mu)$ is a saddle point of L, the saddle point equations yield $\hat x \in U$ and

$$f(\hat x) = L(\hat x,\hat\lambda,\hat\mu) = \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \inf_{x\in U} f(x).$$

(iii) In general, for $x \in \mathbb{R}^n$,

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \begin{cases} f(x), & \text{if } x \in U \\ +\infty, & \text{if } x \notin U \end{cases}$$
$$\Rightarrow \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \inf_{x\in U} f(x)$$

which is finite for $U \neq \varnothing$ and $+\infty$ for $U = \varnothing$.

If $U \neq \varnothing$, then for all $(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k$,

$$\inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \leq \inf_{x\in U} L(x,\lambda,\mu) \leq \inf_{x\in U} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \inf_{x\in U} f(x) < +\infty$$
$$\Rightarrow \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \leq \inf_{x\in\mathbb{R}^n} \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) < +\infty. \tag{7.103}$$

Consider the closed convex cone in 0: $C \overset{\text{def}}{=} A^\top \mathbb{R}^m_+ + B^\top \mathbb{R}^k + Q\,\mathbb{R}^n$. If $-q \notin C$, there exists a minimizer $(\hat\lambda,\hat\mu,\hat x) \in \mathbb{R}^m_+ \times \mathbb{R}^k \times \mathbb{R}^n$ of $\|q + A^\top\lambda + B^\top\mu + Qx\|^2$ with respect to $(\lambda,\mu,x) \in \mathbb{R}^m_+ \times \mathbb{R}^k \times \mathbb{R}^n$. From Corollary 1(iii) to Theorem 3.2

$$\forall(\lambda,\mu,x) \in \mathbb{R}^m_+ \times \mathbb{R}^k \times \mathbb{R}^n, \quad (q + A^\top\hat\lambda + B^\top\hat\mu + Q\hat x)\cdot(A^\top\lambda + B^\top\mu + Qx) \geq 0$$
$$(q + A^\top\hat\lambda + B^\top\hat\mu + Q\hat x)\cdot(A^\top\hat\lambda + B^\top\hat\mu + Qx) = 0 \tag{7.104}$$
$$\Rightarrow \begin{cases} B\,[q + A^\top\hat\lambda + B^\top\hat\mu + Q\hat x] = 0, \quad Q\,[q + A^\top\hat\lambda + B^\top\hat\mu + Q\hat x] = 0, \\ \text{and } A\,[q + A^\top\hat\lambda + B^\top\hat\mu + Q\hat x] \in \mathbb{R}^m_+ \end{cases} \tag{7.105}$$
$$\Rightarrow \forall n \in \mathbb{N}, \quad z_n \overset{\text{def}}{=} n\,(q + A^\top\hat\lambda + B^\top\hat\mu + Q\hat x) \in \mathbb{R}^m_+ \cap \operatorname{Ker} B \cap \operatorname{Ker} Q.$$

Since $U \neq \varnothing$, choose an $\bar x \in U$ and consider the points $x_n = \bar x - z_n$. Then

$$a_i \cdot (\bar x - z_n) \leq \alpha_i + 0, \quad b_j \cdot (\bar x - z_n) = \beta_j - 0, \text{ and } Qx_n = Q\bar x \quad \Rightarrow \forall n, \quad x_n \in U.$$

By using identity (7.104), we get a sequence $\{x_n\} \subset U$ such that

$$\begin{aligned}
f(x_n) - f(\hat{x}) &= (Q\hat{x}+q)\cdot(x_n-\hat{x}) + \frac{1}{2}Q(x_n-\bar{x})\cdot(x_n-\bar{x}) \\
&= -n\,(Q\hat{x}+q)\cdot(q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x}) \\
&\quad + n^2\,\frac{1}{2}\underbrace{Q(q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x})}_{=0}\cdot(q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x}) \\
&= -n\,q\cdot(q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x}) + n\,\hat{x}\cdot\underbrace{Q(q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x})}_{=0} \\
&= -n\,\|q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x}\|^2 \\
&\quad + n\underbrace{(A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x})\cdot(q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x})}_{=0}.
\end{aligned}$$

As $-q \notin C$, $q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x} \neq 0$ and

$$f(x_n) = f(\hat{x}) = -n\,\|q+A^\top\hat{\lambda}+B^\top\hat{\mu}+Q\hat{x}\|^2 \to -\infty$$

as n goes to infinity. Therefore, $q \notin C$ implies that

$$\sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k}\ \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n}\ \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) = \inf_{x\in U} q\cdot x = -\infty;$$

each member is equal to $-\infty$.

If $-q \in C$, consider the function $x \mapsto L(x,\lambda,\mu) = f(x) + (Ax-\alpha)\cdot\lambda + (Bx-\beta)\cdot\mu$. Since L is quadratic in x and its Hessian matrix with respect to x is $Q \geq 0$, by Theorem 2.2(i), there exists a minimizer $x_{\lambda,\mu} \in \mathbb{R}^n$ if and only if the gradient of L with respect to x is zero, that is, $q + Qx_{\lambda,\mu} + A^\top\lambda + B^\top\mu = 0$. By assumption, there exists $(\bar{\lambda},\bar{\mu},\bar{x}) \in \mathbb{R}^m_+\times\mathbb{R}^k\times\mathbb{R}^n$ such that $-q = Q\bar{x} + A^\top\bar{\lambda} + B^\top\bar{\mu}$ and

$$-\infty < -\frac{1}{2}Q\bar{x}\cdot\bar{x} - \bar{\lambda}\cdot\alpha - \bar{\mu}\cdot\beta = \inf_{x\in\mathbb{R}^n} L(x,\bar{\lambda},\bar{\mu}) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k}\ \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu).$$

For $(\bar{\lambda},\bar{\mu})$ the minimizer $\bar{x}$ is unique up to an element of $\operatorname{Ker} Q$. Indeed, $Q\bar{x}_1 + A^\top\bar{\lambda} + B^\top\bar{\mu} = -q = Q\bar{x}_2 + A^\top\bar{\lambda} + B^\top\bar{\mu}$ implies that $\bar{x}_2 - \bar{x}_1 \in \operatorname{Ker} Q$ and, by symmetry of Q, $Q\bar{x}_2\cdot\bar{x}_2 = Q\bar{x}_2\cdot\bar{x}_1 = Q\bar{x}_1\cdot\bar{x}_1$. For the (λ,μ) for which there is no $x \in \mathbb{R}^n$ such that $A^\top\lambda + B^\top\mu + Qx = -q$,

$$-\infty = \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu).$$

Therefore, we get

$$\begin{aligned}
&-\infty < \sup_{\substack{(x,\lambda,\mu)\in\mathbb{R}^n\times\mathbb{R}^m_+\times\mathbb{R}^k\\ Qx+A^\top\lambda+B^\top\mu=-q}} L(x,\lambda,\mu) = \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k}\ \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \leq \inf_{x\in U} f(x) < +\infty \\
&\Rightarrow\quad -\infty < \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k}\ \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \inf_{x\in\mathbb{R}^n}\ \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} L(x,\lambda,\mu) < +\infty.
\end{aligned}$$

By Theorem 7.10, as $U \neq \varnothing$ and $\inf f(U) > -\infty$, there exist a minimizer $\hat{x} \in U$ and, from part (ii), a saddle point. In both cases, strong duality is obtained.

Finally we consider the case $U = \varnothing$ for which $\inf_{x\in U} f(x) = +\infty$. If $-q \notin C$, then, for all $(x,\lambda,\mu) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$, $\nabla_x L(x,\lambda,\mu) = q + Qx + A^\top\lambda + B^\top\mu \neq 0$ and, by Theorem 3.3,

$$\forall(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k,\ \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = -\infty \quad \Rightarrow \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = -\infty.$$

If $-q \in C$, there exists $(\hat{x},\hat{\lambda},\hat{\mu}) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^k$ such that $\nabla_x L(\hat{x},\hat{\lambda},\hat{\mu}) = q + Q\hat{x} + A^\top\hat{\lambda} + B^\top\hat{\mu} = 0$ and $H_x L(\hat{x},\hat{\lambda},\hat{\mu}) = Q \geq 0$ since L is quadratic in x. By Theorem 3.3, $\hat{x} \in \mathbb{R}^n$ is a minimizer of $L(x,\hat{\lambda},\hat{\mu})$ and

$$-\infty < L(\hat{x},\hat{\lambda},\hat{\mu}) = \inf_{x\in\mathbb{R}^n} L(x,\hat{\lambda},\hat{\mu}) = -\alpha\cdot\hat{\lambda} - \beta\cdot\hat{\mu} - \frac{1}{2}Q\hat{x}\cdot\hat{x},$$

since, as was already noted, $\hat{x}$ is unique up to an element of $\operatorname{Ker} Q$. Since L is quadratic, its Taylor expansion yields

$$\begin{aligned}
L(x,\lambda,\mu) &= L(\hat{x},\hat{\lambda},\hat{\mu}) + (\underbrace{Q\hat{x} + q + A^\top\hat{\lambda} + B^\top\hat{\mu}}_{=0})\cdot(x-\hat{x})\\
&\qquad + (A\hat{x}-\alpha)\cdot(\lambda-\hat{\lambda}) + (B\hat{x}-\beta)\cdot(\mu-\hat{\mu}) + \frac{1}{2}\underbrace{Q(x-\hat{x})\cdot(x-\hat{x})}_{\geq 0}\\
&\geq L(\hat{x},\hat{\lambda},\hat{\mu}) + (A\hat{x}-\alpha)\cdot(\lambda-\hat{\lambda}) + (B\hat{x}-\beta)\cdot(\mu-\hat{\mu})\\
\Rightarrow \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) &\geq L(\hat{x},\hat{\lambda},\hat{\mu}) + (A\hat{x}-\alpha)\cdot(\lambda-\hat{\lambda}) + (B\hat{x}-\beta)\cdot(\mu-\hat{\mu}).
\end{aligned}$$

Since $U = \varnothing$, either $b_j\cdot\hat{x} - \beta_j \neq 0$ for some j or $a_i\cdot\hat{x} - \alpha_i > 0$ for some i. In the first case, choose the following sequences: for $n \geq 1$

$$\begin{gathered}
\lambda^n_i = 0,\ 1 \leq i \leq m, \quad \mu^n_j = \hat{\mu}_j + n\,(b_j\cdot\hat{x} - \beta_j) \quad \text{et} \quad \mu^n_{j'} = 0,\ j' \neq j\\
\Rightarrow \inf_{x\in\mathbb{R}^n} L(x,\lambda^n,\mu^n) \geq L(\hat{x},\hat{\lambda},\hat{\mu}) + n\,(b_j\cdot\hat{x} - \beta_j)^2 \to +\infty.
\end{gathered}$$

In the second case, choose the following sequences: for $n \geq 1$

$$\begin{gathered}
\mu^n_j = 0,\ 1 \leq j \leq k, \quad \lambda^n_i = \hat{\lambda}_i + n\,(a_i\cdot\hat{x} - \alpha_i) \quad \text{et} \quad \lambda^n_{i'} = 0,\ i' \neq i\\
\Rightarrow \inf_{x\in\mathbb{R}^n} L(x,\lambda^n,\mu^n) \geq L(\hat{x},\hat{\lambda},\hat{\mu}) + n\,(a_i\cdot\hat{x} - \alpha_i)^2 \to +\infty.
\end{gathered}$$

Therefore,

$$+\infty = \lim_{n\to\infty} \inf_{x\in\mathbb{R}^n} L(x,\lambda^n,\mu^n) \leq \sup_{(\lambda,\mu)\in\mathbb{R}^m_+\times\mathbb{R}^k} \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu). \qquad \square$$

7.7 Fréchet Differentiable Objective Function

Condition (7.27) is the prototype of the necessary condition of the Lagrange multipliers theorem and the Karush–Khun–Tucker theorem that will be given in Chapter 5 for nonlinear constraints. In fact when the objective function f is Fréchet differentiable and the constraint functions are affine, we get the following necessary condition.

Theorem 7.13. *Let U be as specified by* (7.21) *and $f : \mathbb{R}^n \to \mathbb{R}$. If x is a local minimizer of f in U and f is Fréchet differentiable at x, then there exist $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ and $\mu = (\mu_1, \dots, \mu_k) \in \mathbb{R}^k$ such that*

$$\begin{aligned} &\lambda_i \geq 0, \quad a_i \cdot x - \alpha_i \leq 0, \quad \lambda_i (a_i \cdot x - \alpha_i) = 0, \quad 1 \leq i \leq m, \\ &b_j \cdot h = 0, \quad 1 \leq j \leq k, \\ &\nabla f(x) + \sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j b_j = 0. \end{aligned} \tag{7.106}$$

We shall see in Chapter 5 that for nonaffine constraints, the a_i's and b_j's will be replaced by the gradient of the corresponding constraint function.

7.8 Farkas' Lemma and Its Extension

For U specified by (7.21) and $q \in \mathbb{R}^n$, we have seen the equivalence of the following three conditions:

$$T_U(x) \subset \{x \in \mathbb{R}^n : q \cdot x \geq 0\} \iff \{\lambda_0 q : \lambda_0 \geq 0\} \subset T_U(x)^* \iff q \in T_U(x)^*.$$

This equivalence is an extension of the lemma of J. FARKAS [**1**][19] that can be obtained by replacing the tangent cone $T_U(x)$ by the following closed convex cone in 0:

$$V \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \begin{array}{c} a_i \cdot x \leq 0, \ 1 \leq i \leq m \\ b_j \cdot x = 0, \ 1 \leq j \leq k \end{array} \right\}. \tag{7.107}$$

Lemma 7.1. *Given $q \in \mathbb{R}^n$ and the set V defined in* (7.107)*, the following three conditions are equivalent:*

$$V \subset \{x \in \mathbb{R}^n : q \cdot x \geq 0\} \iff \{\lambda_0 q : \lambda_0 \geq 0\} \subset V^* \iff q \in V^*.$$

Proof. From Theorem 5.6(ii), $C^* = \{\lambda_0 q : \lambda_0 \geq 0\}$. So, if $C^* \subset V^*$, then, for $\lambda_0 = 1$, $q \in V^*$. Conversely, if $q \in V^*$, then for all $\lambda_0 \geq 0$, $\lambda_0 q \in V^*$ since V^* is a cone in 0. So, we have the last two equivalences. As for the first two equivalences, since V and $C = \{x \in \mathbb{R}^n : q \cdot x \geq 0\}$ are both closed convex cones in 0, we have $V \subset C$ if and only if $C^* \subset V^*$ by Theorem 5.5(ii). □

This gives the extension of Farkas's lemma to the case of a finite number of linear equalities and inequalities.

[19] Julius (Gyula) Farkas (1847–1930).

Lemma 7.2. *Let $n \geq 1$, $m \geq 1$, and $k \geq 1$ be three integers and q, $a_1, a_2, \ldots, a_m$, $b_1, b_2, \ldots, b_k$ be vectors in $\mathbb{R}^n$. The following conditions are equivalent:*

(i) $$\left\{x \in \mathbb{R}^n : \begin{matrix} a_i \cdot x \leq 0, \ 1 \leq i \leq m \\ b_j \cdot x = 0, \ 1 \leq j \leq k \end{matrix}\right\} \subset \left\{x \in \mathbb{R}^n : q \cdot x \geq 0\right\}$$

(ii) $$\{\lambda_0 q : \forall \lambda_0 \geq 0\} \subset \left\{-\sum_{i=1}^{m} \lambda_i a_i - \sum_{j=1}^{k} \mu_j b_j : \begin{matrix} \forall \lambda_i \geq 0, \ 1 \leq i \leq m, \\ \forall \mu_j \in \mathbb{R}, \ 1 \leq j \leq k \end{matrix}\right\}$$

(iii) $$\exists \lambda_i \geq 0, \ 1 \leq i \leq m, \ \exists \mu_j \in \mathbb{R}, \ 1 \leq j \leq k, \quad q + \sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j b_j = 0.$$

8 ► Glimpse at Optimality via Subdifferentials

Prior to closing this chapter, we very briefly describe the *subdifferential* approach to characterize a minimizer.

As we have previously seen, the local infimum of a function over a subset of $\mathbb{R}^n$ can always be reformulated as the local infimum of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ over $\mathbb{R}^n$: there exist $\hat{x}$ and a neighborhood $V(\hat{x})$ of $\hat{x}$ such that

$$\forall x \in V(\hat{x}), \quad f(x) \geq f(\hat{x}).$$

Definition 8.1.
Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $x \in \text{dom } f$.

(i) $h \in \mathbb{R}^n$ is a *subgradient* of f at x if there exists a neighborhood $V(x)$ of x such that

$$\forall y \in V(x), \quad f(y) - f(x) \geq h \cdot (y - x).$$

The set of all subgradients of f at x is denoted $\partial f(x)$ and will be referred to as the *subdifferential* of f at x.[20]

(ii) $h \in \mathbb{R}^n$ is a *global subgradient* of f at x if

$$\forall y \in \mathbb{R}^n, \quad f(y) - f(x) \geq h \cdot (y - x).$$

The set of all global subgradients of f at x is denoted $\partial_{global} f(x)$ and will be referred to as the *global subdifferential* of f at x.

By convention, $\partial f(x) = \varnothing$ and $\partial_{global} f(x) = \varnothing$ for all $x \notin \text{dom } f$. □

It is easy to check that both subdifferentials are convex sets and that $\partial_{global} f(x)$ is closed.

We readily get the following trivial characterizations.

Theorem 8.1. *Let function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\text{dom } f \neq \varnothing$.*

(i) *f has a local minimizer if and only if there exists $\hat{x}$ such that* $0 \in \partial f(\hat{x})$.

(ii) *f has a global minimizer if and only if there exists $\hat{x}$ such that* $0 \in \partial_{global} f(\hat{x})$.

[20] We adopt the terminology of J. M. BORWEIN and A. S. LEWIS [**1**, p. 35].

When f is convex, the local infima coincide with the global infimum and $\partial_{global} f(x) = \partial f(x)$ and the subdifferential can be easily characterized.

Theorem 8.2. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\mathrm{dom}\, f \neq \varnothing$, be convex.*

(i) *If $x \in \mathrm{dom}\, f$,*

$$\partial f(x) = \left\{ h \in \mathbb{R}^n : \forall y \in \mathbb{R}^n,\ \underline{d} f(x, y-x) \geq h \cdot (y-x) \right\}. \tag{8.1}$$

(ii) *If $x \in \mathrm{int}\,(\mathrm{dom}\, f)$, for all $v \in \mathbb{R}^n$, $d_H(x;v) = \overline{d}_C f(x;v)$ and*

$$\partial f(x) = \left\{ h \in \mathbb{R}^n : \forall v \in \mathbb{R}^n,\ d_H f(x,v) \geq h \cdot v \right\}. \tag{8.2}$$

This gives the general idea behind subdifferentials. We now specialize to Lipschitzian functions and quote a few basic properties from F. H. CLARKE [**2**, Props. 2.1.1 and 2.1.2, pp. 25–27] for subdifferentials. This completes Theorem 3.5(ii) and section 3.5.3 of Chapter 3 for the Hadamard semidifferential of Lipschitz functions.

Theorem 8.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian at x with Lipschitz constant $c(x)$.*[21]

(i) *The function $v \mapsto \overline{d}_C f(x;v) : \mathbb{R}^n \to \mathbb{R}$ is well defined, positively homogeneous, and subadditive, and it satisfies*

$$\forall v \in \mathbb{R}^n, \quad \left| \overline{d}_C f(x;v) \right| \leq c(x) \|v\|. \tag{8.3}$$

(ii) *The function $(x,v) \mapsto \overline{d}_C f(x;v)$ is usc.*

(iii) $\overline{d}_C f(x;-v) = \overline{d}_C(-f)(x;v)$.

In view of the characterization of $\partial f(x)$ for convex functions, we introduce a weaker subdifferential which is better adapted for Lipschitzian functions.

Definition 8.2.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian at x with Lipschitz constant $c(x)$. The *subdifferential* of f at x is the set

$$\partial_C f(x) \overset{\text{def}}{=} \left\{ h \in \mathbb{R}^n : \overline{d}_C f(x;v) \geq h \cdot v \text{ for all } v \in \mathbb{R}^n \right\}. \tag{8.4}$$

□

Recall from Theorem 4.11 in section 4.4 of Chapter 3 that for semiconvex functions, $\overline{d}_C f(x;v) = d_H f(x;v)$. So, for that class of functions, the advantages of the semidifferential and the subdifferential can be combined.

Theorem 8.4. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian at x with Lipschitz constant $c(x)$.*

(i) *$\partial_C f(x)$ is not empty, convex, compact, and*

$$\forall h \in \partial_C f(x), \quad \|h\| \leq c(x). \tag{8.5}$$

[21] see Definition 3.8(i) of Chapter 3.

(ii) *For every $v \in \mathbb{R}^n$,*

$$\overline{d}_C f(x;v) = \max\{h \cdot v : h \in \partial_C f(x)\}. \tag{8.6}$$

As in the general case, a local minimizer $\hat{x}$ of the function f will satisfy the condition $0 \in \partial_C f(\hat{x})$. So, instead of dealing with semidifferentials $d_H f(x;v)$ that are functions, we deal with subdifferentials $\partial_C f(x)$ that are convex compact sets. The *subdifferential calculus* will involve set-valued functions, lim, liminf, and limsup of families of sets (see, for instance, J.-P. AUBIN and H. FRANKOWSKA [**1**]), but a systematic exposition is beyond the objectives of this book. In the subdifferential (set-valued) calculus, the basic functional operations take the following forms:

$$\begin{gathered}
\partial_C(f+g)(x) \subset \partial_C f(x) + \partial_C g(x) \\
\partial_C(f \vee g)(x) \subset \overline{\text{co}}\,(\partial_C f(x) \cup \partial_C g(x)) \\
\partial_C(fg)(x) \subset g(x)\partial_C f(x) + f(x)\partial_C g(x).
\end{gathered}$$

9 Exercises

Exercise 9.1.
Show that the matrix (2.9) is positive definite. □

Exercise 9.2.
Consider the function

$$f(x_1,x_2) = (x_1+1)^2 + (x_2-2)^2.$$

Find the minimizers of f with respect to the set

$$U = \{(x_1,x_2) : x_1 \geq 0 \text{ and } x_2 \geq 0\}.$$

Justify. □

Exercise 9.3.
Consider the function

$$f(x_1,x_2) \stackrel{\text{def}}{=} \frac{1}{2}\left[(x_1-1)^2 + (x_2+2)^2\right] + \frac{1}{2}x_1x_2 \tag{9.1}$$

to minimize with respect to the set

$$U \stackrel{\text{def}}{=} \{(x_1,x_2) : x_1 \geq 0 \text{ and } x_2 \geq 0\}. \tag{9.2}$$

Find the minimizers. Justify. □

Exercise 9.4.
Let $P^1[0,1]$ be the set of polynomials from $[0,1]$ to $\mathbb{R}$ of degree less than or equal to 1; that is,

$$p(x) = ax + b, \quad \forall a \in \mathbb{R}, \quad \forall b \in \mathbb{R}.$$

Consider the function

$$f(p) \stackrel{\text{def}}{=} \int_0^1 \left(p(x) - (2x-1)^3\right)^2 dx.$$

Characterize and find the polynomials that minimize f with respect to the set

$$U \stackrel{\text{def}}{=} \{p \in P^1[0,1] : \forall x \in [0,1],\ p(x) \geq 0\}.$$

Hint: prove that $b_0(x) = 1 - x$ and $b_1(x) = x$ are basis functions for $P^1[0,1]$ and that

$$U = \{p_0 b_0 + p_1 b_1 : \forall p_0 \geq 0 \text{ and } \forall p_1 \geq 0\}.$$

Then prove that U is a convex cone in 0. □

Exercise 9.5.
Compute the dual cones for the following subsets of $\mathbb{R}^2$.

(i) $U_1 = \{(x,y) \in \mathbb{R}^2 : x = 0 \text{ and } y \geq 0\}$

(ii) $U_2 = \{(x,y) \in \mathbb{R}^2 : x + 5y \geq 0 \text{ and } 2x + 3y \leq 0\}$

(iii) $U_3 = \{(x,y) \in \mathbb{R}^2 : 2x + y \geq 0 \text{ and } 10x + 5y \leq 0\}$

(iv) $U_4 = \{(x,y) \in \mathbb{R}^2 : x + y = 0 \text{ and } 2x + 3y \geq 0\}$

(v) $U_5 = \{(1,2), (3,1)\}$. □

Exercise 9.6.
Prove the equivalence of the standard, canonical, and inequality forms of a linear program in Definition 7.1 by using slack and/or extra variables and by transforming inequalities into equalities and vice versa. □

Exercise 9.7.
Let Q be an $n \times n$ symmetric matrix and B an $m \times n$ $(m \leq n)$ matrix. Let $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ be two vectors. Consider the function

$$f(x) = \frac{1}{2} Qx \cdot x - a \cdot x, \tag{9.3}$$

the set of constraints

$$U = \{x \in \mathbb{R}^n : Bx = b\}, \tag{9.4}$$

and the minimization problem

$$\inf\{f(x) : x \in U\}. \tag{9.5}$$

(i) Prove that the vector $\hat{x}$ in $\mathbb{R}^n$ is the unique solution of problem (9.5) if and only if there exists a vector $\hat{\lambda}$ in $\mathbb{R}^m$ such that

$$\begin{cases} Q\hat{x} + B^\top \hat{\lambda} = a \\ B\hat{x} = b \end{cases} \tag{9.6}$$

$$\forall x \in \operatorname{Ker} B,\ h \neq 0, \quad Qh \cdot h \geq 0. \tag{9.7}$$

(ii) Prove that if the rank of the matrix B is m and the matrix Q is positive definite, then there always exists a unique solution to problem (9.5) for all a and b.

(iii) Assume that the rank of the matrix B is m and that Q is positive definite. Give the expression of the solution $\hat{x}$ as a function of the data A, a, B, and b [show that $BQ^{-1}B^\top$ is invertible and eliminate $\hat{\lambda}$].

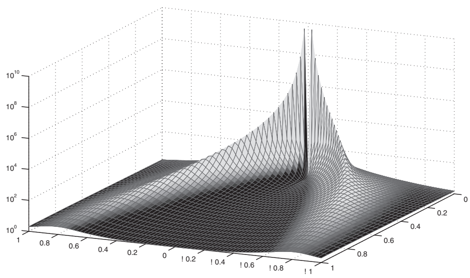
(iv) (Cf. The marginal function of Example 9.7 in Chapter 2.) Assume that the rank of the matrix B is m and that Q is positive definite. Denote by $\hat{x}(b)$ the solution of (9.5) that corresponds to b and define the new function

$$g(b) = f(\hat{x}(b)).$$

Prove that

$$\nabla g(b) = -\hat{\lambda}(b),$$

where $\hat{\lambda}(b)$ is the vector $\hat{\lambda}$ of part (i) [compute $(\hat{x}(b), \hat{\lambda}(b))$ as a function of B, Q, b, and a and use the chain rule to get the semidifferential of the composition $f(\hat{x}(b))$]. □



Chapter 5
Constrained Differentiable Optimization

1 Constrained Problems

A local minimizer $x \in U$ of an objective function $f : \mathbb{R}^n \to \mathbb{R}$ with respect to U, such that f is Fréchet differentiable at x, satisfies the *necessary optimality condition* (6.4) of Theorem 6.1(ii) in Chapter 4:

$$T_U(x) \subset C_f(x) \stackrel{\text{def}}{=} \{h \in \mathbb{R}^n : \nabla f(x) \cdot h \geq 0\}. \tag{1.1}$$

It is equivalent to the *dual optimality condition* (6.14) of Theorem 6.3(iii) in Chapter 4:

$$\nabla f(x) \in T_U(x)^*. \tag{1.2}$$

The cone $T_U(x)$ and its dual $T_U(x)^*$ will have to be characterized.

In this chapter, we consider sets U that are specified by a finite number of equality or/and inequality constraints via functions (called *constraint functions*) that are Fréchet differentiable at $x \in U$ or of class $C^{(1)}$ in a neighborhood of x. This will require the characterization of the cone $T_U(x)$ and its dual $T_U(x)^*$. Section 2 studies the case of U specified by equality constraints

$$U \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : g_j(x) = 0, \quad 1 \leq j \leq m\}, \tag{1.3}$$

for m functions

$$g_j : \mathbb{R}^n \to \mathbb{R}, \quad 1 \leq j \leq m. \tag{1.4}$$

This leads to the *Lagrange multipliers theorem.* Section 3 will consider U given by inequality constraints:

$$U \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : g_j(x) \leq 0, \quad 1 \leq j \leq m\right\}, \tag{1.5}$$

where the g_j are the functions of (1.4). This leads to *Karush–Kuhn–Tucker theorem.* Finally, section 4 will consider sets U specified by both equality and inequality constraints.

2 Equality Contraints: Lagrange Multipliers Theorem

2.1 Tangent Cone of Admissible Directions

Consider the set

$$U \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : g_j(x) = 0, \quad 1 \le j \le m\} \tag{2.1}$$

specified by m functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$. The objective is to prove that for g_j of class $C^{(1)}$ in a neighborhood of x and an assumption on the Jacobian matrix of the vector function $g = (g_1, \dots, g_m) : \mathbb{R}^n \to \mathbb{R}^m$ at a point $x_0 \in U$,

$$T_U(x_0) = \{h \in \mathbb{R}^n : \nabla g_j(x_0) \cdot h = 0, \quad 1 \le j \le m\}. \tag{2.2}$$

Thence, by Theorem 7.2(ii) of Chapter 4, the dual cone is

$$T_U(x_0)^* = \left\{ \sum_{j=1}^m \lambda_j \nabla g_j(x_0) \; : \forall \lambda_j \in \mathbb{R}, \;\; \forall j = 1, \dots, m \right\}. \tag{2.3}$$

By using the dual necessary optimality condition (1.2),

$$\exists (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m \text{ such that } \quad \nabla f(x_0) + \sum_{j=1}^m \lambda_j \nabla g_j(x_0) = 0. \tag{2.4}$$

This is the main conclusion of the Lagrange multipliers theorem.

The central element of the proof is (2.2) that will be established under reasonable assumptions. Identity (2.2) is easy to prove in one direction.

Lemma 2.1. *Let $x_0 \in \mathbb{R}^n$, $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, and*

$$U_0 \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : g_j(x) = g_j(x_0), 1 \le j \le m \right\}.$$

(i) *If the g_j's are Hadamard semidifferentiable at x_0, then*

$$T_{U_0}(x_0) \subset \{h \in \mathbb{R}^n \, : \, d_H g_j(x_0; h) = 0, \, 1 \le j \le m\}. \tag{2.5}$$

(ii) *If the g_j's are Fréchet differentiable at x_0, then*

$$T_{U_0}(x_0) \subset \{h \in \mathbb{R}^n \, : \, \nabla g_j(x_0) \cdot h = 0, \, 1 \le j \le m\}. \tag{2.6}$$

Proof. By definition of an admissible direction h of U_0 in x_0, there exists a sequence $\{t_n > 0\}$, $t_n \searrow 0$ as $n \to \infty$, for which

$$\forall n, \;\; \exists x_n \in U_0, \text{ and } \lim_{n \to \infty} \frac{x_n - x_0}{t_n} = h.$$

(i) Since $x_n \in U_0$, $g_j(x_n) = g_j(x_0)$, $1 \le j \le m$, we can write

$$d_H g_j(x_0;h) = \frac{g_j(x_n) - g_j(x_0)}{t_n} - \left(\frac{g_j(x_n) - g_j(x_0)}{t_n} - d_H g_j(x_0;h)\right)$$
$$= -\left(\frac{g_j\left(x_0 + t_n \frac{x_n - x_0}{t_n}\right) - g_j(x_0)}{t_n} - d_H g_j(x_0;h)\right).$$

Since $(x_n - x_0)/t_n \to h$ and g_j is Hadamard semidifferentiable at x_0, the right-hand side goes to zero and $d_H g_j(x_0;h) = 0$.

(ii) From (i) since $d_H g_j(x_0;h) = \nabla g_j(x_0) \cdot h$. □

When g is affine, that is, $g(x) = Bx - b$ for an $m \times n$ matrix B and a vector $b \in \mathbb{R}^m$, the cone $T_U(x)$ has been characterized by expression (7.3) of Theorem 7.1 in Chapter 4:

$$T_U(x) = \left\{h \in \mathbb{R}^n : Bh = 0\right\} = \operatorname{Ker} B$$

and equality occurs in (2.2). The necessary optimality condition for the existence of an $x_0 \in U$ that minimizes f becomes

$$\exists x_0 \in U \text{ such that } \nabla f(x_0) \in (\operatorname{Ker} B)^* = (\operatorname{Ker} B)^{\perp} = \operatorname{Im} B^{\top}.$$

This is equivalent to

$$\exists x_0 \in \mathbb{R}^n, \exists \lambda \in \mathbb{R}^m \text{ such that } \begin{cases} \nabla f(x_0) + B^{\top}\lambda = 0 \\ Bx_0 = b. \end{cases}$$

We shall see later that this condition is slightly stronger than the one that we will get from the Lagrange multipliers theorem. It does not require that $Dg(x_0) = B$ be surjective or be of maximum rank.

For a general nonlinear vector function g, to prove (2.2) in the other direction is more difficult and we shall resort to the implicit function theorem that we recall in the next section.

2.2 Jacobian Matrix and Implicit Function Theorem

It is convenient to identify the sequence of functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, with the vector-valued function

$$x \mapsto g(x) \stackrel{\text{def}}{=} (g_1(x), \dots, g_n(x)) = \sum_{j=1}^{m} g_j(x) e_j^m : \mathbb{R}^n \to \mathbb{R}^n,$$

where $\{e_j^m : 1 \le j \le m\}$ is the canonical orthonormal basis in $\mathbb{R}^m$. We have seen in section 5 of Chapter 3 the various notions of semidifferentials and differentials of a real-valued and vector-valued functions as well as the notions of *Jacobian map* and *Jacobian matrix* (see Definition 3.7 of Chapter 3). We complete the definitions for vector-valued functions $g : \mathbb{R}^n \to \mathbb{R}^m$.

Definition 2.1.
g is *Fréchet differentiable* at x if each g_i is *Fréchet differentiable* at x. □

Remark 2.1.
This definition is completely equivalent to

(i) for all $v \in \mathbb{R}^n$, $d_H g(x;v)$ exists in $\mathbb{R}^m$ and

(ii) the map $v \mapsto d_H g(x;v) : \mathbb{R}^n \to \mathbb{R}^m$ is linear. □

When g is Fréchet differentiable at x, the map

$$v \mapsto Dg(x)v \stackrel{\text{def}}{=} d_H g(x;v) : \mathbb{R}^n \to \mathbb{R}^m$$

is linear. By introducing the canonical orthonormal bases $\{e_i^m : 1 \le i \le m\}$ of $\mathbb{R}^m$ and $\{e_j^n : 1 \le j \le n\}$ of $\mathbb{R}^n$, the partial derivative of the ith component of g with respect to the jth component of x is given by

$$\partial_j g_i(x) = e_i^m \cdot Dg(x) e_j^n, \quad 1 \le i \le m,\ 1 \le j \le n.$$

We shall also use the notation $\partial g_i / \partial x_j(x)$. It is easy to check that

$$v = \sum_{j=1}^m v_j\, e_j^m \mapsto Dg(x)v = \sum_{i=1}^n [Dg(x)v]_i\, e_i^m = \sum_{i=1}^n \left[\sum_{j=1}^m \partial_j g_i(x)\, v_j \right] e_i^m. \tag{2.7}$$

Associate with$Dg(x)$ the *Jacobian matrix* $[Dg(x)]$ of first-order partial derivatives of the components of g:

$$[Dg(x)]_{ij} \stackrel{\text{def}}{=} [Dg(x)e_j^n]_i, \text{ that is, } [Dg(x)] = \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(x) & \dots & \dots & \dfrac{\partial g_1}{\partial x_n}(x) \\ \dfrac{\partial g_2}{\partial x_1}(x) & \ddots & & \dfrac{\partial g_2}{\partial x_n}(x) \\ \vdots & & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(x) & \dots & \dots & \dfrac{\partial g_m}{\partial x_n}(x) \end{bmatrix}. \tag{2.8}$$

It is an $m \times n$ matrix. For simplicity, we shall use the same notation $Dg(x)$ for the Jacobian map of Definition 3.7 in Chapter 3 and the Jacobian matrix $[Dg(x)]$ of (2.8).

Definition 2.2.
Given $\Omega \subset \mathbb{R}^n$ open, the function g is of *class* $C^{(1)}$ on Ω if each component g_i of g is of class $C^{(1)}$ on Ω. □

Definition 2.3.
When g is Gateaux differentiable in x_0, we say that x_0 is a *regular point* of g if the map $Dg(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ is surjective[1] (or equivalently, if the Jacobian matrix has maximum rank m). Otherwise, we say that x_0 is a *singular point* of g. □

[1] We also say that g is a *submersion* at the point x_0.

Remark 2.2.
In order to have regular points for the map $g : \mathbb{R}^n \to \mathbb{R}^m$ it is necessary that $n \geq m$. In the case $m = 1$, a point x_0 is regular if and only if $\nabla g(x_0) \neq 0$. □

Finally, recall the *implicit function theorem* (see Theorem 2.1 in Appendix A and its proof).

Theorem 2.1 (Implicit function theorem)**.** *Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be of class $C^{(1)}$ in a neighborhood of a regular point $x_0 \in \mathbb{R}^n$ (that is, $Dg(x_0)$ is surjective). Then for each $h \in \mathbb{R}^n$, there exist $t_0 > 0$ and a function $x :]-t_0, t_0[\to \mathbb{R}^n$ of class $C^{(1)}$ such that*

$$\begin{cases} x(0) = x_0 \\ x'(0) = h \end{cases} \text{and } g(x(t)) = g(x_0) + t\,Dg(x_0)h, \quad -t_0 < t < t_0. \tag{2.9}$$

2.3 Lagrange Multipliers Theorem

We now use Theorem 2.1 to complete Lemma 2.1 of section 2.1.

Lemma 2.2. *Let $x_0 \in \mathbb{R}^n$, $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \leq j \leq m$, be functions of class $C^{(1)}$ on a neighborhood of x_0 and define the set*

$$U_0 \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : g_j(x) = g_j(x_0),\ 1 \leq j \leq m\right\}.$$

(i) *If x_0 is a regular point of g, then*

$$T_{U_0}(x_0) = \{h : Dg(x_0)h = 0\} = \{h : \nabla g_j(x_0) \cdot h = 0,\ 1 \leq j \leq m\}. \tag{2.10}$$

(ii) *If $x_0 \in \mathbb{R}^n$ is a singular point of g, there exists $\lambda = (\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$ such that*

$$Dg(x_0)^\top \lambda = \sum_{j=1}^m \lambda_j \nabla g_j(x_0) = 0. \tag{2.11}$$

Proof. (i) The right-hand side of (2.10) is equal to Ker $Dg(x_0)$. By Lemma 2.1, $T_{U_0}(x_0) \subset$ Ker $Dg(x_0)$. We now prove that any $h \in$ Ker $Dg(x_0)$ is an admissible direction. Since x_0 is a regular point, by Theorem 2.1,

$$\begin{gathered} \exists t_0 > 0,\ \exists x :]-t_0, t_0[\to \mathbb{R}^n \text{ of class } C^{(1)} \text{ such that} \\ x(0) = x_0,\ x'(0) = h, \text{ and } g(x(t)) = g(x_0) + t\,Dg(x_0)h, \quad -t_0 < t < t_0. \end{gathered}$$

Therefore, if $h \in$ Ker $Dg(x_0)$, then $Dg(x_0)h = 0$,

$$\forall t \in]-t_0, t_0[,\ g(x(t)) = g(x_0) \quad \Rightarrow x(t) \in U_0$$

and

$$\lim_{t \to 0} \frac{x(t) - x(0)}{t} = x'(0) = h.$$

By definition of an element of $T_{U_0}(x_0)$, $h \in T_{U_0}(x_0)$ and Ker $Dg(x_0) \subset T_{U_0}(x_0)$.

(ii) When x_0 is a singular point of g, the map $Dg(x_0)$ is not surjective and there exists $\lambda = (\lambda_1,\ldots,\lambda_m) \neq (0,\ldots,0)$ such that

$$\forall h \in \mathbb{R}^n, \quad (\lambda_1,\ldots,\lambda_m)\cdot Dg(x_0)h = 0 \quad \Rightarrow \sum_{j=1}^{m} \lambda_j \nabla g_j(x_0) = 0,$$

or equivalently $Dg(x_0)^\top \lambda = 0$. $\square$

We now give several versions of the main theorem.

Theorem 2.2 (Lagrange[2] multipliers theorem)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *and*

$$U \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : g_j(x) = 0, \quad 1 \le j \le m\right\} \tag{2.12}$$

specified by functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$. *Assume that*

(a) f *has a local minimum at* $x_0 \in U$ *with respect to* U,

(b) f *is Fréchet differentiable at* x_0,

(c) $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, *are of class* $C^{(1)}$ *on a neighborhood of* x_0.

Then, we have the following properties.

(i) *If* x_0 *is a regular point of* g, *then there exists* $\lambda \in \mathbb{R}^m$ *such that*

$$\nabla f(x_0) + Dg(x_0)^\top \lambda = 0 \;\text{ and }\; g(x_0) = 0. \tag{2.13}$$

(ii) *If* x_0 *is a singular point of* g, *then there exists* $0 \neq \lambda \in \mathbb{R}^m$ *such that*

$$Dg(x_0)^\top \lambda = 0 \;\text{ and }\; g(x_0) = 0. \tag{2.14}$$

(iii) *There exists* $(\lambda_0,(\lambda_1,\ldots,\lambda_m)) \in \mathbb{R}\times\mathbb{R}^m$, *not all zero, such that* $\lambda_0 \ge 0$ *and*

$$\lambda_0 \nabla f(x_0) + Dg(x_0)^\top \lambda = 0 \;\text{ and }\; g(x_0) = 0. \tag{2.15}$$

(iv) *There exist* $\lambda_0 \ge 0$ *and* $(\lambda_1,\ldots,\lambda_m) \in \mathbb{R}^m$, *not all zero, such that*

$$\lambda_0 \nabla f(x_0) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x_0) = 0 \;\text{ and }\; g_j(x_0) = 0, \;\; 1 \le j \le m, \tag{2.16}$$

or, equivalently,

$$\frac{\partial L}{\partial x_i}(x_0,\lambda) = 0, \;\; 1 \le i \le n, \;\text{ and }\; \frac{\partial L}{\partial \lambda_j}(x_0,\lambda) = 0, \;\; 1 \le j \le m, \tag{2.17}$$

by introducing the Lagrangian

$$L(x,\lambda) \stackrel{\text{def}}{=} \lambda_0 f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) \tag{2.18}$$

for $x \in \mathbb{R}^n$ *and* $\lambda = (\lambda_0,\lambda_1,\ldots,\lambda_m) \in \mathbb{R}\times\mathbb{R}^m$.

[2] J. L. Lagrange [**2**].

(v) *The conditions of parts* (i)–(iv) *are also verified in point* x_0 *of* U *that achieves a local maximum of* f *with respect to* U.

Remark 2.3.
As stated in part (v) of Theorem 2.2, its application yields not only local minima but also local maxima. So condition $\lambda_0 \geq 0$ is *artificial*. Indeed, by multiplying the first equation in (2.16) by -1 and and introducing the new multipliers $\mu_i = -\lambda_i$, we get the existence of multipliers, not all zero, such that $\mu_0 \leq 0$. This would correspond to a local maximizer $x_0 \in U$ of f with respect to U. □

Proof. (i) By Theorem 6.3(ii) of Chapter 4, if f is Fréchet differentiable on a neighborhood of a local minimizer $x_0 \in U$ of f, the necessary condition

$$\nabla f(x_0) \in T_U(x_0)^* \tag{2.19}$$

is verified. Since $x_0 \in U$ is a regular point of $g = (g_1, \dots, g_m)$, then $g(x_0) = 0$, and we get from identity (2.10) of Lemma 2.2(i),

$$T_U(x_0) = \operatorname{Ker} Dg(x_0) = \{h : \nabla g_j(x_0) \cdot h = 0,\ 1 \leq j \leq m\}. \tag{2.20}$$

So, $T_U(x_0)^* = \operatorname{Ker} Dg(x_0)^* = \operatorname{Ker} Dg(x_0)^\perp = \operatorname{Im} Dg(x_0)^\top$ from the characterization of the dual cone in Theorem 7.2(i) of Chapter 4. Therefore, there exists $\alpha \in \mathbb{R}^m, \alpha \neq 0$, such that

$$\nabla f(x_0) - Dg(x_0)^\top \alpha = 0.$$

Choose $\lambda = -\alpha$ to get the first identity in (2.13). Since $x_0 \in U$, $g(x_0) = 0$, and we get the second identity in (2.15).

(ii) If x_0 is a singular point of g, by Lemma 2.2(ii), there exists $\lambda = (\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$ such that

$$Dg(x_0)^\top \lambda = 0.$$

(iii) Combine (i) with $\lambda_0 = 1$ in identities (2.15) and (ii) with $\lambda_0 = 0$ in identities (2.15).

(iv) We explicit identites (2.15) of part (iii) by using the explicit expression of $Dg(x_0)^\top$:

$$\forall \lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m, \quad Dg(x_0)^\top \lambda = \sum_{j=1}^m \lambda_j \nabla g_j(x_0).$$

(v) If $x_0 \in U$ is a local maximizer of f in U, it is sufficient to replace f by $-f$ to come back to the case of a local minimizer. For instance, from part (iii), there exist $\lambda_0 \geq 0$ and $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, not all zero, such that

$$-\lambda_0 \nabla f(x_0) + \sum_{j=1}^m \lambda_j \nabla g_j(x_0) = 0 \ \text{ and } \ g_j(x_0) = 0,\ \ 1 \leq j \leq m. \tag{2.21}$$

Multiplying by -1,

$$\lambda_0 \nabla f(x_0) + \sum_{j=1}^m (-\lambda_j) \nabla g_j(x_0) = 0 \ \text{ and } \ g_j(x_0) = 0,\ \ 1 \leq j \leq m. \tag{2.22}$$

Hence, the condition (for a local minimizer) of part (iii) is verified with the multipliers $\lambda_0' = \lambda_0 \geq 0$ and $\lambda_j' = -\lambda_j \in \mathbb{R}$, $1 \leq j \leq m$, for a local maximizer. □

Remark 2.4.
It is important to notice that when x_0 is a regular point of g, we can always choose $\lambda_0 = 1$. Also note that the conditions of the theorem will always be verified for points of U that are singular points of g whether they correspond to local minima or maxima of f in U or not. □

Example 2.1.
Consider the objective function $f : \mathbb{R}^2 \to \mathbb{R}$ and the constraint function $g : \mathbb{R}^2 \to \mathbb{R}$ of class $C^{(1)}$:

$$x = (x_1, x_2) \mapsto f(x) \stackrel{\text{def}}{=} x_1(1+x_2) \text{ and } x = (x_1, x_2) \mapsto g(x) \stackrel{\text{def}}{=} x_1 + x_2^2,$$
$$U \stackrel{\text{def}}{=} \{x = (x_1, x_2) : g(x) = 0\}$$
$$L(x, \lambda) = \lambda_0 f(x) + \lambda_1 g(x).$$

The gradient $\nabla g(x)$ is

$$\frac{\partial g}{\partial x_1}(x) = 1, \quad \frac{\partial g}{\partial x_2}(x) = 2x_2 \quad \text{and} \quad \nabla g(x) \cdot v = \begin{bmatrix} 1 \\ 2x_2 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

and the map $v \mapsto \nabla g(x) \cdot v : \mathbb{R}^2 \to \mathbb{R}$ is surjective since

$$\forall x = (x_1, x_2), \quad \nabla g(x) = \begin{bmatrix} 1 \\ 2x_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

So we can set $\lambda_0 = 1$ and apply the Lagrange multipliers theorem:

$$L(x, \lambda_1) = f(x) + \lambda_1 g(x) = x_1(1+x_2) + \lambda_1(x_1 + x_2^2)$$
$$\begin{cases} \dfrac{\partial L}{\partial x_1}(x, \lambda_1) = (1+x_2) + \lambda_1 = 0, \quad \dfrac{\partial L}{\partial x_2}(x, \lambda_1) = x_1 + 2\lambda_1 x_2 = 0 \\ \dfrac{\partial L}{\partial \lambda_1}(x, \lambda_1) = x_1 + x_2^2 = 0. \end{cases}$$

This yields

$$\lambda_1 = -(1+x_2) \quad \Rightarrow x_1 = 2x_2(1+x_2) \quad \Rightarrow 2x_2(1+x_2) + x_2^2 = 0$$
$$\Rightarrow 3x_2^2 + 2x_2 = 0 \quad \Rightarrow x_2 = 0 \text{ or } x_2 = -2/3.$$

The two candidates for a local minimum are

$$\begin{cases} a = (0,0), & b = (-\frac{4}{9}, -\frac{2}{3}) \\ f(a) = 0, & f(b) = -\frac{4}{27}. \end{cases}$$

The potential candidate for a global minimum is a. However, at this juncture we cannot conclude since Lagrange gives both local minimizers and maximizers.

Pushing the analysis further, eliminate the variable x_1 from $f(x_1, x_2)$ by using the constraint, to obtain a function of a single real variable

$$h(x_2) = f(-x_2^2, x_2) = -x_2^2(1+x_2)$$

to minimize on $\mathbb{R}$. It is readily seen that

$$\begin{aligned} h(x_2) &\to -\infty \ \text{ when } \ x_2 \to +\infty \\ &\quad +\infty \ \text{ when } \ x_2 \to -\infty. \end{aligned}$$

As a result, there is no finite global minimum or maximum. Since h is of class $C^{(2)}$, the local minima and maxima are characterized as follows:

$$\begin{array}{ll} \underline{\text{minima}} & \underline{\text{maxima}} \\ \dfrac{dh}{dx_2}(x_2) = 0 & \dfrac{dh}{dx_2}(x_2) = 0 \\ \dfrac{d^2h}{dx_2^2}(x_2) \geq 0 & \dfrac{d^2h}{dx_2^2}(x_2) \leq 0. \end{array}$$

We get,

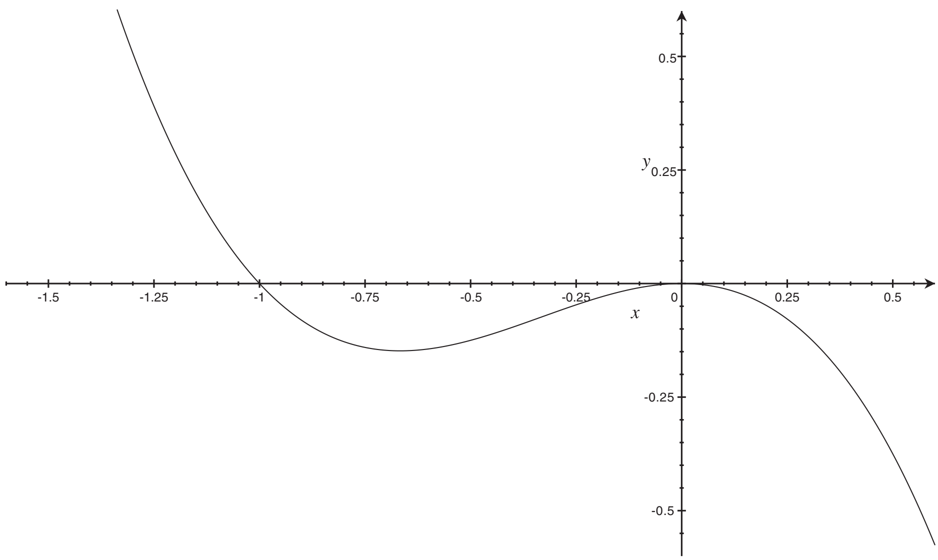


Figure 5.1. *Function $y = h(x)$.*

$$\frac{dh}{dx_2}(x_2) = -x_2(2+3x_2), \quad \frac{d^2h}{dx_2^2}(x_2) = -2(3x_2+1)$$

and

$$\begin{array}{lll} \underline{\text{local minimum}} & x_2^* = -\dfrac{2}{3}, & \dfrac{d^2h}{dx_2^2}(x_2) = +2 \\ \underline{\text{local maximum}} & x_2^* = 0, & \dfrac{d^2h}{dx_2^2}(x_2^*) = -2. \end{array}$$

Indeed, it is easy to see that

$$\frac{d^2h}{dx_2^2}(x_2)\begin{cases} >0, & x_2 < -\dfrac{1}{3} \\ =0, & x_2 = -\dfrac{1}{3} \\ <0, & x_2 > -\dfrac{1}{3} \end{cases}.$$

□

Example 2.2.
Consider the $C^{(1)}$ objective and constraint functions $f, g : \mathbb{R}^n \to \mathbb{R}$

$$f(x) = a \cdot x,\, a \in \mathbb{R}^n \text{ and } g(x) = \|x\|^2 - 1$$

and the maximization problem

$$\sup_{\substack{x\in\mathbb{R}^n\\ g(x)=0}} f(x) \text{ or } \sup_{x\in U} f(x) \text{ or } \inf_{x\in U} -f(x), \quad U \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : g(x) = 0\}.$$

We have changed f to $-f$ to go back to a minimization problem. Since f is continuous and the set U is the sphere of radius one that is compact, by the Weierstrass theorem, there exist at least a minimizer and a maximizer in U.

All the points of the constraint $g(x) = 0$ are regular points of g since

$$\nabla g(x) = 2x \neq 0 \text{ for } 0 = g(x) = \|x\|^2 - 1$$

and we can set $\lambda_0 = 1$:

$$0 = -\nabla f(x) + \lambda \nabla g(x) = -a + 2\lambda x \text{ and } \|x\|^2 = 1.$$

If $a = 0$, then $f(x) = 0$ for all points of U that are both global minimizers and maximizers. If $a \neq 0$, then

$$0 \neq \|a\| = 2\,|\lambda|\,\|x\| = 2\,|\lambda| \quad \Rightarrow \lambda \neq 0$$

$$x = \frac{1}{2\lambda}a \quad \Rightarrow x = \pm\frac{a}{\|a\|} \quad \Rightarrow f(a/\|a\|) = \|a\| \text{ and } f(-a/\|a\|) = -\|a\|.$$

Observe that

$$\begin{aligned} |f(x)| = |a\cdot x| \le \|a\|\,\|x\| \; &\text{ by Schwartz's inequality} \\ \Rightarrow \forall x \text{ such that } \|x\| = 1,\; &|f(x)| \le \|a\|. \end{aligned}$$

We conclude that

$$\begin{aligned} x &= -\frac{a}{\|a\|} \quad && \text{minimizes } f(x) \text{ for } \|x\|^2 = 1 \\ x &= \frac{a}{\|a\|} && \text{maximizes } f(x) \text{ for } \|x\|^2 = 1. \end{aligned}$$

□

Example 2.3.
Consider the constraint function

$$g(x,y) = (x^2+y^2-1)^2.$$

Then

$$\nabla g(x,y) = 2(x^2+y^2-1)\begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

and for all

$$(x,y) \in U \stackrel{\text{def}}{=} \{(x,y) : g(x,y) = 0\}$$

we have

$$\forall (x,y) \in U, \quad \nabla g(x,y) = 0,$$

and no point of U is regular for g. Given an arbitrary objective function f, the conclusions of the Lagrange multipliers theorem are verified for all points of the set of constraints U. Thus, squaring a constraint results in no information on the local minimizers or maximizers of f. □

Example 2.4.
Find the shortest and largest distance from the origin $(0,0)$ to the ellipse

$$U = \left\{ x = (x_1,x_2) \in \mathbb{R}^2 : g(x) \stackrel{\text{def}}{=} 5x_1^2 + 4x_1x_2 + x_2^2 - 10 = 0 \right\}.$$

Choose as objective function the square of the distance from the origin to x,

$$f(x) = \|x\|^2,$$

where x is constrained by the equation

$$g(x) = Ax \cdot x - 10 = 0 \;\text{ where }\; A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

to be on the ellipse.
<u>Existence</u>: The function f is continuous. U is bounded and hence compact. Indeed,

$$10 = 5x_1^2 + 4x_1x_2 + x_2^2 = (x_2+2x_1)^2 + x_1^2 \quad \Rightarrow |x_1| \le \sqrt{10} \text{ and } |x_2| \le 3\sqrt{10}.$$

<u>Regular points of g</u>:

$$\nabla g(x) = 2Ax$$
$$x \text{ is a regular point if } \; 2Ax \neq 0.$$

In particular, any point $x \in U$ verifies $2Ax \cdot x = 20$ and $Ax \neq 0$. So, all the points of the constraint g are regular.
<u>Lagrange multipliers theorem</u>:

$$\nabla f(x) + \lambda \nabla g(x) = 0 \quad \Rightarrow 2x + 2\lambda Ax = 0 \quad \Rightarrow [I+\lambda A]x = 0$$
$$Ax \cdot x - 10 = 0.$$

Then

$$f(x) = \|x\|^2 = -\lambda Ax \cdot x = -10\lambda$$

where

$$[I + \lambda A]x = 0, \quad Ax \cdot x = 10. \tag{2.23}$$

Since A is invertible, the solution $x = 0$ is not in U. So we are looking for points $x \neq 0$ such that $[I + \lambda A]x = 0$. This is possible only if

$$\det(I + \lambda A) = 0$$

$$\begin{vmatrix} 1+5\lambda & 2\lambda \\ 2\lambda & 1+\lambda \end{vmatrix} = (1+5\lambda)(1+\lambda) - 4\lambda^2 = 0 \quad \Rightarrow \lambda = -3 \pm 2\sqrt{2}.$$

Since the function f is continuous and U is compact, we have the existence of a global minimum and a global maximum:

$$\begin{array}{llll} \text{maximum} & \lambda = -3 - 2\sqrt{2}, & f(x) = 10(3 + 2\sqrt{2}) \simeq 58.28, & \text{distance } \simeq 7.63 \\ \text{minimum} & \lambda = -3 + 2\sqrt{2}, & f(x) = 10(3 - 2\sqrt{2}) \simeq 1.72, & \text{distance } \simeq 1.31. \end{array}$$

To compute the corresponding points x_1 and x_2, use identity (2.23),

$$2\lambda x_1 + (1+\lambda)x_2 = 0,$$

and substitute

$$x_1 = -\frac{(1+\lambda)}{2\lambda} x_2$$

in the constraint function. The problem reduces to finding solution x_2 of the equation

$$\left[5\left(\frac{1+\lambda}{2\lambda}\right)^2 - 4\left(\frac{1+\lambda}{2\lambda}\right) + 1\right] x_2^2 = 10 \quad \Rightarrow \frac{\lambda^2 + 2\lambda + 5}{4\lambda^2} x_2^2 = 10.$$

Therefore, in each case, we have two solutions symmetric with respect to the origin as could have been expected for an ellipse. □

Example 2.5.
Find the minimum distance between the circle C and the line L:

$$C \stackrel{\text{def}}{=} \left\{(x_1, y_1) \in \mathbb{R}^2 : x_1^2 + y_1^2 = 1\right\}, \quad L \stackrel{\text{def}}{=} \left\{(x_2, y_2) \in \mathbb{R}^2 : x_2 + y_2 = 5\right\}.$$

Choose as objective function the square of the distance

$$f(x_1, y_1, x_2, y_2) \stackrel{\text{def}}{=} (x_1 - x_2)^2 + (y_1 - y_2)^2$$

between points $(x_1, y_1) \in C$ and $(x_2, y_2) \in L$.

Do we have minimizers? The set of constraints

$$U \stackrel{\text{def}}{=} \{(x_1, y_1, x_2, y_2) \in \mathbb{R}^4 : (x_1, y_1) \in C \text{ and } (x_2, y_2) \in L\}$$

is not compact. However, it can be shown that the function f has a bounded lower section in U. Indeed, choose the point

$$\bar{x} = (\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{x}_2) = (1,0,0,5) \in U \quad \Rightarrow f(\bar{x}) = (1-0)^2 + (0-5)^2 = 26.$$

By construction, this lower section

$$U_{26} = \{x \in U \,:\, f(x) \le 26\} \ni (1,0,0,5)$$

is not empty. By definition,

$$x_1^2 + y_1^2 = 1.$$

Moreover,

$$\begin{aligned}(x_2^2 + y_2^2)^{1/2} &\le [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2} + (x_1^2 + y_1^2)^{1/2} \\ &\le (26)^{1/2} + 1 \le 7 \\ &\Rightarrow \quad x_1^2 + y_1^2 + x_2^2 + y_2^2 \le 1 + 49 = 50.\end{aligned}$$

The set U_{26} is nonempty (by construction) and bounded. The function f that is continuous has a compact lower section. By Theorem 5.3 of Chapter 2, we conclude that f has a global minimizer in U.

The constraint functions are

$$\begin{aligned}g_1(x_1, y_1, x_2, y_2) &= x_1^2 + y_1^2 - 1 = 0 \\ g_2(x_1, y_1, x_2, y_2) &= x_2 + y_2 - 5 = 0\end{aligned}$$

Regular points:

$$x \mapsto Dg(x) = Dg(x_1, y_1, x_2, y_2) = \begin{bmatrix} 2x_1 & 2y_1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} : \mathbb{R}^4 \to \mathbb{R}^2$$

Dg is surjective if $(x_1, y_1) \neq (0,0)$. Since $(0,0,x_2,y_2)$ is not in the first constraint, all points of U are regular.

Gradient of f:

$$\nabla f(x) = \nabla f(x_1, y_1, x_2, y_2) = \begin{bmatrix} 2(x_1 - x_2) \\ 2(y_1 - y_2) \\ -2(x_1 - x_2) \\ -2(y_1 - y_2) \end{bmatrix}.$$

Computations:

$$\nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0$$

$$2(x_1 - x_2) + \lambda_1 2x_1 = 0 \tag{2.24}$$

$$2(y_1 - y_2) + \lambda_1 2y_1 = 0 \tag{2.25}$$

$$-2(x_1 - x_2) + \lambda_2 = 0 \tag{2.26}$$

$$-2(y_1 - y_2) + \lambda_2 = 0 \tag{2.27}$$

and the contraints

$$x_1^2 + y_1^2 - 1 = 0 \tag{2.28}$$

$$x_2 + y_2 - 5 = 0. \tag{2.29}$$

If $\lambda_1 = 0$, we get from (2.24) and (2.26) and then from (2.25) and (2.27)

$$0 = 2(x_1 - x_2) = \lambda_2 = 2(y_1 - y_2) \quad \Rightarrow x_1 = x_2 \text{ and } y_1 = y_2 = 5 - x_2 = 5 - x_1$$

$$\Rightarrow \left| \begin{array}{l} 1 = x_1^2 + y_1^2 = x_1^2 + (5-x_1)^2 = x_1^2 + x_1^2 - 10x_1 + 25 \\ 1 = 2\left[x_1^2 - 5x_1 + \dfrac{25}{2}\right] = 2\left[\left(x_1 - \dfrac{5}{2}\right)^2 - \dfrac{25}{4} + \dfrac{25}{2}\right] \\ 1 = 2\left[\left(x_1 - \dfrac{5}{2}\right)^2 + \dfrac{25}{4}\right] \geq \dfrac{25}{2} \end{array} \right.$$

and finally a contradiction. Since $\lambda_1 \neq 0$, by adding (2.24) and (2.26) on one hand and (2.25) and (2.27) on the other hand, we now get

$$\begin{cases} 2\lambda_1 x_1 + \lambda_2 = 0 \\ 2\lambda_1 y_1 + \lambda_2 = 0 \end{cases} \Rightarrow y_1 = x_1.$$

From (2.24) and (2.25),

$$(1+\lambda_1)x_1 = x_2 \text{ and } (1+\lambda_1)y_1 = y_2 \quad \Rightarrow y_2 = x_2. \tag{2.30}$$

Therefore,

$$\begin{aligned} x_1 = y_1 \quad &\Rightarrow x_1 = y_1 = \pm\frac{\sqrt{2}}{2} \text{ from (2.28)} \\ x_2 = y_2 \quad &\Rightarrow x_2 = y_2 = \frac{5}{2} \text{ from (2.29)} \end{aligned}$$

and there are two cases: either

$$x_1 = y_1 = \frac{\sqrt{2}}{2}, \qquad x_2 = y_2 = \frac{5}{2},\ f = 2\left(\frac{5-\sqrt{2}}{2}\right)^2 \quad \Rightarrow \text{distance } = \frac{5\sqrt{2}-2}{2}$$

or

$$x_1 = y_1 = -\frac{\sqrt{2}}{2}, \qquad x_2 = y_2 = \frac{5}{2},\ f = 2\left(\frac{5+\sqrt{2}}{2}\right)^2 \quad \Rightarrow \text{distance } = \frac{5\sqrt{2}+2}{2}.$$

The Lagrange multipliers theorem yields two candidates for local minima. Since we know that there exists a global minimizer, it is achieved by the point

$$x_1 = y_1 = \frac{\sqrt{2}}{2},\ x_2 = y_2 = \frac{5}{2} \text{ and } f = 2\left(\frac{5-\sqrt{2}}{2}\right)^2 \quad \Rightarrow \text{ distance } = \frac{5\sqrt{2}-2}{2}.$$

The other solution is neither a local minimum nor a local maximum. □

Example 2.6.
Consider $n \geq 1$ energy plants. The plant i produces p_i Watts at the cost of $c_i(p_i)$\$ as a function of the output power p_i in Watts. The cost $c_i = c_i(p_i)$ is a function of the

production p_i (see Figure 5.2). Let P be the total demand in Watts. In practice, the curves c_i are monotonically increasing. We seek a distribution $(p_1, \ldots, p_n)$ of the production that will exactly meet the global demand P, that is,

$$g(p) \stackrel{\text{def}}{=} P - \sum_{i=1}^{n} p_i = 0$$

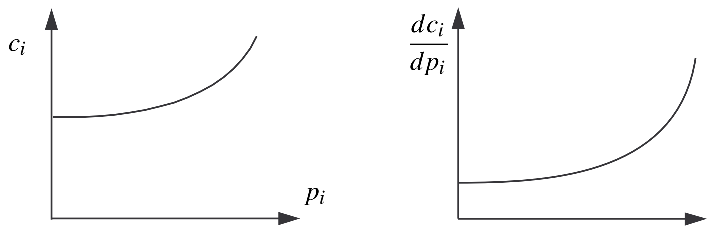


Figure 5.2. *Production cost $c_i(p_i)$ as a function of the output power p_i.*

while minimizing the total production cost

$$f(p_1, \ldots, p_n) \stackrel{\text{def}}{=} \sum_{i=1}^{n} c_i(p_i).$$

The Lagrangian associated with this problem is

$$L(p, \lambda) = \lambda_0 \sum_{i=1}^{n} c_i(p_i) + \lambda_1 \left(P - \sum_{i=1}^{n} p_i \right).$$

It is readily seen that all points of the constraint set U are regular points of g since

$$\nabla g(p) = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} \neq 0 \quad \text{for } p \in U.$$

So we can choose $\lambda_0 = 1$:

$$\frac{\partial L}{\partial p_i} = \frac{dc_i}{dp_i}(p_i) - \lambda_1 = 0 \quad \Rightarrow \lambda_1 = \frac{dc_i}{dp_i}(p_i), \; 1 \leq i \leq n.$$

The necessary condition to minimize the cost is

$$\lambda_1 = \frac{dc_1}{dp_1}(p_1) = \frac{dc_2}{dp_2}(p_2) = \cdots = \frac{dc_n}{dp_n}(p_n).$$

In practice, the curves $c_i(p_i)$ and dc_i/dp_i are known. So we draw the n curves dc_i/dp_i on the same graph (see Figure 5.3). Given a level λ_1, compute the p_i's for each i. Then adjust the value (level) of λ_1 so that the sum $\sum_{i=1}^{n} p_i$ is equal to P. $\square$

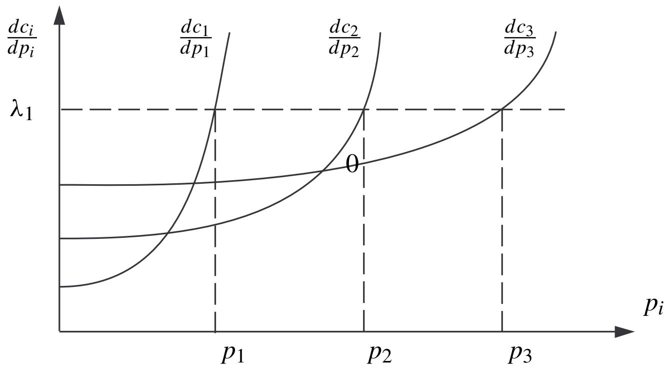


Figure 5.3. *Determination of the output power p_i as a function of λ_1.*

3 Inequality Contraints: Karush–Kuhn–Tucker Theorem

Consider now U specified by inequality constraints:

$$U \overset{\text{def}}{=} \left\{x \in \mathbb{R}^n : g_j(x) \le 0, \quad 1 \le j \le m\right\}. \tag{3.1}$$

For $x \in U$, let

$$I(x) \overset{\text{def}}{=} \left\{j : 1 \le j \le m \text{ and } g_j(x) = 0\right\} \tag{3.2}$$

be the set of indices of the *active* constraints ($g_j(x) = 0$). The other indices correspond to *inactive* constraints ($g_j(x) < 0$).

As for equality constraints, if f is Fréchet differentiable at a point x_0 achieving a local minimum of f on U, by Theorem 6.3(iii) of Chapter 4, the dual necessary optimality condition is equivalent to

$$\nabla f(x_0) \in T_U(x_0)^*. \tag{3.3}$$

So, we first characterize $T_U(x_0)$.

Lemma 3.1. *Given the functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, consider the set U of constraints given by* (3.1)*, and $x_0 \in U$. Assume that the functions g_j are Fréchet differentiable at x_0 and introduce the cone*

$$C \overset{\text{def}}{=} \left\{h \in \mathbb{R}^n : \forall j \in I(x_0), \nabla g_j(x_0) \cdot h < 0\right\}, \tag{3.4}$$

where $I(x_0)$ is the set (3.2) *of indices of the active constraints at x_0.*

(i) *Then*

$$C \subset T_U(x_0) \subset \left\{h \in \mathbb{R}^n : \forall j \in I(x_0), \nabla g_j(x_0) \cdot h \le 0\right\}. \tag{3.5}$$

(ii) *If $C \neq \varnothing$ or $I(x_0) = \varnothing$, then*

$$\overline{C} = T_U(x_0) = \left\{h \in \mathbb{R}^n : \forall j \in I(x_0), \nabla g_j(x_0) \cdot h \le 0\right\}. \tag{3.6}$$

(iii) *If $C = \varnothing$ and $I(x_0) \neq \varnothing$, then there exists $\{\lambda_j\}_{j\in I(x_0)}$, $\lambda_j \geq 0$, not all zero, such that*

$$\sum_{j\in I(x_0)} \lambda_j \nabla g_j(x_0) = 0. \tag{3.7}$$

Remark 3.1.
It is readily seen from the proof of part (i) of Lemma 3.1 that

$$\left\{h \in \mathbb{R}^n : \forall j \in I(x_0), d_H g_j(x_0;h) < 0\right\} \subset T_U(x_0)$$
$$T_U(x_0) \subset \left\{h \in \mathbb{R}^n : \forall j \in I(x_0), d_H g_j(x_0;h) \leq 0\right\}$$

with the weaker assumption that the g_j's be only Hadamard semidifferentiable at x_0. However, this does not seem to be sufficient to prove parts (ii) and (iii). □

Proof. We first prove that $T_U(x_0) = T_V(x_0)$, where

$$V \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : \forall j \in I(x_0),\ g_j(x) \leq g_j(x_0)\right\}$$

and we prove the lemma with $T_V(x_0)$ in place of $T_U(x_0)$. Indeed, since $g_j(x_0) = 0$ for all $j \in I(x_0)$, $U \subset V$ and $T_U(x_0) \subset T_V(x_0)$. In the other direction, for all $j \notin I(x_0)$, we have $g_j(x_0) < 0$. Since the g_j's are Fréchet differentiable at x_0, they are continuous at x_0. So, there exists $r > 0$ such that

$$B_r(x_0) \subset \left\{x \in \mathbb{R}^n : g_j(x) < 0,\ \ 1 \leq j \leq m,\ j \notin I(x_0)\right\}.$$

By definition, for all $h \in T_V(x_0)$,

$$\exists \{t_n > 0\}, t_n \searrow 0, \quad \exists \{x_n\} \subset V, \ \lim_{n\to\infty} \frac{x_n - x_0}{t_n} = h.$$

Since $x_n \to x_0$, there exists N such that, for all $n > N$, $x_n \in B_r(x_0)$ and hence $x_n \in B_r(x_0) \cap V$. We have for $n > N$,

$$\begin{cases} g_j(x_n) \leq g_j(x_0) = 0, & \text{for } j \in I(x_0), \\ g_j(x_n) < 0, & \text{for } j \notin I(x_0), \end{cases}$$

which implies that $x_n \in U$, $h \in T_U(x_0)$, and $T_V(x_0) \subset T_U(x_0)$. So we can use V in place of U for the proof.

(i) When $I(x_0) = \varnothing$, $V = \mathbb{R}^n$, $T_V(x_0) = \mathbb{R}^n$ and, moreover, the right-hand side of (3.5) is $\mathbb{R}^n$. For $I(x_0) \neq \varnothing$, let $h \in T_V(x_0)$. Then, by definition of $T_V(x_0)$,

$$\exists \{t_n > 0\}, t_n \searrow 0, \quad \exists \{x_n\} \subset V, \ \lim_{n\to\infty} \frac{x_n - x_0}{t_n} = h.$$

By definition of V, for all $j \in I(x_0)$, $g_j(x_n) - g_j(x_0) \leq 0$ and

$$\begin{aligned}\nabla g_j(x_0)\cdot h &= \frac{g_j(x_n) - g_j(x_0)}{t_n} - \left[\frac{g_j(x_n) - g_j(x_0)}{t_n} - \nabla g_j(x_0)\cdot h\right] \\ &\leq -\left[\frac{g_j\left(x_0 + t_n \frac{x_n - x_0}{t_n}\right) - g_j(x_0)}{t_n} - \nabla g_j(x_0)\cdot h\right].\end{aligned}$$

Since g_j is Fréchet (Hadamard) differentiable at x_0 and $(x_n - x_0)/t_n \to h$, the right-hand side goes to zero and

$$\forall j \in I(x_0), \quad \nabla g_j(x_0) \cdot h \le 0.$$

So the second part of relation (3.5) is verified.

It remains to prove that $C \subset T_V(x_0)$. If $C = \varnothing$, there is nothing to prove. Otherwise, for all $h \in C$, we have for all $j \in I(x_0)$, $\nabla g_j(x_0) \cdot h < 0$ and

$$\alpha \stackrel{\text{def}}{=} \sup_{j \in I(x_0)} \nabla g_j(x_0) \cdot h < 0.$$

Associate with h the trajectory $x(t) = x_0 + th$. Since each g_j is Fréchet (Hadamard) differentiable at x_0, there exists $\delta > 0$ such that for all $0 < t < \delta$,

$$\forall j \in I(x_0), \quad \left| \frac{g_j(x_0 + th) - g_j(x_0)}{t} - \nabla g_j(x_0) \cdot h \right| < -\frac{\alpha}{2}.$$

Hence, for $0 < t < \delta$,

$$\begin{aligned} g_j(x(t)) - g_j(x_0) &= t \, \nabla g_j(x_0) \cdot h + t \left[\frac{g_j(x(t)) - g_j(x_0)}{t} - \nabla g_j(x_0) \cdot h \right] \\ &< t\,\alpha - t\,\frac{\alpha}{2} = t\,\frac{\alpha}{2} < 0 \\ \Rightarrow \forall 0 < t < \delta, &\quad x(t) \in V \text{ and } \frac{x(t) - x_0}{t} = h \quad \Rightarrow h \in T_V(x_0) \end{aligned}$$

and $C \subset T_U(x_0)$.

(ii) For $I(x_0) = \varnothing$, we have seen in (i) that $T_V(x_0) = \mathbb{R}^n$. Moreover, by definition, $C = \mathbb{R}^n$. Thus, identity (3.6) is verified. For $I(x_0) \neq \varnothing$ and for each $j \in I(x_0)$, define

$$C_j \stackrel{\text{def}}{=} \{h \in \mathbb{R}^n : \nabla g_j(x_0) \cdot h < 0\}.$$

Since, by assumption, $C \neq \varnothing$, there exists $h^* \in \mathbb{R}^n$ such that

$$\forall j \in I(x_0), \quad h^* \in C_j \quad \Rightarrow \nabla g_j(x_0) \cdot h^* < 0 \text{ and } \nabla g_j(x_0) \neq 0.$$

As a result, $C_j \neq \varnothing$ since $-\nabla g_j(x_0) \in C_j$ and

$$\overline{C_j} = \left\{ h \in \mathbb{R}^n : \nabla g_j(x_0) \cdot h \le 0 \right\}.$$

We now prove that

$$\bigcap_{j \in I(x_0)} \overline{C_j} \subset \overline{\bigcap_{j \in I(x_0)} C_j} = \overline{C}. \tag{3.8}$$

From this, we will be able to conclude from (3.5) in part (i), and from the fact that $T_V(x_0)$ is closed, that

$$C = \bigcap_{j \in I(x_0)} C_j \subset T_V(x_0) \subset \bigcap_{j \in I(x_0)} \overline{C_j} \quad \Rightarrow \bigcap_{j \in I(x_0)} \overline{C_j} \subset \overline{C} \subset T_V(x_0) \subset \bigcap_{j \in I(x_0)} \overline{C_j},$$

and hence identity (3.6).

Indeed, for all $h \in \bigcap_{j\in I(x_0)} \overline{C_j}$, we have

$$h \in \bigcap_{j\in I(x_0)} \overline{C_j} \quad \Rightarrow \forall j \in I(x_0), \quad \nabla g_j(x_0)\cdot h \le 0.$$

Associate with h the sequence

$$h_n = h + \frac{1}{n} h^* \to h, \quad n \ge 1.$$

It is easy to check that for each $n \ge 1$ and each $j \in I(x_0)$,

$$\begin{aligned}\nabla g_j(x_0)\cdot h_n = \nabla g_j(x_0)\cdot h + \frac{1}{n}\nabla g_j(x_0)\cdot h^* < \nabla g_j(x_0)\cdot h \le 0 \\ \Rightarrow \nabla g_j(x_0)\cdot h_n < 0 \quad \Rightarrow h_n \in C_j \\ \Rightarrow h_n \in \bigcap_{j\in I(x_0)} C_j \quad \Rightarrow h_n \to h \in \overline{\bigcap_{j\in I(x_0)} C_j}.\end{aligned}$$

This completes the proof of part (ii).

(iii) If there exists an index $j \in I(x_0)$ such that $\nabla g_j(x_0) = 0$, then the result is true by taking $\lambda_j = 1$ and $\lambda_{j'} = 0$, $j' \ne j$. It remains to deal with the case where, for all $j \in I(x_0)$, $\nabla g_j(x_0) \ne 0$. In the proof of part (ii), we have seen that in that case, for all $j \in I(x_0)$, $C_j \ne \varnothing$. Relabel the sets C_j, $j \in I(x_0)$, in the increasing order of the indices from 1 to r. If $\bigcap_{j=2}^r C_j = \varnothing$, go to $\bigcap_{j=3}^r C_j$. If this intersection is empty, go to $\bigcap_{j=4}^r C_j$. Let s be the first index such that

$$\bigcap_{j=s}^r C_j = \varnothing \text{ and } K \overset{\text{def}}{=} \bigcap_{j=s+1}^r C_j \ne \varnothing \quad \Rightarrow K \cap C_s = \varnothing.$$

This index s exists since $C_j \ne \varnothing$, $1 \le j \le r$, and for $j = r$,

$$\bigcap_{j=r}^r C_j = C_r \ne \varnothing.$$

Hence $K \subset \mathbb{R}^n \backslash C_s$ or, more explicitly,

$$\left\{h \in \mathbb{R}^n : \nabla g_j(x_0)\cdot h < 0, s+1 \le j \le r\right\} \subset \left\{h \in \mathbb{R}^n : \nabla g_s(x_0)\cdot h \ge 0\right\}. \tag{3.9}$$

Since

$$K = \bigcap_{j=s+1}^r C_j \ne \varnothing,$$

we know from the proof of part (ii) that

$$\overline{K} = \left\{h \in \mathbb{R}^n : \nabla g_j(x_0)\cdot h \le 0, s+1 \le j \le r\right\}$$

and from Theorem 5.4(ii) that $\overline{K}^* = K^*$. Then, by taking the dual of both sides of (3.9), we get

$$\left\{h \in \mathbb{R}^n : \nabla g_s(x_0)\cdot h \ge 0\right\}^* \subset K^* = \overline{K}^*$$

or

$$\left\{h \in \mathbb{R}^n : \nabla g_s(x_0) \cdot h \geq 0\right\}^* \subset \left\{h \in \mathbb{R}^n : -\nabla g_j(x_0) \cdot h \geq 0, s+1 \leq j \leq r\right\}^*.$$

But we have already characterized those dual cones in Theorems 5.6(ii) and 7.2(ii) of Chapter 4. Hence,

$$\{\lambda_s \nabla g_s(x_0) : \forall \lambda_s \geq 0\} \subset \left\{ \sum_{j=s+1}^{r} \lambda_j \left(-\nabla g_j(x_0)\right) : \forall \lambda_j \geq 0, s+1 \leq j \leq r \right\}.$$

Therefore, for $\lambda_s = 1$, there exist $\lambda_{s+1} \geq 0, \dots, \lambda_r \geq 0$ such that

$$\nabla g_s(x_0) = \sum_{j=s+1}^{r} \lambda_j \left(-\nabla g_j(x_0)\right).$$

By choosing $\lambda_j = 0$ for all $j \notin I(x_0)$, we finally get (3.7) where $\lambda_s = 1$. This concludes the proof of part (iii). □

In general, when $C = \varnothing$, the three sets in (3.5) of Lemma 3.1(i) are different as shown in the next example.

Example 3.1.
Consider the following constraint functions $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ and the set U:

$$g_1(x) \overset{\text{def}}{=} -x_1, \quad g_2(x) \overset{\text{def}}{=} -x_2, \quad g_3(x) = -(1-x_1)^3 + x_2$$
$$U \overset{\text{def}}{=} \left\{ x = (x_1, x_2) \in \mathbb{R}^2 : \begin{matrix} -x_1 \leq 0, \ -x_2 \leq 0, \\ -(1-x_1)^3 + x_2 \leq 0 \end{matrix} \right\}.$$

Observe that $x_2 \geq 0$ implies that $(1-x_1)^3 \geq x_2 \geq 0$ and $1 - x_1 \geq 0$. So we get $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 1$. At the point $x_0 = (1,0)$, we have $I(1,0) = \{2,3\}$ and

$$\nabla g_1(1,0) = (-1,0), \ \ \nabla g_2(1,0) = (0,-1), \ \ \nabla g_3(1,0) = (0,1).$$

Then, $C = C_2 \cap C_3$ and

$$\begin{aligned} C_2 \cap C_3 &= \{h \in \mathbb{R}^2 : \nabla g_2(1,0) \cdot h < 0 \text{ and } \nabla g_3(1,0) \cdot h < 0\} = \varnothing \\ \overline{C_2} \cap \overline{C_3} &= \{h \in \mathbb{R}^2 : \nabla g_2(1,0) \cdot h \leq 0 \text{ and } \nabla g_3(1,0) \cdot h \leq 0\} \\ &= \{(h_1, 0) : \forall h_1 \in \mathbb{R}\}. \end{aligned} \tag{3.10}$$

Indeed, the two semidifferentials on the right-hand side of (3.10) give

$$\nabla g_2(1,0) \cdot h = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -h_2, \quad \nabla g_3(1,0) \cdot h = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = h_2,$$

which yields $C_2 \cap C_3 = \varnothing$ and $h_2 = 0$ for $\overline{C_2} \cap \overline{C_3}$. It can also be proved that

$$T_U(1,0) = \{(h_1, 0) : \forall h_1 \leq 0\}. \tag{3.11}$$

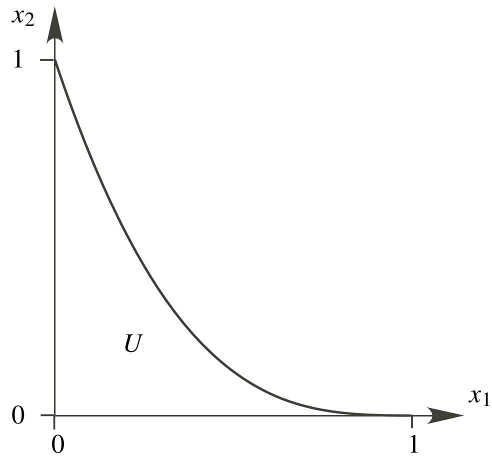


Figure 5.4. *Region U in Example* 3.1.

Indeed, we know that the elements of $T_U(1,0)$ are of the form $(h_1,0)$ for $h_1 \in \mathbb{R}$. We first prove that if $h_1 \le 0$, then $(h_1,0) \in T_U(0,1)$, and in a second step that if $h_1 > 0$, then $(h_1,0) \notin T_U(0,1)$. For $h_1 = 0$, $(0,0) \in T_U(0,1)$. For $h_1 < 0$, choose $t_0 = -1/h_1$ and for $0 < t < t_0$ define $x(t) = (1,0) + t(h_1,0)$. By construction, $0 = 1 + t_0 h_1 < 1 + t h_1 < 1$ and hence $x(t) \in U$. Therefore,

$$\frac{x(t) - (1,0)}{t} = (h_1,0) \quad \Rightarrow (h_1,0) \in T_U(1,0).$$

Finally, by contradiction, if $h_1 > 0$ and $(h_1,0) \in T_U(1,0)$, then

$$\exists \{t_n > 0\}, t_n \searrow 0, \quad \exists \{x_n\} \subset U, \ \lim_{n\to\infty} \frac{x_n - (1,0)}{t_n} = (h_1,0).$$

Therefore, there exists N such that for all $n > N$,

$$\left\| \frac{x_n - (1,0)}{t_n} - (h_1,0) \right\| < \frac{h_1}{2}.$$

In particular, for all $n > N$,

$$\left| \frac{(x_n)_1 - 1}{t_n} - h_1 \right| < \frac{h_1}{2} \quad \Rightarrow -\frac{h_1}{2} < \frac{(x_n)_1 - 1}{t_n} - h_1 \quad \Rightarrow \frac{h_1}{2} < \frac{(x_n)_1 - 1}{t_n} \le 0,$$

since $(x_n)_1 \le 1$. This yields the contradiction and proves (3.11). This simple example shows that, in general, the three sets appearing in (3.5) are all distinct. □

We shall see later that it is only when property (3.5) of Lemma 3.1 is verified that the theorem of Karush–Kuhn–Tucker gives information about the local minima. We now introduce some terminology.

Definition 3.1.
Given functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, let U be defined by (3.1). Assume that the functions g_j are Fréchet differentiable at a point $x_0 \in U$. We say that there is *qualification of the constraints* at x_0 if

$$T_U(x_0) = \left\{h \in \mathbb{R}^n \, : \, \forall i \in I(x_0),\, \nabla g_i(x_0)\cdot h \le 0\right\}, \tag{3.12}$$

where $I(x_0)$ is the set (3.2) of the indices of the active constraints at x_0. □

As we have seen in Lemma 3.1, the condition $C \neq \varnothing$ implies the qualification of the constraints. The question remains open in the case $C = \varnothing$.

Theorem 3.1 (Karush[3]–Kuhn[4]–Tucker[5]). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be the objective function,*

$$U \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : \; g_j(x) \le 0, \quad 1 \le j \le m\} \tag{3.13}$$

the set specified by the functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, and

$$I(x) \stackrel{\text{def}}{=} \{j : 1 \le j \le m \text{ and } g_j(x) = 0\} \tag{3.14}$$

the set of indices of active constraints *at x. Assume that*

(a) *f has a local minimum at $x_0 \in U$ with respect to U,*

(b) *f is Fréchet differentiable at x_0,*

(c) *the functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, are Fréchet differentiable at x_0.*

Then, the following conditions are verified.

(i) *If there is* qualification of the constraints *at x_0 (see Definition* 3.1*), then there exists $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ such that*

$$\lambda_j \ge 0, \;\; g_j(x_0) \le 0, \;\; \lambda_j g_j(x_0) = 0, \quad 1 \le j \le m, \tag{3.15}$$

$$\nabla f(x_0) + \sum_{j=1}^m \lambda_j \nabla g_j(x_0) = 0. \tag{3.16}$$

(ii) *If there is no* qualification of the constraints *at x_0, then there exist reals $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, not all zero, such that*

$$\lambda_j \ge 0, \;\; g_j(x_0) \le 0, \;\; \lambda_j g_j(x_0) = 0, \quad 1 \le j \le m, \tag{3.17}$$

$$\sum_{j=1}^m \lambda_j \nabla g_j(x_0) = 0. \tag{3.18}$$

---

[3] The necessary conditions for inequality-constrained problem were first published in the Masters thesis of William Karush [**1**] in 1939, although they became renowned after a seminal conference paper by Harold W. Kuhn and Albert W. Tucker [**1**] in 1951.

[4] Harold William Kuhn (1925–).

[5] Albert William Tucker (1905–1995).

(iii) *There exist* $\lambda_0, (\lambda_1, \dots, \lambda_m) \in \mathbb{R} \times \mathbb{R}^m$*, not all zero, such that*

$$\lambda_0 \geq 0, \quad \lambda_j \geq 0,\ g_j(x_0) \leq 0,\ \lambda_j g_j(x_0) = 0, \quad 1 \leq j \leq m, \tag{3.19}$$

$$\lambda_0 \nabla f(x_0) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x_0) = 0. \tag{3.20}$$

(iv) *By introducing the* Lagrangian

$$L(x,\lambda) \stackrel{\text{def}}{=} \lambda_0 f(x) + \sum_{j=1}^{m} \lambda_j g_j(x), \tag{3.21}$$

there exist $(\lambda_0, (\lambda_1, \dots, \lambda_m)) \in \mathbb{R} \times \mathbb{R}^m$*, not all zero, such that*

$$\lambda_0 \geq 0, \quad \lambda_j \geq 0,\ \frac{\partial L}{\partial \lambda_j}(x_0,\lambda) \leq 0,\ \lambda_j \frac{\partial L}{\partial \lambda_j}(x_0,\lambda) = 0, \quad 1 \leq j \leq m, \tag{3.22}$$

$$\frac{\partial L}{\partial x_i}(x_0,\lambda) = 0, \quad 1 \leq i \leq n. \tag{3.23}$$

Proof. (i) Since $x_0 \in U$ is a local minimizer of f in U,

$$\nabla f(x_0) \in T_U(x_0)^*.$$

As we have qualification of the constraints at x_0,

$$T_U(x_0) = \left\{ h \in \mathbb{R}^n : -\nabla g_j(x_0) \cdot h \geq 0,\ \forall j \in I(x_0) \right\}.$$

By Theorem 7.2(ii) of Chapter 4, the dual of $T_U(x_0)$ is

$$T_U(x_0)^* = \left\{ \sum_{j \in I(x_0)} -\lambda_j \nabla g_j(x_0) : \lambda_j \geq 0,\ \forall j \in I(x_0) \right\}.$$

So there exist $\{\lambda_j\}_{j \in I(x_0)}$, $\lambda_j \geq 0$, such that

$$\nabla f(x_0) + \sum_{j \in I(x_0)} \lambda_j \nabla g_j(x_0) = 0.$$

Since $g_j(x_0) = 0$ for all $j \in I(x_0)$, we have $\lambda_j g_j(x_0) = 0$. For $j \notin I(x_0)$, choose $\lambda_j = 0$. Thence, the properties (3.15) and (3.16).

(ii) When the constraints are not qualified, then by Lemma 3.1(ii), $C = \varnothing$ and $I(x_0) \neq \varnothing$. By Lemma 3.1(iii), there exist $\{\lambda_j\}_{j \in I(x_0)}$, $\lambda_j \geq 0$, not all zero, such that

$$\sum_{j \in I(x_0)} \lambda_j \nabla g_j(x_0) = 0. \tag{3.24}$$

Since $g_j(x_0) = 0$ for all $j \in I(x_0)$, we have $\lambda_j g_j(x_0) = 0$. For $j \notin I(x_0)$, choose $\lambda_j = 0$. Thence, properties (3.17) and (3.18).

(iii) If we have qualification of the constraints, from part (i) we can choose $\lambda_0 = 1$. If we do not have qualification of the constraints, from part (ii) choose $\lambda_0 = 0$ but we know that there exists at least one λ_j, $1 \leq j \leq m$, such that $\lambda_j > 0$.

(iv) It is another way to rewrite the conditions of part (iii). □

The assumption (3.12) of *qualification of the constraints* allows the choice $\lambda_0 = 1$ as coefficient of $\nabla f(x_0)$ and thus reduces the search of multipliers. However, failure of the qualification of the constraints does not prevent choosing $\lambda_0 = 1$ as illustrated in the next example.

Example 3.2.
Consider the objective function

$$f(x_1,x_2) \stackrel{\text{def}}{=} (x_1-1)^2 + x_2^2$$

and the set U specified by the three constraint functions of Example 3.1:

$$g_1(x) = -x_1, \quad g_2(x) = -x_2, \quad g_3(x) = -(1-x_1)^3 + x_2$$
$$U = \{(x_1,x_2) : g_1(x) \le 0,\, g_2(x) \le 0,\, g_3(x) \le 0\}.$$

We have seen that there is no qualification of the constraints at $(1,0) \in U$ since

$$T_U(1,0) = \{(h_1,0) : \forall h_1 \ge 0\}$$
$$\left\{h \in \mathbb{R}^2 : \nabla g_2(1,0)\cdot h \le 0 \text{ and } \nabla g_3(1,0)\cdot h \le 0\right\} = \{(h_1,0) : \forall h_1 \in \mathbb{R}\}.$$

By choosing $\lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 1$, it is easy to check that

$$\lambda_1 \nabla g_1(1,0) + \lambda_2 \nabla g_2(1,0) + \lambda_3 \nabla g_3(1,0) = 0\begin{bmatrix}-1\\0\end{bmatrix} + 1\begin{bmatrix}0\\-1\end{bmatrix} + 1\begin{bmatrix}0\\1\end{bmatrix} = 0.$$

However,

$$\nabla f(x_1,x_2) = 2(x_1-1,x_2) \quad \Rightarrow \nabla f(1,0) = (0,0)$$

and we can choose $\lambda_0 = 1$ that yields

$$\nabla f(1,0) + \lambda_1 \nabla g_1(1,0) + \lambda_2 \nabla g_2(1,0) + \lambda_3 \nabla g_3(1,0) = 0. \qquad \square$$

Since the qualification of the constraints (3.12) allows the choice of $\lambda_0 = 1$, it is interesting to look for simple sufficient conditions to get it. The next theorem gives three such conditions. The first one is nothing but $C \ne \varnothing$ and the other two imply $C \ne \varnothing$.

Theorem 3.2. *Given functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le m$, let U be defined by* (3.1)*. Assume that the functions g_j are Fréchet differentiable at a point $x_0 \in U$. Then, condition $C \ne \varnothing$ of Lemma* 3.1 *and condition* (3.12) *of qualification of the constraints are verified if any one of the following conditions is verified.*

(i) *There exists $h^* \in \mathbb{R}^n$ such that*

$$\forall j,\, 1 \le j \le m, \quad \nabla g_j(x_0)\cdot h^* < 0. \tag{3.25}$$

(ii) *(Slater's condition) The functions g_j, $1 \le j \le m$, are convex and*

$$\exists x^* \in \mathbb{R}^n \text{ such that } \quad \forall j,\, 1 \le j \le m, \;\; g_j(x^*) < 0. \tag{3.26}$$

(iii) *The point x_0 is a regular point of $g = (g_1,\dots,g_m)$.*

Proof. (i) Since the property is true for all j, $1 \le j \le m$, it is true for all $I(x_0)$ and by Lemma 3.1(ii), $C \neq \varnothing$ where C is defined in (3.4).

(ii) For all $j \in I(x_0)$, $g_j(x_0) = 0$. By convexity of g_j, we have from Theorem 4.1 in Chapter 3

$$g_j(x^*) \ge g_j(x_0) + \nabla g_j(x_0) \cdot (x^* - x_0) = \nabla g_j(x_0) \cdot (x^* - x_0)$$

and

$$\forall j \in I(x), \quad 0 > g_j(x^*) \ge \nabla g_j(x_0) \cdot (x^* - x_0), \quad \Rightarrow h^* \stackrel{\text{def}}{=} x^* - x_0 \in C.$$

Hence $C \neq \varnothing$ and the constraints are qualified by Lemma 3.1(ii).

(iii) Since x_0 is a regular point of g, $Dg(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ is surjective. So there exists $h \in \mathbb{R}^n$ such that

$$Dg_{x_0}(x_0)h = -\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad \Rightarrow \forall j, 1 \le j \le m, \quad \nabla g_j(x_0) \cdot h = -1 < 0 \quad \Rightarrow C \neq \varnothing.$$

The conclusion follows from part (i). □

Example 3.3.
Find the minimum distance from the point $(5,3) \in \mathbb{R}^2$ to the set

$$U \stackrel{\text{def}}{=} \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_2 \le 0,\ x_1 + 1 \le 0 \text{ and } x_1 + x_2 \le -2 \right\}$$

(see Figure 5.5). The objective function is the distance

$$\sqrt{(x_1 - 5)^2 + (x_2 - 3)^2}$$

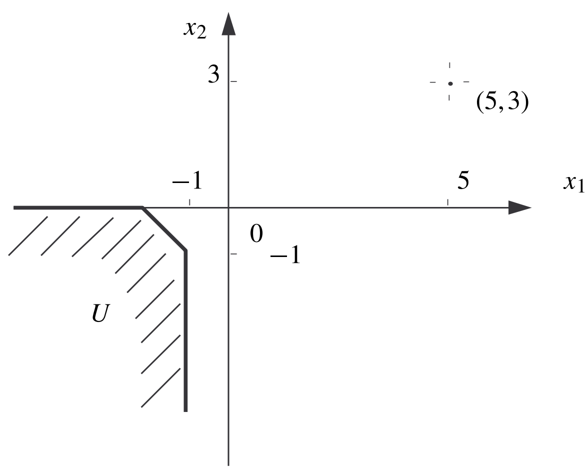


Figure 5.5. *The set of constraints U of Example* 3.3.

that has the same minimizers as the square of the distance

$$f(x) = \frac{1}{2}[(x_1-5)^2+(x_2-3)^2]$$

with respect to U. There exists a global minimum since f is continuous and has the growth property at infinity and U is closed. Let

$$g_1(x) = x_2, \quad g_2(x) = x_1+1, \quad g_3(x) = x_1+x_2+2$$

for which

$$U = \bigcap_{j=1}^{3} \{x : g_j(x) \le 0\}.$$

Since all functions are differentiable and even $C^{(1)}$, we are in the conditions of application of the Karush–Kuhn–Tucker theorem (Theorem 3.1): there exist $\lambda_j \ge 0$, $0 \le j \le 3$, not all zero, such that

$$\begin{aligned}
0 &= \lambda_1\, g_1(x) = \lambda_1\, x_2, && x_2 \le 0\\
0 &= \lambda_2\, g_2(x) = \lambda_2\,(x_1+1), && x_1+1 \le 0\\
0 &= \lambda_3\, g_3(x) = \lambda_3\,(x_1+x_2+2), && x_1+x_2+2 \le 0,
\end{aligned}$$

and

$$\lambda_0 \nabla f(x) + \sum_{j=1}^{3} \lambda_j \nabla g_j(x) = 0.$$

The last identity is equivalent to

$$\begin{aligned}
\lambda_0(x_1-5)+\lambda_2+\lambda_3 &= 0\\
\lambda_0(x_2-3)+\lambda_1+\lambda_3 &= 0.
\end{aligned}$$

We first test the value $\lambda_0 = 0$. In that case, we would have

$$0 \le \lambda_1 = \lambda_2 = -\lambda_3 \le 0 \quad \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0$$

that yields a contradiction. Set $\lambda_0 = 1$.

If $x_2 = 0$, then $x_1 \le -2$, which implies that

$$x_1 + 1 \le -1 \quad \Rightarrow \lambda_2 = 0.$$

We are left with

$$\lambda_1 \ge 0, \;\; \lambda_3(x_1+2) = 0, \;\; x_1+2 \le 0, \;\; x_1-5+\lambda_3 = 0, \;\; -3+\lambda_1+\lambda_3 = 0.$$

If $\lambda_3 = 0$, then $x_1 = 5$ and $\lambda_1 = 3$. But the condition $x_1 \le -2$ is violated. If $\lambda_3 > 0$, then $x_1 = -2$ and

$$\lambda_3 = 7, \;\; \lambda_1 = 3-\lambda_3 = 3-7 = -4 < 0$$

and this solution is not acceptable. So, by elimnation, $x_2 < 0$ and $\lambda_1 = 0$. We then get

$$x_2 < 0,\ \ \lambda_2(x_1+1) = 0,\ \ x_1+1 \le 0,\ \ \lambda_3(x_1+x_2+2) = 0,$$
$$x_1+x_2+2 \le 0,\ \ x_1-5+\lambda_2+\lambda_3 = 0,\ \ x_2-3+\lambda_3 = 0.$$

If $\lambda_3 = 0$, then $x_2 = 3$ that contradicts $x_2 < 0$ and the assumption that $\lambda_3 = 0$. Therefore, $\lambda_3 > 0$ and necessarily

$$x_1+x_2+2 = 0. \tag{3.27}$$

If $\lambda_2 = 0$, then

$$x_1-5+\lambda_3 = 0 \ \text{ and } \ x_2-3+\lambda_3 = 0 \quad \Rightarrow x_1+x_2-8+2\lambda_3 = 0$$

and combining this with (3.27), we get

$$-2-8+2\lambda_3 = 0 \quad \Rightarrow \lambda_3 = 5$$
$$\Rightarrow x_2 = 3-\lambda_3 = 3-5 = -2,\ \ x_1 = 5-\lambda_3 = 0.$$

Unfortunately,

$$x_1+1 = 1 > 0 \ \text{ while } \ x_1+1 \le 0.$$

So $\lambda_2 > 0$ and $x_1 = -1$. Then, since $\lambda_3 > 0$,

$$x_1+x_2+2 = 0 \quad \Rightarrow x_2 = -1.$$

It can be checked that $\lambda_3 = 4$ and $\lambda_2 = 2$ ($\lambda_1 = 0$) and that

$$x_1 = -1,\ \ x_2 = -1.$$

Thus the constraint $g_1(x) = x_2 \le 0$ is inactive.

Since the constraint functions are linear or affine and the objective function is quadratic and strictly convex, the necessary condition with $\lambda_0 = 1$ is also sufficient by Theorem 7.12(ii) of Chapter 4. □

Example 3.4.
Find the minimum of

$$f(x,y) = \frac{1}{2}[(x-2)^2+(y-1)^2]$$

with respect to the set

$$U = \left\{(x,y) \in \mathbb{R}^2 : x^2 \le y \le x\right\}$$

(see Figure 5.6).

Define the constraint functions

$$g_1(x,y) = x^2 - y \quad \text{and} \quad g_2(x,y) = y - x.$$

The set U is bounded. Indeed, for $(x,y) \in U$,

$$0 \le x^2 \le y \le x \quad \Rightarrow x^2 \le x \quad \Rightarrow 0 \le x(1-x)$$
$$0 \le x(1-x) \quad \Rightarrow 0 \le x \le 1 \quad \Rightarrow 0 \le x^2 \le y \le x \le 1 \quad \Rightarrow 0 \le y \le 1$$

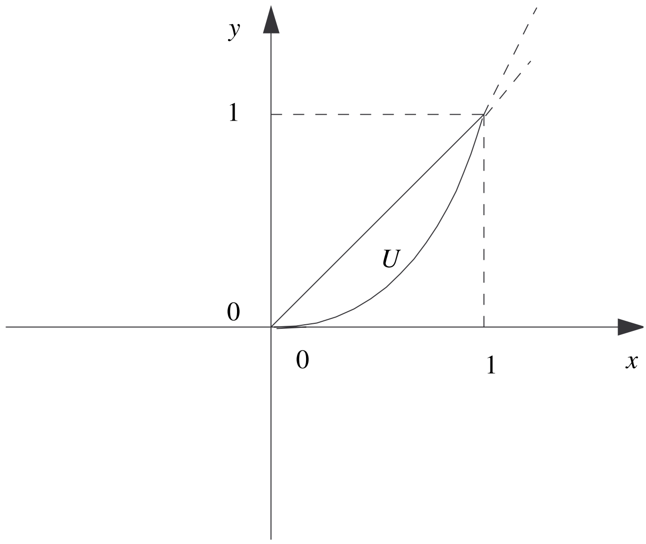


Figure 5.6. *The set of constraints U of Example* 3.4.

and U is bounded. It is also closed (and hence compact) since it is specified by two inequalities involving the continuous functions g_1 and g_2. Since f is continuous and U is compact, there exists a global minimizer in U. Since all the functions are $C^{(1)}$ on $\mathbb{R}^2$, we can apply the Karush–Kuhn–Tucker theorem (Theorem 3.1):

$$\begin{gathered}\exists\lambda_0 \geq 0,\ \ \lambda_1 \geq 0,\ \ \lambda_2 \geq 0,\ \ \lambda_0+\lambda_1+\lambda_2 > 0 \ \text{ such that}\\ \lambda_1(x^2-y)=0,\ \ \lambda_1 \geq 0,\ \ x^2-y \leq 0\\ \lambda_2(y-x)=0,\ \ \lambda_2 \geq 0,\ \ y-x \leq 0\\ \lambda_0(x-2)+2\lambda_1 x-\lambda_2=0\\ \lambda_0(y-1)-\lambda_1+\lambda_2=0.\end{gathered}$$

If $\lambda_0 = 0$, the first two equations yield $\lambda_1 = \lambda_2 > 0$ (since $\lambda_1+\lambda_2 > 0$) and $x = 1/2$. Moreover, since $\lambda_1 > 0$ and $\lambda_2 > 0$, we get

$$x^2 = y \ \text{ and } \ y = x \quad \Rightarrow \frac{1}{4} = \frac{1}{2},$$

which is a contradiction. Therefore, we can take $\lambda_0 = 1$.

If $\lambda_2 = 0$, then we get from the last two equations

$$x = \frac{2}{1+2\lambda_1} \ \text{ and } \ y = 1+\lambda_1.$$

But

$$x^2 \leq y \ \text{ and } \ y \leq x \quad \Rightarrow x^2 \leq x \quad \Rightarrow 0 \leq x \leq 1$$

and

$$x^2 \leq y \leq x \quad \Rightarrow 0 \leq y \leq 1.$$

From this,

$$1+\lambda_1 = y \le 1 \ \text{ and } \ \lambda_1 \ge 0 \quad \Rightarrow \lambda_1 = 0$$
$$\Rightarrow y = 1, \ \ x = 2 > 1$$

which is impossible since $x \le 1$. So we must take

$$\lambda_2 > 0 \quad \Rightarrow y = x.$$

If $\lambda_1 = 0$, we have

$$x = 2+\lambda_2, \ \ y = 1-\lambda_2.$$

But

$$\lambda_2 \ge 0 \quad \Rightarrow x = 2+\lambda_2 > 1$$

that is not possible. Therefore,

$$\lambda_1 > 0 \ \text{ and } \ y = x^2.$$

Then

$$x^2 = x \quad \Rightarrow x = 0 \text{ or } 1 \quad \Rightarrow (x,y) = (0,0) \text{ or } (1,1).$$

Substitute into the equations

$$x-2+2\lambda_1 x-\lambda_2 = 0, \ \ \lambda_1 > 0, \ \ \lambda_2 > 0,$$
$$y-1-\lambda_1+\lambda_2 = 0.$$

For $(0,0)$, we get

$$-2-\lambda_2 = 0, \ \ -1-\lambda_1+\lambda_2 = 0 \ \Longrightarrow \ \lambda_2 = -2 < 0.$$

For $(1,1)$, we get

$$1-2+2\lambda_1-\lambda_2 = 0$$
$$1-1-\lambda_1+\lambda_2 = 0 \quad \Rightarrow \lambda_1 = \lambda_2 = 1 > 0.$$

Finally, the solution is

$$\lambda_0 = \lambda_1 = \lambda_2 = 1, \ \ x = y = 1.$$

Since the continuous function f has a global minimizer in U over the compact set U, it will be $(1,1)$ which is the only point verifying the conditions of the theorem. □

Remark 3.2 (Slater's condition)**.**
Slater's condition could have been used in Examples 3.3 and 3.4. In Example 3.3, the constraint functions are linear and hence convex and the point $(x_1,x_2) = (-10,-10)$ is an interior point. In Example 3.4, the constraint function g_2 is linear. The constraint function g_1 is convex as the sum of a convex function and a linear function. It can also be checked that

$$x^2 \le y \le x \quad \Rightarrow 0 \le x \le 1.$$

By taking

$$x = \frac{1}{2} \quad \Rightarrow \frac{1}{4} \le y \le \frac{1}{2}$$

we can choose as interior point

$$y = \frac{1}{3} \quad \Rightarrow \frac{1}{2} = x > y = \frac{1}{3}, \ \ \frac{1}{4} = x^2 < y = \frac{1}{3}.$$

In both cases, Slater's condition could have been used. □

4 Simultaneous Equality and Inequality Constraints

In the last two sections, we have separately considered the cases of equality and inequality constraints. We now turn to the mixed case and generalize both the Lagrange and Karush–Kuhn–Tucker theorems.

Let $n,m,k \geq 1$ be three integers, $f : \mathbb{R}^n \to \mathbb{R}$ be the objective function, and

$$U \overset{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \begin{array}{l} g_i(x) \leq 0,\ 1 \leq i \leq m \\ q_j(x) = 0,\ 1 \leq j \leq k \end{array} \right\} \tag{4.1}$$

be the set of constraints specified by the functions $g_1,\dots,g_m$ and $q_1,\dots,q_k$ from $\mathbb{R}^n$ to $\mathbb{R}$. A possibility would be to rewrite the set U in terms of inequalities,

$$U \overset{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \begin{array}{l} g_i(x) \leq 0,\ 1 \leq i \leq m \\ q_j(x) \leq 0 \text{ and } -q_j(x) \leq 0,\ 1 \leq j \leq k \end{array} \right\},$$

and apply the Karush–Kuhn–Tucker theorem (Theorem 3.1). So, assume that f and the functions $g_1,\dots,g_m$ and $q_1,\dots,q_k$ are Fréchet differentiable at a point $x_0 \in U$ that is a local minimizer of f with respect to U. Then there exist

$$\lambda_0 \in \mathbb{R}, (\lambda_1,\dots,\lambda_m), \in \mathbb{R}^m, (\lambda_1^+,\dots,\lambda_k^+) \in \mathbb{R}^k, (\lambda_1^-,\dots,\lambda_k^-) \in \mathbb{R}^k \tag{4.2}$$

that are not all zero such that

$$\left| \begin{array}{l} \lambda_0 \geq 0, \quad \forall i,\ 1 \leq i \leq m, \quad \lambda_i \geq 0,\ g_i(x_0) \leq 0,\ \lambda_i g_i(x_0) = 0, \\ \forall j,\ 1 \leq j \leq k, \quad \left\{ \begin{array}{l} \lambda_j^+ \geq 0,\ q_j(x_0) \leq 0,\ \lambda_j^+ q_j(x_0) = 0, \\ \lambda_j^- \geq 0,\ -q_j(x_0) \leq 0,\ \lambda_j^+(-q_j(x_0)) = 0, \end{array} \right. \end{array} \right. \tag{4.3}$$

$$\lambda_0 \nabla f(x_0) + \sum_{i=1}^m \lambda_i \nabla g_i(x_0) + \sum_{j=1}^k (\lambda_j^+ - \lambda_j^-) \nabla q_j(x_0) = 0. \tag{4.4}$$

But, this necessary condition is verified at every point $x_0 \in U$ whether x_0 is a local minimizer or not. In other words, for each point $x_0 \in U$, we can find multipliers that are not all zero for which conditions (4.3) and (4.4) are verified: choose $\lambda_0 = 0, \lambda_i = 0, 1 \leq i \leq m, \lambda_j^+ = 1 = \lambda_j^-$, $1 \leq j \leq k$, to get

$$\forall i,\ 1 \leq i \leq m, \quad g_i(x_0) \leq 0, \quad \forall j,\ 1 \leq j \leq k, \quad q_j(x_0) = 0, \tag{4.5}$$

$$0 \nabla f(x_0) + \sum_{i=1}^m 0 \nabla g_i(x_0) + \sum_{j=1}^k (1-1) \nabla q_j(x_0) = 0$$

and the remaining condition (4.5) is verified for all points $x_0 \in U$. With that choice of multipliers, all conditions (4.3)–(4.4) are verified for all points of U. So, applying the Karush–Kuhn–Tucker theorem yields no information on the local minimizers.

In order to extend the previous theorems to the mixed case, we need to generalize Lemma 3.1 and the notion of constraints qualification of Definition 3.1 and use Theorem 7.3 of Chapter 4. The proof of the final theorem is analogous to the one of Theorem 3.1.

Notation 4.1.
In what follows, it will be convenient to deal with the k equality constraints globally by introducing the notation

$$x \mapsto g_{m+1}(x) \stackrel{\text{def}}{=} (q_1(x),\dots,q_k(x)) = \sum_{j=1}^{k} q_j(x)\,e_j^k : \mathbb{R}^n \to \mathbb{R}^k$$

$$\lambda_{m+1} \stackrel{\text{def}}{=} (\mu_1,\dots,\mu_k) = \sum_{j=1}^{k} \mu_j e_j^k \in \mathbb{R}^k.$$

The following identities will be used to make expressions more compact and streamline the proof of the lemmas and theorems.

By definition, with the above notation, the inner product gives

$$\lambda_{m+1} \cdot g_{m+1}(x) = \sum_{j=1}^{k} \mu_j\, q_j(x)$$

and, for all directions $h \in \mathbb{R}^n$,

$$\lambda_{m+1} \cdot (Dg_{m+1}(x)h) = \sum_{j=1}^{k} \mu_j\, \nabla q_j(x) \cdot h \quad \Rightarrow \; Dg_{m+1}(x)^\top \lambda_{m+1} = \sum_{j=1}^{k} \mu_j\, \nabla q_j(x).$$

We also have the equivalence,

$$Dg_{m+1}(x)h = 0 \quad \iff \quad \nabla q_j(x) \cdot h = 0, \quad \forall j,\, 1 \le j \le k. \tag{4.6}$$

□

Lemma 4.1. *Let $m \ge 1$, $n \ge 1$, and $k \ge 1$ be three integers, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \le i \le m$, $g_{m+1} = (q_1,\dots,q_k) : \mathbb{R}^n \to \mathbb{R}^k$,*

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : g_i(x) \le 0,\ 1 \le i \le m, \text{ and } g_{m+1}(x) = 0 \right\} \tag{4.7}$$

be the set of constraints, and $x_0 \in U$. Assume that $g_1,\dots,g_m$ are Fréchet differentiable at x_0 and that $q_1,\dots,q_k$ are of class $C^{(1)}$ on a neighborhood of x_0. Define the cone

$$C^0 \stackrel{\text{def}}{=} \left\{ h \in \mathbb{R}^n : \forall i \in I(x_0),\ \nabla g_i(x_0) \cdot h < 0,\ Dg_{m+1}(x_0)h = 0 \right\}, \tag{4.8}$$

where $I(x_0)$ is the set (3.2) *of indices of the active constraints $\{g_i\}$.*

(i) *Then*

$$T_U(x_0) \subset \left\{ h \in \mathbb{R}^n : \forall i \in I(x_0),\ \nabla g_i(x_0) \cdot h \le 0,\ Dg_{m+1}(x_0)h = 0 \right\}. \tag{4.9}$$

(ii) *If x_0 is a regular point of g_{m+1} and either $C^0 \ne \varnothing$ or $I(x_0) = \varnothing$, then*

$$\overline{C^0} = T_U(x_0) = \left\{ h \in \mathbb{R}^n : \forall i \in I(x_0),\ \nabla g_i(x_0) \cdot h \le 0,\ Dg_{m+1}(x_0)h = 0 \right\}. \tag{4.10}$$

(iii) *If x_0 is a singular point of g_{m+1} or if $C^0 = \varnothing$ and $I(x_0) \neq \varnothing$, then there exist $\{\lambda_i\}_{i\in I(x_0)}$, $\lambda_i \geq 0$, and $\{\mu_j\}_{j=1,\dots,k}$, $\mu_j \in \mathbb{R}$, not all zero, such that*

$$\sum_{i\in I(x_0)} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_j(x_0) = 0. \tag{4.11}$$

In general,

$$C^0 \subset T_U(x_0) \subset \left\{h \in \mathbb{R}^n : \forall i \in I(x_0),\ \nabla g_i(x_0)\cdot h \leq 0,\ Dg_{m+1}(x_0)h = 0\right\}.$$

Remark 4.1.
It is readily seen from the proof of part (i) that

$$\left\{h \in \mathbb{R}^n : \forall j \in I(x_0),\, d_H g_j(x_0;h) < 0,\ Dg_{m+1}(x_0)h = 0\right\} \subset T_U(x_0)$$
$$T_U(x_0) \subset \left\{h \in \mathbb{R}^n : \forall j \in I(x_0),\, d_H g_j(x_0;h) \leq 0,\ Dg_{m+1}(x_0)h = 0\right\}$$

with the weaker assumption that the g_j's be only Hadamard semidifferentiable at x_0. However, this does not seem to be sufficient to prove parts (ii) and (iii). □

Proof. We first prove that $T_U(x_0) = T_V(x_0)$ with

$$V \overset{\text{def}}{=} \left\{x \in \mathbb{R}^n : g_{m+1}(x) = g_{m+1}(x_0) \text{ and } \forall i \in I(x_0),\ g_i(x) \leq g_i(x_0)\right\},$$

and we prove the lemma with $T_V(x_0)$ in place of $T_U(x_0)$. Indeed, for all $i \notin I(x_0)$, $g_i(x_0) < 0$. Since the g_i's are Fréchet differentiable at x_0, they are continuous at x_0 and there exists a ball $B_r(x_0)$ of radius $r > 0$ such that

$$B_r(x_0) \subset \left\{x \in \mathbb{R}^n : g_i(x) < 0,\ 1 \leq i \leq m,\ i \notin I(x_0)\right\}.$$

By going from U to V, we only retain the active constraints. We first show that $U \cap B_r(x_0) = V \cap B_r(x_0)$, then $T_V(x_0) = T_U(x_0)$. Indeed, by definition, $U \subset V$ and hence $U \cap B_r(x_0) \subset V \cap B_r(x_0)$ and $T_U(x_0) \subset T_V(x_0)$. Conversely, for $x \in V \cap B_r(x_0)$ we have for all $j \in I(x_0)$, $g_j(x) \leq g_j(x_0) = 0$ and for all $j \notin I(x_0)$, $g_j(x) < 0$ since $x \in B_r(x_0)$. Therefore, x belongs to U and, a fortiori, to $U \cap B_r(x_0)$. By definition, for $h \in T_V(x_0)$,

$$\exists \{t_n > 0\},\, t_n \searrow 0, \text{ such that } \forall n, \exists x_n \in V,\ \lim_{n\to\infty} \frac{x_n - x_0}{t_n} = h.$$

Since $x_n \to x_0$, there exists N such that for all $n > N$, $x_n \in B_r(x_0)$ and hence $x_n \in B_r(x_0) \cap V = B_r(x_0) \cap U \subset U$. Therefore, always by definition, $h \in T_U(x_0)$ and $T_V(x_0) \subset T_U(x_0)$. So we can replace $T_U(x_0)$ by $T_V(x_0)$ to prove the lemma.

(i) When $I(x_0) = \varnothing$, $V = \{x \in \mathbb{R}^n : g_{m+1}(x) = g_{m+1}(x_0)\}$ and we have by Lemma 2.1 that $T_V(x_0) \subset \operatorname{Ker} Dg_{m+1}(x_0)$ and, on the other hand, the right-hand side of (4.9) is equal to $\operatorname{Ker} Dg_{m+1}(x_0)$. For $I(x_0) \neq \varnothing$ and $h \in T_V(x_0)$,

$$\exists \{t_n > 0\},\, t_n \searrow 0, \text{ such that } \forall n, \exists x_n \in V,\ \lim_{n\to\infty} \frac{x_n - x_0}{t_n} = h.$$

By definition of V,

$$\forall i \in I(x_0),\ g_i(x_n) - g_i(x_0) \leq 0$$

and

$$\nabla g_i(x_0)\cdot h = \frac{g_i(x_n)-g_i(x_0)}{t_n} - \left[\frac{g_i(x_n)-g_i(x_0)}{t_n} - \nabla g_i(x_0)\cdot h\right]$$
$$\leq -\left[\frac{g_i\left(x_0+t_n\frac{x_n-x_0}{t_n}\right)-g_i(x_0)}{t_n} - \nabla g_i(x_0)\cdot h\right].$$

Since g_i is Fréchet (Hadamard) at x_0 and $(x_n-x_0)/t_n \to h$, the right-hand side goes to zero and

$$\forall i \in I(x_0), \quad \nabla g_i(x_0)\cdot h \leq 0.$$

Now, after a certain rank, x_n belongs to the neighborhood V and

$$g_{m+1}(x_n) = g_{m+1}(x_0) = 0.$$

But the functions q_j are of class $C^{(1)}$ on a neighborhood of x_0. By Theorem 3.8 of Chapter 3, there exists $\alpha_j \in]0,1[$ such that

$$0 = q_j(x_n) - q_j(x_0) = \nabla q_j(x_0+\alpha_j(x_n-x_0))\cdot(x_n-x_0).$$

By dividing by $t_n > 0$ and by going to the limit,

$$0 = \nabla q_j(x_0)\cdot h \quad \Rightarrow Dg_{m+1}(x_0)h = 0.$$

This establishes the inclusion (4.9).

(ii) For $I(x_0)=\varnothing$, since x_0 is a regular point of g_{m+1}, we know by Lemma 2.2 that $T_V(x_0) = \operatorname{Ker} Dg_{m+1}(x_0)$ and, on the other hand, the right-hand side of (4.10) is equal to $\operatorname{Ker} Dg_{m+1}(x_0)$. This proves (4.10) in this case.

For $I(x_0)\neq\varnothing$ and $C^0\neq\varnothing$, we first prove that $C^0\subset T_V(x_0)$. For $h\in C^0$,

$$\forall i \in I(x_0),\ \nabla g_i(x_0)\cdot h < 0, \text{ and } Dg_{m+1}(x_0)h = 0$$

(and hence $h\neq 0$). Since the functions q_j are of class $C^{(1)}$ on a neighborhood of x_0 and x_0 is a regular point of g_{m+1}, we know by Theorem 2.1 that

$$\exists \beta>0,\ \exists x :]-\beta,\beta[\to\mathbb{R}^n,\ x(0)=x_0,\ x'(0)=h,$$
$$g_{m+1}(x(t)) = g_{m+1}(x_0) + t\,Dg_{m+1}(x_0)h = 0,$$

where x is of class $C^{(1)}$. To prove that, as t goes to zero, $x(t)\in V$, it remains to prove that, as t goes to zero, $g_i(x(t))\leq g_i(x_0)$ for all $i\in I(x_0)$. Since $h\in C^0$,

$$\alpha \overset{\text{def}}{=} \max_{i\in I(x_0)} \nabla g_i(x)\cdot h < 0.$$

Consider the quotient

$$\frac{g_i(x(t))-g_i(x_0)}{t} = \nabla g_i(x_0)\cdot h + \left[\frac{g_i(x(t))-g_i(x_0)}{t} - \nabla g_i(x_0)\cdot h\right]$$
$$\leq \max_{i\in I(x_0)} \nabla g_i(x)\cdot h + \max_{i\in I(x_0)}\left|\frac{g_i(x(t))-g_i(x_0)}{t} - \nabla g_i(x_0)\cdot h\right|.$$

Since the g_i's are Fréchet (Hadamard) differentiable at x_0 and $(x(t)-x_0)/t \to h$, there exists β_1, $0 < \beta_1 \le \beta$ such that

$$\forall t, 0 < t < \beta_1, \quad \max_{i \in I(x_0)} \left| \frac{g_i\left(x_0 + \frac{x(t)-x_0}{t}\right) - g_i(x_0)}{t} - \nabla g_i(x_0) \cdot h \right| < -\frac{\alpha}{2}.$$

By definition of α, for all t, $0 < t < \beta_1$,

$$\forall i \in I(x_0), \quad \frac{g_i(x(t)) - g_i(x_0)}{t} < \alpha - \frac{\alpha}{2} = \frac{\alpha}{2} < 0$$

and, since $t > 0$, for all $i \in I(x_0)$, $g_i(x(t)) - g_i(x_0) < 0$, and $h \in T_V(x_0)$. This proves that $C^0 \subset T_V(x_0)$.

Now for each $i \in I(x_0)$, define

$$\begin{aligned} C_i^0 &\stackrel{\text{def}}{=} \left\{h \in \mathbb{R}^n : \nabla g_i(x_0) \cdot h < 0 \text{ and } Dg_{m+1}(x_0)h = 0\right\} \\ &= \{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_i(x_0) \cdot h < 0\}. \end{aligned}$$

Since, by assumption, $C^0 \neq \varnothing$, then

$$\exists h^* \in \mathbb{R}^n \text{ such that } \forall i \in I(x_0), \nabla g_i(x_0) \cdot h^* < 0 \text{ and } Dg_{m+1}(x_0)h^* = 0.$$

Since, for each $i \in I(x_0)$, $C_i^0 \neq \varnothing$, we then get $\nabla g_i(x_0) \neq 0$ and

$$\overline{C_i^0} = \left\{h \in \mathbb{R}^n : \nabla g_i(x_0) \cdot h \le 0 \text{ and } Dg_{m+1}(x_0)h = 0\right\}.$$

We now prove that

$$\bigcap_{i \in I(x_0)} \overline{C_i^0} \subset \overline{\bigcap_{i \in I(x_0)} C_i^0} = \overline{C^0}.$$

For all $h \in \cap_{i \in I(x_0)} \overline{C_i^0}$, we have, by definition,

$$\forall i \in I(x_0), \nabla g_i(x_0) \cdot h \le 0 \text{ and } Dg_{m+1}(x_0)h = 0.$$

For each $n \ge 1$, define

$$h_n \stackrel{\text{def}}{=} h + \frac{1}{n} h^* \to h$$

that yields

$$\begin{gathered} Dg_{m+1}(x_0)h_n = Dg_{m+1}(x_0)h + \frac{1}{n} Dg_{m+1}(x_0)h^* = 0 \\ \nabla g_i(x_0) \cdot h_n = \nabla g_i(x_0) \cdot h + \frac{1}{n} \nabla g_i(x_0) \cdot h^* < \nabla g_i(x_0) \cdot h \le 0 \\ \Rightarrow h_n \in \bigcap_{i \in I(x_0)} C_i^0 \quad \Rightarrow h_n \to h \in \overline{\bigcap_{i \in I(x_0)} C_i^0}. \end{gathered}$$

Finallly, from the first part, since $C^0 \subset T_V(x_0)$ and $T_V(x_0)$ is closed,

$$\bigcap_{i\in I(x_0)} \overline{C_i^0} \subset \overline{\bigcap_{i\in I(x_0)} C_i^0} \subset T_U(x_0) \subset \bigcap_{i\in I(x_0)} \overline{C_i^0} \quad \Rightarrow T_U(x_0) = \bigcap_{i\in I(x_0)} \overline{C_i^0}.$$

This completes the argument.

(iii) If x_0 is a singular point of g_{m+1}, then $Dg_{m+1}(x_0) : \mathbb{R}^n \to \mathbb{R}^k$ is not surjective and there exists $\lambda_{m+1} \in \mathbb{R}^k$, $\lambda_{m+1} \neq 0$, such that $Dg_{m+1}(x_0)^\top \lambda_{m+1} = 0$; that is, there exist $\mu_1, \dots, \mu_k$, $\mu_j \in \mathbb{R}$, not all zero, such that

$$\sum_{j=1}^k \mu_j \nabla q_j(x_0) = 0.$$

The result follows by choosing $\lambda_i = 0$ for all $i \in I(x_0)$.

If $x_0 \in U$ is a regular point of g_{m+1}, then $I(x_0) \neq \varnothing$ and $C^0 = \varnothing$. Assume that there exists an index $i \in I(x_0)$ such that $\nabla g_i(x_0) \in [\operatorname{Ker} Dg_{m+1}(x_0)]^\perp$ or, equivalently, $\nabla g_i(x_0) \in \operatorname{Im}[Dg_{m+1}(x_0)]^\top$. Then, there exists $\lambda_{m+1} = -(\mu_1, \dots, \mu_k) \in \mathbb{R}^k$ such that

$$\begin{aligned}\nabla g_i(x_0) = [Dg_{m+1}(x_0)]^\top \lambda_{m+1} = -\sum_{j=1}^k \mu_j \nabla q_j(x_0)\\ \Rightarrow \nabla g_i(x_0) + \sum_{j=1}^k \mu_j \nabla q_j(x_0) = 0.\end{aligned}$$

The lemma is then true by choosing $\lambda_i = 1$ and $\lambda_{i'} = 0$, $i' \neq i$.

The last case is when x_0 is a regular point of g_{m+1}, $I(x_0) \neq \varnothing$, for all $i \in I(x_0)$, $\nabla g_i(x_0) \notin \operatorname{Ker} Dg_{m+1}(x_0)^\perp$, and $C^0 = \varnothing$. So, we are back to the proof of Lemma 3.1(iii) with the cones C_i^0 satisfying

$$\forall i \in I(x_0), C_i^0 \neq \varnothing \quad \text{and} \quad C^0 = \bigcap_{i\in I(x_0)} C_i^0 = \varnothing \tag{4.12}$$

instead of the cones $C_i = \{h \in \mathbb{R}^n : \nabla g_i(x_0) \cdot h < 0\}$, where the linear subspace Ker $Dg_{m+1}(x_0)$ now plays the role of $\mathbb{R}^n$. Finally, since $\operatorname{Ker} Dg_{m+1}(x_0)$ is a linear subspace, we readily get that if $\nabla g_i(x_0) \notin \operatorname{Ker} Dg_{m+1}(x_0)^\perp$, then $C_i^0 \neq \varnothing$, $\nabla g_i(x_0) \neq 0$, and

$$\overline{C_i^0} = \{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_i(x_0) \cdot h \leq 0\}.$$

Relabel the sets C_i^0, $i \in I(x_0)$, from 1 to r in the increasing order of the indices, where r is the total number of elements in $I(x_0)$. If $\cap_{i=2}^r C_i^0 = \varnothing$, go to $\cap_{i=3}^r C_i^0$. If this last intersection is empty, go to $\cap_{i=4}^r C_i^0$. Let s be the first index such that

$$\bigcap_{i=s}^r C_i^0 = \varnothing \text{ and } K^0 \stackrel{\text{def}}{=} \bigcap_{i=s+1}^r C_i^0 \neq \varnothing \quad \Rightarrow K^0 \cap C_s^0 = \varnothing.$$

This index s exists since $C_i^0 \neq \varnothing$, $1 \leq i \leq r$, and for $i = r$,

$$\bigcap_{i=r}^r C_i^0 = C_r^0 \neq \varnothing.$$

Then, we get $K^0 \subset \operatorname{Ker} Dg_{m+1}(x_0)\backslash C_s^0$ or, explicitly,

$$\begin{aligned}&\{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_i(x_0)\cdot h < 0, s+1 \le i \le r\}\\ &\subset \{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_s(x_0)\cdot h \ge 0\}.\end{aligned} \tag{4.13}$$

Since x_0 is a regular point of g_{m+1} and

$$K^0 = \bigcap_{i=s+1}^{r} C_i^0 \neq \varnothing,$$

we know from the proof of part (ii) that

$$\overline{K^0} = \left\{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_j(x_0)\cdot h \le 0, s+1 \le j \le r\right\}$$

and from Theorem 5.4(ii) of Chapter 4 that $(\overline{K^0})^* = (K^0)^*$. Then, by taking the dual of the inclusion (4.13), we get

$$\{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_s(x_0)\cdot h \ge 0\}^* \subset (K^0)^* = (\overline{K^0})^*$$

or still

$$\begin{aligned}&\{h \in \operatorname{Ker} Dg_{m+1}(x_0) : \nabla g_s(x_0)\cdot h \ge 0\}^*\\ &\subset \left\{h \in \operatorname{Ker} Dg_{m+1}(x_0) : -\nabla g_j(x_0)\cdot h \ge 0, s+1 \le j \le r\right\}^*.\end{aligned}$$

But the dual cones are known from Theorem 7.3 of Chapter 4, and hence

$$\begin{aligned}&\left\{\lambda_s \nabla g_s(x_0) + [Dg_{m+1}(x_0)]^\top \lambda_{m+1} : \forall \lambda_s \ge 0 \text{ and } \forall \lambda_{m+1} \in \mathbb{R}^k\right\}\\ &\subset \left\{\sum_{j=s+1}^{r} \lambda_j \left(-\nabla g_j(x_0)\right) + [Dg_{m+1}(x_0)]^\top \lambda'_{m+1} : \begin{array}{l}\forall \lambda_j \ge 0, s+1 \le j \le r\\ \text{and } \forall \lambda'_{m+1} \in \mathbb{R}^k\end{array}\right\}.\end{aligned}$$

Therefore, for $\lambda_s = 1$, there exist $\lambda_{s+1} \ge 0, \dots, \lambda_r \ge 0$ and $(\mu_1, \dots, \mu_k) = \lambda_{m+1} - \lambda'_{m+1} \in \mathbb{R}^k$ such that

$$\nabla g_s(x_0) + \sum_{i=s+1}^{r} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_j(x_0) = 0.$$

By choosing $\lambda_1 = \cdots = \lambda_{s-1} = 0$, we finally get identity (4.11) where $\lambda_s = 1$, that concludes the proof of part (iii). □

We shall see later that it is only when property (4.10) of Lemma 4.1 is verified that the extension of the theorem of Karush–Kuhn–Tucker gives information on the local minimizers. As in the case of inequalities, we introduce the following terminology.

Definition 4.1 (Qualification of the constraints)**.**
Let $m \ge 1$, $n \ge 1$, and $k \ge 1$ be three integers, U the set of constraints in (4.7) specified by the real-valued functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \le i \le m$, and the vector-valued function $g_{m+1} = (q_1, \dots, q_k) : \mathbb{R}^n \to \mathbb{R}^k$. Assume that the functions $g_1, \dots, g_m$ are Fréchet differentiable at

a point $x_0 \in U$ and that $q_1, \dots, q_k$ are of class $C^{(1)}$ on a neighborhood of x_0. We say that we have *qualification of the constraints* at x_0 if

$$T_U(x_0) = \left\{ h \in \mathbb{R}^n : \forall i \in I(x_0), \quad \nabla g_i(x_0) \cdot h \le 0,\ Dg_{m+1}(x_0)h = 0 \right\}, \tag{4.14}$$

where $I(x_0)$ is the set (3.2) of indices of the active constraints g_i at x_0. □

As we have seen in Lemma 4.1, we have qualification of the constraints when x^0 is a regular point of g_{m+1} and $C^0 \neq \varnothing$. The question is open in the other case.

Theorem 4.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be the objective function and let*

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \begin{matrix} g_i(x) \le 0,\ 1 \le i \le m \\ q_j(x) = 0,\ 1 \le j \le k \end{matrix} \right\} \tag{4.15}$$

be the set of constraints specified by the functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \le i \le m$, and $q_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le k$.

Assume that

(a) *f has a local minimum at $x_0 \in U$ with respect to U,*

(b) *f is Fréchet differentiable at x_0,*

(c) *the functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \le i \le m$, are Fréchet differentiable at x_0 and that the functions $q_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le k$, are of class $C^{(1)}$ on a neighborhood of x_0.*

Then, we have the following properties.

(i) *If we have qualification of the constraints at x_0, then there exist $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ and $(\mu_1, \dots, \mu_k) \in \mathbb{R}^k$ such that*

$$\lambda_i \ge 0,\ g_i(x_0) \le 0,\ \lambda_i g_i(x_0) = 0, \quad 1 \le i \le m, \tag{4.16}$$

$$\begin{cases} q_j(x_0) = 0, \quad 1 \le j \le k, \\ \nabla f(x_0) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_j(x_0) = 0. \end{cases} \tag{4.17}$$

(ii) *If we do not have qualification of the constraints at x_0, then there exist $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ and $(\mu_1, \dots, \mu_k) \in \mathbb{R}^k$, not all zero, such that*

$$\lambda_i \ge 0,\ g_i(x_0) \le 0,\ \lambda_i g_i(x_0) = 0, \quad 1 \le i \le m, \tag{4.18}$$

$$\begin{cases} q_j(x_0) = 0, \quad 1 \le j \le k, \\ \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_j(x_0) = 0. \end{cases} \tag{4.19}$$

(iii) *There exist $\lambda_0 \in \mathbb{R}$, $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, and $(\mu_1, \dots, \mu_k) \in \mathbb{R}^k$, not all zero, such that*

$$\lambda_0 \ge 0, \quad \lambda_i \ge 0,\ g_i(x_0) \le 0,\ \lambda_i g_i(x_0) = 0, \quad 1 \le i \le m, \tag{4.20}$$

$$\begin{cases} q_j(x_0) = 0, \quad 1 \le j \le k, \\ \lambda_0 \nabla f(x_0) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_j(x_0) = 0. \end{cases} \tag{4.21}$$

(iv) *By introducing the* Lagrangian

$$L(x,\lambda,\mu) \stackrel{\text{def}}{=} \lambda_0 f(x) + \sum_{i=1}^{m} \lambda_j g_i(x) + \sum_{j=1}^{k} \lambda_j q_j(x), \tag{4.22}$$

there exist $\lambda_0 \in \mathbb{R}$, $(\lambda_1,\dots,\lambda_m) \in \mathbb{R}^m$, *and* $(\mu_1,\dots,\mu_k) \in \mathbb{R}^k$, *not all zero, such that*

$$\begin{cases} \lambda_0 \geq 0, \\ \lambda_j \geq 0, \ \dfrac{\partial L}{\partial \lambda_j}(x_0,\lambda,\mu) \leq 0, \ \lambda_j \dfrac{\partial L}{\partial \lambda_j}(x_0,\lambda,\mu) = 0, \quad 1 \leq j \leq m, \end{cases} \tag{4.23}$$

$$\frac{\partial L}{\partial \mu_j}(x_0,\lambda,\mu) = 0, \quad 1 \leq j \leq k, \quad \frac{\partial L}{\partial x_i}(x_0,\lambda,\mu) = 0, \quad 1 \leq i \leq n. \tag{4.24}$$

Proof. (i) As $x_0 \in U$ is a local minimizer of f with respect to U,

$$\nabla f(x_0) \in T_U(x_0)^*.$$

Since we have qualification of the constraints at x_0,

$$T_U(x_0) = \left\{ h \in \mathbb{R}^n : \begin{array}{l} \nabla g_i(x_0) \cdot h \leq 0, \quad i \in I(x_0) \\ \text{and } \nabla q_j(x_0) \cdot h = 0, \quad 1 \leq j \leq k \end{array} \right\}.$$

By Theorem 7.3 of Chapter 4, the dual of $T_U(x_0)$ is given by

$$T_U(x_0)^* = \left\{ \sum_{i \in I(x_0)} -\lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} -\mu_j \nabla q_i(x_0) : \begin{array}{l} \forall \lambda_i \geq 0, 1 \leq i \leq m \\ \forall \mu_j \in \mathbb{R}, 1 \leq j \leq k \end{array} \right\}.$$

So, there exist $\{\lambda_i\}_{i \in I(x_0)}$, $\lambda_i \geq 0$, and $(\mu_1,\dots,\mu_k) \in \mathbb{R}^k$ such that

$$\nabla f(x_0) + \sum_{i \in I(x_0)} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_i(x_0) = 0.$$

Since $g_i(x_0) = 0$ for all $i \in I(x_0)$, we have $\lambda_i g_i(x_0) = 0$. For $i \notin I(x_0)$, choose $\lambda_i = 0$. Thence, properties (4.16) and (4.17).

(ii) When we do not have qualification of the constraints, then by Lemma 4.1(ii), $C_0 = \varnothing$ and $I(x_0) \neq \varnothing$ or the point x_0 is not a regular point of g_{m+1} and by Lemma 4.1(iii), there exist $\{\lambda_i\}_{i \in I(x_0)}$, $\lambda_i \geq 0$, and $(\mu_1,\dots,\mu_k) \in \mathbb{R}^k$, not all zero, such that

$$\sum_{i \in I(x_0)} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{k} \mu_j \nabla q_i(x_0) = 0. \tag{4.25}$$

Since $g_j(x_0) = 0$ for all $j \in I(x_0)$, we have $\lambda_j g_j(x_0) = 0$. For $j \notin I(x_0)$, choose $\lambda_j = 0$. Thence, properties (4.18) and (4.19).

(iii) If we have qualification of the constraints, we know from part (i) that we can choose $\lambda_0 = 1$. If we do not have qualification of the constraints, we use part (ii) with $\lambda_0 = 0$, but, in that case, we know that there exists at least one λ_i, $1 \leq i \leq m$, or a μ_j, $1 \leq j \leq k$, that is non zero. This completes the argument.

(iv) This is a straightforward rewriting of part (iii). $\square$

Again, we can always choose $\lambda_0 = 1$ when we have qualification of the constraints (4.14). As in Theorem 3.2, there are sufficient conditions to check that condition.

Theorem 4.2. *Let $m \geq 1$, $n \geq 1$, and $k \geq 1$ be three integers, let*

$$U \overset{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \begin{array}{l} g_i(x) \leq 0,\ 1 \leq i \leq m \\ q_j(x) = 0,\ 1 \leq j \leq k \end{array} \right\} \tag{4.26}$$

be the set of constraints specified by the functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq m$, and $g_{m+1} = (q_1, \dots, q_k) : \mathbb{R}^n \to \mathbb{R}^k$. Assume that the functions $g_1, \dots, g_m$ are Fréchet differentiable at a point $x_0 \in U$ and that the functions $q_1, \dots, q_k$ are of class $C^{(1)}$ on a neighborhood of x_0.

Condition (4.14) *of qualification of the constraints in x_0 is verified if any one of the following conditions is verified.*

(i) *The point x_0 is a regular point of g_{m+1} and there exists $h^* \in \mathbb{R}^n$ such that*

$$\forall i,\ 1 \leq i \leq m,\ \nabla g_i(x_0) \cdot h^* < 0,\ \text{and}\ Dg_{m+1}(x_0)h^* = 0. \tag{4.27}$$

(ii) *The point x_0 is a regular point of g_{m+1}, the functions g_i, $i \in I(x_0)$, are convex, and there exists $x^* \in \mathbb{R}^n$ such that*

$$\forall i,\ 1 \leq i \leq m,\ g_i(x^*) < 0,\ \text{and}\ Dg_{m+1}(x_0)(x^* - x_0) = 0. \tag{4.28}$$

(iii) *The point x_0 is a regular point of the map*

$$\begin{aligned} g(x) &\overset{\text{def}}{=} (g_1(x), \dots, g_m(x), q_1(x), \dots, q_k(x)) \\ &= \sum_{i=1}^{m} g_i(x) e_i^{m+k} + \sum_{j=1}^{k} q_j(x) e_{m+j}^{m+k} \end{aligned} \tag{4.29}$$

where $\{e_\ell^{m+k}\}_{\ell=1}^{m+k}$ is the canonical orthonormal basis of $\mathbb{R}^{m+k}$.

Proof. (i) Since the property is verified for all i, $1 \leq i \leq m$, it is verified for all $I(x_0)$. Therefore, by Lemma 4.1(ii), $C^0 \neq \varnothing$.

(ii) For all $i \in I(x_0)$, $g_i(x_0) = 0$. By convexity of g_i,

$$g_i(x^*) \geq g_i(x_0) + \nabla g_i(x_0) \cdot (x^* - x_0)$$

and

$$\forall i \in I(x_0), \quad 0 > g_i(x^*) \geq \nabla g_i(x_0) \cdot (x^* - x_0).$$

Since, by assumption, we also have $Dg_{m+1}(x_0)(x^* - x_0) = 0$, the direction $x^* - x_0$ belongs to C^0 and hence $C^0 \neq \varnothing$. By Lemma 4.1(ii), the constraints are qualified.

(iii) Since x_0 is a regular point of g, $Dg(x_0) : \mathbb{R}^n \to \mathbb{R}^{m+k}$ is surjective. So there exists $h \in \mathbb{R}^n$ such that

$$\begin{aligned} &Dg(x_0)h = ((-1, \dots, -1), (0, \dots, 0)) \\ \Rightarrow &\left| \begin{array}{ll} \forall i,\ 1 \leq i \leq m, & \nabla g_i(x_0) \cdot h = -1 < 0 \\ \forall 1 \leq j \leq k, & \nabla q_j(x_0) \cdot h = 0 \end{array} \right. \end{aligned}$$

and hence $C^0 \neq \varnothing$. Again the conclusion follows from part (i). □

Example 4.1.
Find the minimum of

$$f(x_1,x_2,x_3) \overset{\text{def}}{=} x_1^2+2x_2^2+4x_3^2$$

with respect to

$$U \overset{\text{def}}{=} \left\{(x_1,x_2,x_3) : \begin{matrix} x_i \geq 0, \;\; 1 \leq i \leq 3 \\ x_1+x_2+x_3 = 1 \end{matrix}\right\}.$$

There exists a minimizer since f is continuous and U is compact, $0 \leq x_i \leq 1$, $1 \leq i \leq 3$. Introduce the constraint functions

$$g_i(x) = -x_i, \;\; 1 \leq i \leq 3$$
$$g_4(x) = x_1+x_2+x_3-1.$$

All functions $f, g_1, \dots, g_4$ are $C^{(1)}$ in $\mathbb{R}^3$ and the points of U_4 are regular points of g_4 since

$$\nabla g_4(x) = (1,1,1) \neq (0,0,0)$$

and $h \mapsto \nabla g_4(x) \cdot h$ is surjective for all $x \in \mathbb{R}^3$. On the other hand, it is easy to check that the set

$$\{h \in \mathbb{R}^3 : \nabla g_i(x) \cdot h < 0, \;\; i = 1,2,3, \;\; \text{and} \;\; \nabla g_4(x) \cdot h = 0\}$$

is empty for all $x \in \mathbb{R}^3$ since

$$\nabla g_i(x) \cdot h = -h_i, \;\; 1 \leq i \leq 3, \quad \text{and} \quad \nabla g_4(x) \cdot h = h_1+h_2+h_3$$

would imply

$$h_i > 0, \;\; i = 1,2,3, \;\; \text{and} \;\; h_1+h_2+h_3 = 0$$

that is impossible. So, condition (ii) of Theorem 4.2 does not appply. In addition, there are no regular points in dimension 3 for a vector function with 4 components and only 3 variables. Condition (i) of Theorem 4.2 does not apply either.

So, we have no a priori information on λ_0. By Theorem 4.1(iii),

$$\exists(\lambda_0,\lambda_1,\lambda_2,\lambda_3,\lambda_4) \neq (0,0,0,0,0)$$

such that $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i = 1,2,3$, $\lambda_4 \in \mathbb{R}$,

$$\lambda_i x_i = 0 \;\; \text{and} \;\; x_i \geq 0, \;\; 1 \leq i \leq 3 \tag{4.30}$$
$$x_1+x_2+x_3 = 1 \tag{4.31}$$
$$\lambda_0 \nabla f(x) + \sum_{i=1}^{4} \lambda_i \nabla g_i(x) = 0.$$

The last equation reduces to

$$\lambda_0(2x_1,4x_2,8x_3) - (\lambda_1,\lambda_2,\lambda_3) + \lambda_4(1,1,1) = 0. \tag{4.32}$$

If $\lambda_0 = 0$, then from (4.32),

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$$

and from (4.30),

$$0 = \sum_{i=1}^{3} \lambda_i x_i = \lambda_4(x_1 + x_2 + x_3) = \lambda_4 \quad \Rightarrow \lambda_4 = 0.$$

So all the multipliers are zero, which contradicts the conclusion of the theorem. So, we can choose $\lambda_0 = 1$. Inner product equation (4.32) and $x = (x_1, x_2, x_3)$

$$2x_1^2 + 4x_2^2 + 8x_3^2 - \lambda_1 x_1 - \lambda_2 x_2 - \lambda_3 x_3 + \lambda_4(x_1 + x_2 + x_3) = 0.$$

But, from (4.30) and (4.31),

$$2(x_1^2 + 2x_2^2 + 4x_3^2) + \lambda_4 = 0.$$

We draw two conclusions:

$$f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 4x_3^2 = -\lambda_4/2,$$

and since

$$x_1 + x_2 + x_3 = 1, \;\; x_1^2 + 2x_2^2 + 4x_3^2 > 0 \implies \lambda_4 < 0.$$

Now, from (4.32), if $x_i = 0$ for some i, then $\lambda_i = \lambda_4 < 0$; this contradicts the conclusion of the theorem that says that $\lambda_i \geq 0$. So

$$\forall i = 1, 2, 3, \;\; x_i > 0 \;\text{ and }\; \lambda_i = 0 \;\; (\text{from (4.30)}).$$

Therefore, from (4.32),

$$x_1 = -\lambda_4/2, \;\; x_2 = -\lambda_4/4, \;\; x_3 = -\lambda_4/8.$$

Finally, from (4.31),

$$1 = x_1 + x_2 + x_3 = -\frac{\lambda_4}{8}[4 + 2 + 1] = -\frac{7}{8}\lambda_4 \quad \Rightarrow \lambda_4 = -\frac{8}{7}$$

and

$$x_1 = \frac{4}{7}, \;\; x_2 = \frac{2}{7}, \;\; x_3 = \frac{1}{7}, \;\; -\frac{\lambda_4}{2} = \frac{4}{7}.$$

So, the minimum of the function f with respect to U is $4/7$ and the minimizer is $x = (4/7, 2/7, 1/7)$.

Since the constraint functions are linear or affine and the objective function is quadratic and strictly convex, the necessary condition with $\lambda_0 = 1$ is also sufficient by Theorem 7.12(ii) of Chapter 4. □

Example 4.2.
To show that $n \geq 1$,

$$\left(\prod_{i=1}^{n} x_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n} x_i, \;\; \forall x_i \geq 0, \;\; i = 1, \ldots, n. \tag{4.33}$$

Choose as objective function

$$f(x) = \prod_{i=1}^{n} x_i \tag{4.34}$$

and constraint functions

$$g_i(x) \stackrel{\text{def}}{=} -x_i,\ 1 \le i \le n, \text{ and } g_{n+1}(x) \stackrel{\text{def}}{=} \sum_{i+1}^{n} x_i - a \tag{4.35}$$

for a given $a > 0$.

The function f is continuous and the set

$$U = \left\{ x \in \mathbb{R}^n : g_i(x) \le 0,\ 1 \le i \le n, \text{ and } g_{n+1}(x) = 0 \right\}$$

is compact ($0 \le x_i \le a$). So there exists a maximizer of f in U that is also a minimzer of $-f$ in U. All the functions $f, g_1, \dots, g_n, g_{n+1}$ are $C^{(1)}$ in $\mathbb{R}^n$. As in Example 4.1, it is easy to check that

$$\nabla g_{n+1}(x) = (1, 1, \dots, 1)$$

and that all the points of U_{n+1} are regular points of g_{n+1}. Apply Theorem 4.1(iii) directly:

$$\exists (\lambda_0, \lambda_1, \dots, \lambda_{n+1}) \ne (0, 0, \dots, 0) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \tag{4.36}$$

such that $\lambda_0 \ge 0$, $\lambda_i \ge 0$, $1 \le i \le n$, $\lambda_{n+1} \in \mathbb{R}$,

$$\lambda_i x_i = 0,\ \ x_i \ge 0,\ \ 1 \le i \le n, \tag{4.37}$$

$$\sum_{i=1}^{n} x_i = a \tag{4.38}$$

and

$$-\lambda_0 \nabla f(x) + \sum_{i=1}^{n+1} \lambda_i \nabla g_i(x) = 0$$

or

$$-\lambda_0 \prod_{\substack{j=1 \\ j \ne i}}^{n} x_j - \lambda_i + \lambda_{n+1} = 0,\ \ 1 \le i \le n. \tag{4.39}$$

If $\lambda_0 = 0$,

$$\lambda_i = \lambda_{n+1},\ \ 1 \le i \le n$$

and from (4.37),

$$0 = \sum_{i=1}^{n} \lambda_i x_i = \lambda_{n+1} \sum_{i=1}^{n} x_i = \lambda_{n+1} a \quad \Rightarrow \lambda_{n+1} = 0.$$

So all the multipliers are zero and this contradicts the conclusions of the theorem. So we can choose $\lambda_0 = 1$. Multiply (4.39) by x_i and sum up with respect to i,

$$-n \prod_{j=1}^{n} x_j - \sum_{i=1}^{n} \lambda_i x_i + \lambda_{n+1} \sum_{i=1}^{n} x_i = 0,$$

and from (4.37) and (4.38),

$$-n\prod_{j=1}^{n} x_j + \lambda_{n+1}a = 0 \implies \prod_{j=1}^{n} x_j = \frac{a}{n}\lambda_{n+1} \tag{4.40}$$

$$\Rightarrow \lambda_{n+1} = \frac{n}{a}\prod_{j=1}^{n} x_j \geq 0. \tag{4.41}$$

By multiplying (4.39) by x_i and using identity (4.41), we have

$$-\prod_{j=1}^{n} x_j + \lambda_i x_i + \left[\frac{n}{a}\prod_{j=1}^{n} x_j\right] x_i = 0 \quad \Rightarrow \prod_{j=1}^{n} x_j \left[\frac{n}{a}x_i - 1\right] = 0, \ \forall i = 1,\ldots,n.$$

So, two cases must be considered

$$f(x) = \prod_{j=1}^{n} x_j \neq 0 \quad \implies x_i = \frac{a}{n}, \ \ 1 \leq i \leq n, \quad \implies f(x) = \left(\frac{a}{n}\right)^n > 0$$

and

$$f(x) = \prod_{j=1}^{n} x_j = 0 \ \ \text{(at least one of the } x_j \text{ is zero).}$$

Since we seek a global maximum f, the last case can be dropped. There is only one possibility left,

$$x_i = \frac{a}{n} > 0, \ \ 1 \leq i \leq n, \ \ \text{and} \ \ f(x) = \left(\frac{a}{n}\right)^n > 0,$$

which is the desired global maximum of f.

Now go back to our initial problem. If

$$\sum_{i=1}^{n} x_i = 0 \ \text{ with } \ \forall i, \ \ x_i \geq 0,$$

then, for all $i, x_i = 0$, and

$$0 = \prod_{i=1}^{n} x_i = \left[\frac{1}{n}\sum_{i=1}^{n} x_i\right]^n \quad \Rightarrow \left[\prod_{i=1}^{n} x_i\right]^{\frac{1}{n}} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

If

$$\sum_{i=1}^{n} x_i = a > 0, \ \ \forall i, \ \ x_i \geq 0,$$

then we have proved that

$$\prod_{i=1}^{n} x_i \leq \left[\frac{1}{n}\sum_{i=1}^{n} x_i\right]^n \quad \Rightarrow \left[\prod_{i=1}^{n} x_i\right]^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{n} x_i.$$

This completes the proof of (4.33). □

Example 4.3.
Two plants produce power at the costs $c_1(p_1)$ and $c_2(p_2)$ (the costs $c_i(p_i)$ are expressed in $K\$$ and the power $p_i \geq 0$ in MW):

$$c_1(p_1) = \begin{Bmatrix} 2, & 0 \leq p_1 \leq 1 \\ 2+(p_1-1)^2, & p_1 > 1 \end{Bmatrix}, c_2(p_2) = \begin{Bmatrix} 1, & 0 \leq p_2 \leq 2 \\ 1+2(p_2-2)^2, & p_2 > 2 \end{Bmatrix}.$$

Note that there is an initial cost associated to each plant whether it is producing or not. The units are

$$1\, MW = 1\ 000\ 000 \text{ Watts}, \quad 1\, K\$ = 1\ 000\$.$$

The objective is to distribute the productions $p_1 \geq 0$ and $p_2 \geq 0$ to satisfy the total demand $P \geq 0$ in MW while minimizing the sum of the production costs

$$f(p_1, p_2) \stackrel{\text{def}}{=} c_1(p_1) + c_2(p_2).$$

Since $p_1 \geq 0$, $p_2 \geq 0$, and $p_1 + p_2 = P \geq 0$, we have $0 \leq p_1 \leq P$ and $0 \leq p_2 \leq P$ and the set

$$U \stackrel{\text{def}}{=} \{(p_1, p_2) : 0 \leq p_1, 0 \leq p_2, p_1 + p_2 = P\}$$

is compact. Since the objective function $f(p_1, p_2)$ is continuous, the minimization problem

$$\inf_{(p_1,p_2)\in U} f(p_1, p_2)$$

has global minimizers.

Introduce the constraint functions

$$g_1(p_1, p_2) = -p_1 \leq 0, \quad g_2(p_1, p_2) = -p_2 \leq 0, \quad g_3(p_1, p_2) = P - p_1 - p_2 = 0.$$

Since all those functions and the functions c_1 and c_2 are of class $C^{(1)}$, Theorem 4.1(iii) can be used with the Lagrangian

$$L(p_1, p_2, \lambda_0, \lambda_1, \lambda_2, \mu) \stackrel{\text{def}}{=} \lambda_0 [c_1(p_1) + c_2(p_2)] - \lambda_1 p_1 - \lambda_2 p_2 + \mu [P - p_1 - p_2].$$

There exist multipliers, not all zero, such that

$$\begin{gathered} p_1 \geq 0,\, \lambda_1 \geq 0,\, p_1 \lambda_1 = 0 \\ p_2 \geq 0,\, \lambda_2 \geq 0,\, p_2 \lambda_2 = 0 \\ p_1 + p_2 = P \\ \lambda_0 \frac{dc_1}{dp_1}(p_1) - \lambda_1 - \mu = 0 \\ \lambda_0 \frac{dc_2}{dp_2}(p_2) - \lambda_2 - \mu = 0. \end{gathered}$$

If $\lambda_0 = 0$, then

$$\lambda_1 = \lambda_2 = -\mu \neq 0 \quad \Rightarrow \lambda_1 > 0 \text{ and } \lambda_2 > 0.$$

This yields $p_1 = 0 = p_2$ and hence $P = 0$. For $P > 0$, we get a contradiction.

So, when $P > 0$, we can choose $\lambda_0 = 1$ and the system becomes

$$p_1 \geq 0, \lambda_1 \geq 0, p_1 \lambda_1 = 0 \tag{4.42}$$
$$p_2 \geq 0, \lambda_2 \geq 0, p_2 \lambda_2 = 0 \tag{4.43}$$
$$p_1 + p_2 = P \tag{4.44}$$
$$\frac{dc_1}{dp_1}(p_1) - \lambda_1 - \mu = 0 \tag{4.45}$$
$$\frac{dc_2}{dp_2}(p_2) - \lambda_2 - \mu = 0. \tag{4.46}$$

Multiply (4.45) by p_1 and (4.46) by p_2:

$$\left. \begin{array}{l} p_1 \dfrac{dc_1}{dp_1}(p_1) - p_1\lambda_1 - p_1\mu = 0 \\ p_2 \dfrac{dc_2}{dp_2}(p_2) - p_2\lambda_2 - p_2\mu = 0 \end{array} \right| \quad \Rightarrow \; p_1 \frac{dc_1}{dp_1}(p_1) + p_2 \frac{dc_2}{dp_2}(p_2) = P\mu.$$

At this juncture, note that the functions c_1 and c_2 have nonnegative derivatives:

$$\frac{dc_1}{p_1}(p_1) = \begin{Bmatrix} 0, & 0 \leq p_1 \leq 1 \\ 2(p_1 - 1), & p_1 > 1 \end{Bmatrix}, \; c_2(p_2) = \begin{Bmatrix} 0, & 0 \leq p_2 \leq 2 \\ 4(p_2 - 2), & p_2 > 2 \end{Bmatrix}.$$

As a result,

$$\mu \geq 0, \quad p_1 \frac{dc_1}{dp_1}(p_1) \geq 0, \quad p_2 \frac{dc_2}{dp_2}(p_2) \geq 0.$$

Thus we have four cases to consider:

$$\text{(a)} \left| \begin{array}{l} p_1 \dfrac{dc_1}{dp_1}(p_1) > 0 \\ p_2 \dfrac{dc_2}{dp_2}(p_2) > 0 \end{array} \right. \quad \text{(b)} \left| \begin{array}{l} p_1 \dfrac{dc_1}{dp_1}(p_1) = 0 \\ p_2 \dfrac{dc_2}{dp_2}(p_2) > 0 \end{array} \right. \quad \text{(c)} \left| \begin{array}{l} p_1 \dfrac{dc_1}{dp_1}(p_1) > 0 \\ p_2 \dfrac{dc_2}{dp_2}(p_2) = 0 \end{array} \right. \quad \text{(d)} \left| \begin{array}{l} p_1 \dfrac{dc_1}{dp_1}(p_1) = 0 \\ p_2 \dfrac{dc_2}{dp_2}(p_2) = 0. \end{array} \right.$$

Case (a).

$$\begin{array}{lll} p_1 \dfrac{dc_1}{dp_1}(p_1) > 0 & p_2 \dfrac{dc_2}{dp_2}(p_2) > 0 & \Rightarrow \mu > 0 \\ p_1 > 1 & p_2 > 2 & \\ \lambda_1 = 0 & \lambda_2 = 0 & \\ 2(p_1 - 1) - \mu = 0 & 4(p_2 - 2) - \mu = 0 & \\ p_1 = 1 + \dfrac{\mu}{2} & p_2 = 2 + \dfrac{\mu}{4} & P = p_1 + p_2 = 3 + \dfrac{3}{4}\mu > 3 \\ 0 < \mu = \dfrac{4}{3}(P - 3) & \Rightarrow P > 3 \text{ and} & \begin{cases} p_1 = 1 + \dfrac{2}{3}(P - 3) \\ p_2 = 2 + \dfrac{1}{3}(P - 3). \end{cases} \end{array}$$

Case (b).

$$\begin{array}{lll}
p_1 \dfrac{dc_1}{dp_1}(p_1) = 0 & p_2 \dfrac{dc_2}{dp_2}(p_2) > 0 & \Rightarrow \mu > 0 \\
0 \leq p_1 \leq 1 & p_2 > 2 & \\
\dfrac{dc_1}{dp_1}(p_1) = 0 & \lambda_2 = 0 & \\
-\lambda_1 - \mu = 0 & 4(p_2 - 2) - \mu = 0 & \\
0 > -\mu = \lambda_1 \geq 0 & p_2 - 2 > 0 & \\
\Rightarrow \text{contradiction.} & &
\end{array}$$

Case (c).

$$\begin{array}{lll}
p_1 \dfrac{dc_1}{dp_1}(p_1) > 0 & p_2 \dfrac{dc_2}{dp_2}(p_2) = 0 & \Rightarrow \mu > 0 \\
p_1 > 1 & 0 \leq p_2 \leq 2 & \\
\lambda_1 = 0 & \dfrac{dc_2}{dp_2}(p_2) = 0 & \\
2(p_1 - 1) - \mu = 0 & -\lambda_2 - \mu = 0 & \\
p_1 - 1 > 0 & 0 > -\mu = \lambda_2 \geq 0 & \\
 & \Rightarrow \text{contradiction.} &
\end{array}$$

Case (d).

$$\begin{array}{lll}
p_1 \dfrac{dc_1}{dp_1}(p_1) = 0 & p_2 \dfrac{dc_2}{dp_2}(p_2) = 0 & \Rightarrow \mu = 0 \\
0 \leq p_1 \leq 1 & 0 \leq p_2 \leq 2 & \Rightarrow P = p_1 + p_2 \leq 1 + 2 = 3 \\
\dfrac{dc_1}{dp_1}(p_1) = 0 & \dfrac{dc_2}{dp_2}(p_2) = 0 & \\
-\lambda_1 - \mu = 0 & -\lambda_2 - \mu = 0 & \Rightarrow \lambda_1 = \lambda_2 = -\mu = 0.
\end{array}$$

Case (d) is complementary to case (a), that is, $0 < P \leq 3$. In this case, the objective function is constant and is equal to $1 + 2 = 3$. The solution is unique for $P = 3$ with $p_1 = 1$ and $p_2 = 2$, but for $0 < P < 3$, any combination of $0 \leq p_1 \leq 1$ and $0 \leq p_2 \leq 2$ such that $p_1 + p_2 = P$ is a minimizer.

Since the constraint functions are linear or affine, we have the stronger necessary condition with $\lambda_0 = 1$ by Theorem 7.13 of Chapter 4. □

5 Exercises

Exercise 5.1.
Find the global minima of

$$f(x) = x_1 + (x_2 - 1)^2$$

with respect to the set

$$U = \{x = (x_1, x_2) \in \mathbb{R}^2 \,:\, g(x) = x_1^2 = 0\}$$

by using the Lagrange multipliers theorem. If this is not possible or if it provides no information, use the method of your choice. □

Exercise 5.2.
Find the maximum and the minimum of the function

$$f(x,y,z) = -4xy - z^2$$

with respect to the set

$$U = \left\{(x,y,z) \,:\, x^2 + y^2 + z^2 = 1\right\}$$

if they exist. □

Exercise 5.3.
Consider the parabola P and the line L:

$$P \stackrel{\text{def}}{=} \left\{(x_1, y_1) \,:\, y_1 = 2x_1^2\right\}, \quad L \stackrel{\text{def}}{=} \{(x_2, y_2) : y_2 = -1 + x_2\}.$$

(i) First prove the existence of a solution to the problem of finding the shortest distance between P and L.

(ii) Find the points of P and L that achieve that distance and compute that distance. □

Exercise 5.4. (i) Find the local minima of the function

$$f(x) = (x_1 - 3)^2 + (x_2 - 4)^2$$

with respect to the constraint

$$U_0 = \left\{(x_1, x_2) \in \mathbb{R}^2 \,:\, x_1^2 + x_2^2 \le 1\right\}.$$

(ii) Introduce the extra variable x_3 and transform the constraint U_0 into

$$U_1 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 + x_3^2 = 1\right\}.$$

Find the local minima of the function

$$f(x_1, x_2, x_3) = (x_1 - 3)^2 + (x_2 - 4)^2$$

with respect to U_1.

(iii) What can you conclude by comparing the results of (i) and (ii)? □

Exercise 5.5.
Find the local minima of the function

$$f(x) = 2x_1^2 + x_2^2 - 2x_1 + 4x_2, \;\; x = (x_1, x_2) \in \mathbb{R}^2$$

with respect to

$$\begin{aligned} &U_0 = \mathbb{R}^2, \\ &U_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 \,:\, g(x_1, x_2) = 1 - x_1^2 + x_2 \le 0\right\}, \\ &U_2 = \{(x_1, x_2) \,|\, x_1 \ge 0, \; x_2 \ge 0\}. \end{aligned} \qquad \square$$

Exercise 5.6.
Find the minimizers of the function

$$f(x) = |x|^5 - 5x$$

with respect to the set

$$U = \left\{x \in \mathbb{R} : g(x) = x^2 - 8x + 15 \le 0\right\}. \qquad \square$$

Exercise 5.7.
Let A be a positive definite symmetric 2×2 matrix, b a vector of $\mathbb{R}^2$ of norm 1, and $c \in \mathbb{R}$ a scalar. Define the ellipse E and the line L:

$$E \stackrel{\text{def}}{=} \left\{x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : Ax \cdot x \le 1\right\}, \quad L \stackrel{\text{def}}{=} \left\{y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : b \cdot y = c\right\}. \tag{5.1}$$

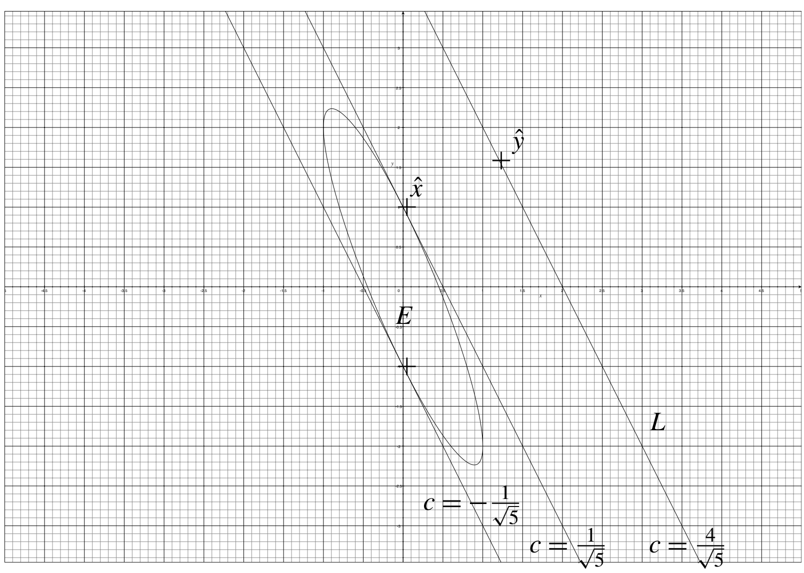


Figure 5.7. *The ellipse E and the line L as a function of c.*

(i) Prove that there exist points $(x, y) \in E \times L$ that minimize the distance $\|x - y\|$ between E and L.

(ii) Prove that, if $E \cap L = \varnothing$, the minimizing points are given by the expressions

$$x = \frac{c}{|c|} \frac{1}{\sqrt{A^{-1}b \cdot b}} A^{-1}b \in E \quad \text{and} \quad y = x + \frac{c}{|c|}\left(|c| - \sqrt{A^{-1}b \cdot b}\right) b \in L,$$

that $|c| > \sqrt{A^{-1}b \cdot b}$, and that the minimum distance is given by

$$\|x - y\| = |c| - \sqrt{A^{-1}b \cdot b}.$$

Figure 5.7 shows the ellipse and the line for the parameters

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}, \quad b = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad c = \frac{4}{\sqrt{5}}$$

$$A^{-1} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}, \quad A^{-1}b \cdot b = \frac{1}{5}, \quad \sqrt{A^{-1}b \cdot b} = \frac{1}{\sqrt{5}}$$

$$\hat{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \hat{y} = \hat{x} + \frac{3}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{2}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \|\hat{y} = \hat{x}\| = \frac{3}{5}\sqrt{5}$$

and the two cases $c = \pm 1/\sqrt{5}$ when the line is tangent to the ellipse. □

Appendix A. Inverse and Implicit Function Theorems

1 Inverse Function Theorem

Given an arbitrary function $g : \mathbb{R}^n \to \mathbb{R}^m$ and $y \in \mathbb{R}^m$, consider the problem of finding $x \in \mathbb{R}^n$ such that

$$g(x) = y, \tag{1.1}$$

that is, the inverse image of $g^{-1}\{y\}$ of y. This amounts *to solve equation* (1.1). If we now consider equation (1.1) for all y, it is possible that, for each y, equation (1.1) has a unique solution $x = x(y)$. In that case, starting with the solution of equation (1.1), we can define the function $y \mapsto h(y) = x(y)$ of y. This function is called the *implicit function* defined by equation (1.1). It is characterized by the property

$$g(h(y)) = y. \tag{1.2}$$

Naturally, such good conditions seldom occur. It might happen that for some values of y, there are no solution x, and that, for other values of y, there are several solutions x, even an infinity. The special case of interest to us is the following: assuming that we have a solution of equation (1.1); that is, for $y = y_0$, there exists $x = x_0$ such that $y_0 = g(x_0)$. Our objective is to find out if for points sufficiently close to y_0, equation (1.1) has a unique solution near x_0. Under those conditions we indeed define an implicit function $x = h(y)$ from the equation *at least in a neighborhood of* (x_0, y_0). The geometrical interpretation is simple. The equation

$$g(x) - y = 0$$

defines a manifold in $\mathbb{R}^n \times \mathbb{R}^m$, and we want to express this curve in the usual resolved form by computing x as a function of y, in a neighborhood of the point (x_0, y_0).

Theorem 1.1 (Inverse function theorem[6]). *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be a function of class $C^{(1)}$ on a neighborhood W of a regular point x_0 of g and $y_0 = g(x_0)$. Then*

(i) *there exist open subsets U and V of $\mathbb{R}^n$ such that $x_0 \in U$, $y_0 \in V$ and a function g bijective from U to V, that is, g injective on U and $g(U) = V$.*

[6]Cf. for instance W. Rudin [**1**, Thm. 9.17, p. 193].

(ii) *If g^{-1} is the inverse function of g (it exists according to* (i)*) defined on V by*

$$g^{-1}(g(x)) = x, \quad x \in U, \tag{1.3}$$

then g^{-1} is of class $C^{(1)}$ on V and $Dg^{-1}(y) = [Dg(g^{-1}(y))]^{-1}$ or in the more compact form

$$Dg^{-1} = [Dg]^{-1} \circ g^{-1} \text{ on } V.$$

Proof. See, for instance, W. RUDIN [**1**, Thm. 9.17, p. 193]. □

Remark 1.1.
It can also be shown that, under the assumptions of Theorem 1.1, g^{-1} is of class $C^{(k)}$, $k \geq 2$, if g is of class $C^{(k)}$. □

2 Implicit Function Theorem

There are several versions of the implicit function theorem (see, for instance, A. DONTCHEV and R. T. ROCKAFELLAR [**1**]). This section gives a version suited for our purposes directly from the inverse function theorem.

Theorem 2.1 (Implicit function theorem)**.** *Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be a function of class $C^{(1)}$ in a neighborhood of a regular point x_0 (that is, $Dg(x_0)$ is surjective). Then, for all $h \in \mathbb{R}^n$, there exist $t_0 > 0$ and a function $x :]-t_0, t_0[\to \mathbb{R}^n$ of class $C^{(1)}$ such that*

$$\begin{cases} x(0) = x_0 \\ x'(0) = h \end{cases} \text{ and } g(x(t)) = g(x_0) + t\,Dg(x_0)h, \quad -t_0 < t < t_0. \tag{2.1}$$

Proof. Given $\alpha \in \mathbb{R}^m$ and $t \in \mathbb{R}$, define the function $\gamma : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ as

$$\gamma(t,\alpha) \stackrel{\text{def}}{=} g\left(x_0 + \sum_{j=1}^{m} \alpha_j \nabla g_j(x_0) + th \right) - g(x_0) - t\,Dg(x_0)h. \tag{2.2}$$

Recall that

$$Dg(x_0)^\top \alpha = \sum_{j=1}^{m} \alpha_j \nabla g_j(x_0). \tag{2.3}$$

By construction, $\gamma(0,0) = 0$ and

$$D_\alpha \gamma(t,\alpha) = Dg(x_0 + Dg(x_0)^\top \alpha + th) Dg(x_0)^\top, \quad D_\alpha \gamma(0,0) = Dg(x_0) Dg(x_0)^\top.$$

Finally, the matrix $g'(x_0) g'(x_0)^\top$ associated with $Dg(x_0) Dg(x_0)^\top$ is positive definite (and hence invertible and of maximal rank m),

$$g'(x_0) g'(x_0)^\top > 0,$$

since $Dg(x_0)$ is surjective and hence $Dg(x_0)^\top$ is injective. It can also be checked that $(t,\alpha) \mapsto \gamma(t,\alpha)$ is of class $C^{(1)}$.

Define

$$G(t,\alpha) \overset{\text{def}}{=} (t,\gamma(t,\alpha)). \tag{2.4}$$

It is a transformation of $\mathbb{R}\times\mathbb{R}^m$ into itself such that

$$G(0,0) = (0,0)$$

with Jacobian matrix

$$G'(0,0) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & g'(x_0)g'(x_0)^\top & \\ 0 & & & \end{bmatrix}.$$

From the previous considerations, $G'(0,0)$ is invertible (hence of maximal rank) and the linear transformation $DG(0,0)$ is also invertible (and hence surjective). So, the conditions of application of Theorem 1.1 are fulfilled:

$$\exists N =]-e,e[,\ e>0, \text{ and a ball } V = B_r(0),\ r>0$$

such that G has an inverse

$$G^{-1} : G(N\times V) \to N\times V$$

of class $C^{(1)}$ and

$$\forall (t,\alpha)\in N\times V,\ \ G^{-1}G(t,\alpha) = (t,\alpha).$$

Since

$$G(0,0) = (0,0),$$

then $G(N\times V)$ is a neighborhood of $(0,0)$. Therefore, for all t, $|t|<e$,

$$\exists \alpha(t)\in V \text{ such that } (t,\alpha(t)) = G^{-1}(t,0) \text{ and } \alpha(0)=0. \tag{2.5}$$

Since G^{-1} is of class $C^{(1)}$, we conclude that the function α is of class $C^{(1)}$ from N into $\mathbb{R}^m$. Define

$$x(t) \overset{\text{def}}{=} x_0 + Dg(x_0)^\top\alpha(t) + th \quad \Rightarrow x(0) = x_0$$

and note that x is of class $C^{(1)}$ from N to $\mathbb{R}^n$. Moreover,

$$x'(t) = Dg(x_0)^\top\alpha'(t) + h \quad \Rightarrow x'(0) = Dg(x_0)^\top\alpha'(0) + h.$$

We also have from (2.5) and (2.4)

$$0 = \gamma(t,\alpha(t)) = g(x(t)) - g(x_0) - tDg(x_0)h$$

and this yields identity (2.1). Moreover,

$$\begin{aligned} 0 = \frac{d}{dt}\gamma(t,\alpha(t)) &= Dg(x(t))x'(t) - Dg(x_0)h \\ &= Dg(x(t))[Dg(x_0)^\top\alpha'(t)+h] - Dg(x_0)h \\ &= Dg(x(t))Dg(x_0)^\top\alpha'(t) + [Dg(x(t)) - Dg(x_0)]h. \end{aligned}$$

Letting t go to 0, we get

$$0 = Dg(x_0)Dg(x_0)^\top\alpha'(0) \quad \Rightarrow \alpha'(0) = 0 \quad \Rightarrow x'(0) = h. \qquad \square$$

Appendix B. Answers to Exercises

1 Exercises of Chapter 2

Exercise 10.1

Let $f,g:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be two functions that are lsc at a point $a \in \mathbb{R}^n$. Show that the following functions are lsc at a:

$$\text{(i)}\quad (f+g)(x) \stackrel{\text{def}}{=} f(x)+g(x) \quad \text{and} \quad \text{(ii)}\quad \forall \alpha \geq 0, \quad (\alpha f)(x) \stackrel{\text{def}}{=} \alpha f(x).$$

Answer. (i) To show that

$$\forall h \in \mathbb{R} \text{ such that } h < (f+g)(a) = f(a)+g(a)$$

there exists a neighborhood $V(a)$ of a such that

$$\forall x \in V(a), \quad h < (f+g)(x) = f(x)+g(x).$$

Consider several cases.

If $f(a) < +\infty$ and $g(a) < +\infty$, as $f(a)+g(a) > h$, choose

$$h_1 = f(a) - \frac{f(a)+g(a)-h}{2} < f(a)$$
$$h_2 = g(a) - \frac{f(a)+g(a)-h}{2} < g(a)$$

for which $h_1 + h_2 = h$. By the lower semicontinuity assumption at a for f and g,

$$\exists V_1(a) \text{ neighborhood of } a \text{ such that } \forall x \in V_1(a), \ f(x) > h_1$$
$$\exists V_2(a) \text{ neighborhood of } a \text{ such that } \forall x \in V_2(a), \ g(x) > h_2.$$

But $V(a) = V_1(a) \cap V_2(a)$ is also a neighborhood of a and

$$\forall x \in V(a), \ (f+g)(x) = f(x)+g(x) > h_1 + h_2 = h.$$

This shows that $(f+g)$ is lsc at a.

If $f(a) = +\infty$ or $g(a) = +\infty$. Assume that $f(a) = +\infty$. Since $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, there exists h_1 such that $h_1 < g(a)$ and, since g is lsc at a, there exists $V_g(a)$ such that

$$\forall x \in V_g(a), \quad h_1 < g(x).$$

As $h - h_1 \in \mathbb{R}$, we get $h - h_1 < +\infty = f(a)$. Since f is lsc at a, there exists $V_f(a)$ such that

$$\forall x \in V_f(a), \quad h - h_1 < f(x).$$

Hence, for all $x \in V_g(a) \cap V_f(a)$,

$$h = (h - h_1) + h_1 < f(x) + g(x)$$

and $f + g$ is lsc at a.

(ii) If $\alpha = 0$, by convention, $(\alpha f)(x) = \alpha f(x) = 0$ for all $x \in \mathbb{R}^n$. For each $h \in \mathbb{R}$ such that $h < \alpha f(a) = 0$, the set $V = \{x : h < \alpha f(x) = 0\} = \mathbb{R}^n$ is a neighborhood of a and αf is trivially lsc at a. For $\alpha > 0$, let $h \in \mathbb{R}$ be such that

$$h < (\alpha f)(a) = \alpha f(a).$$

So $f(a) > h/\alpha$. As f is lsc at a, there exists a neighborhood $V(a)$ of a such that

$$\forall x \in V(a), \quad \frac{h}{\alpha} < f(x) \quad \Rightarrow \forall x \in V(a), \quad h < \alpha f(x) = (\alpha f)(x).$$

Hence αf is lsc at a. $\square$

Exercise 10.2

Given $f : \mathbb{R}^n \to \mathbb{R}$, prove that

$$f \text{ is continuous at } a \quad \iff \quad f \text{ is lsc and usc at } a.$$

Answer. ($\Rightarrow$) By assumption on $f : \mathbb{R}^n \to \mathbb{R}$, we have $f(a) \in \mathbb{R}$. We first show that f is lsc at a. Let h be such that $h < f(a)$. Pick

$$\varepsilon \stackrel{\text{def}}{=} \frac{f(a) - h}{2} > 0.$$

By continuity, there exists $\delta > 0$ such that for all x, $\|x - a\| < \delta$, we have $|f(x) - f(a)| < \varepsilon$

$$\Rightarrow f(a) - \varepsilon < f(x) \quad \Rightarrow f(x) > f(a) - \varepsilon = \frac{f(a) + h}{2} > h.$$

Next, we show that f is usc at a. For all k such that $k > f(a)$, pick

$$\varepsilon \stackrel{\text{def}}{=} \frac{k - f(a)}{2} > 0.$$

By continuity, there exists $\delta > 0$ such that for all x, $\|x - a\| < \delta$, we have $|f(x) - f(a)| < \varepsilon$

$$\Rightarrow f(x) < f(a) + \varepsilon \quad \Rightarrow f(x) < f(a) + \varepsilon = \frac{f(a) + k}{2} > k.$$

($\Leftarrow$) For $\varepsilon > 0$, choose $h = f(a) - \varepsilon < f(a)$ and $k = f(a) + \varepsilon > f(a)$. Then, since f is lsc and usc at a,

$$\exists V_1(a) \text{ neighborhood of } a \text{ such that } \forall x \in V_1(a),\ f(x) > f(a) - \varepsilon$$
$$\exists V_2(a) \text{ neighborhood of } a \text{ such that } \forall x \in V_2(a),\ f(x) < f(a) + \varepsilon.$$

Therefore, $V(a) = V_1(a) \cap V_2(a)$ is a neighborhood of a and

$$\forall x \in \overbrace{V_1(a) \cap V_2(a)}^{V(a)},\ -\varepsilon < f(x) - f(a) < \varepsilon \quad \Rightarrow |f(x) - f(a)| < \varepsilon,$$

which proves the continuity of f at a. $\square$

Exercise 10.3

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be two lsc functions at a. Show that the function

$$(f \wedge g)(x) \stackrel{\text{def}}{=} \inf\{f(x), g(x)\}$$

is lsc at a.

Answer. Pick $h \in \mathbb{R}$ such that

$$h < (f \wedge g)(a) = \inf\{f(a), g(a)\}.$$

Then $h < f(a)$ and $h < g(a)$. Since f is lsc at a and g is lsc at a,

$$\exists V_1, \text{ a neighborhood of } a \text{ such that } \forall x \in V_1,\ h < f(x)$$
$$\exists V_2, \text{ a neighborhood of } a \text{ such that } \forall x \in V_2,\ h < g(x)$$

and we get

$$\forall x \in V = V_1 \cap V_2,\ (f \wedge g)(x) = \inf\{f(x), g(x)\} > h.$$

Since $V = V_1 \cap V_2$ is a neighborhood of a, $(f \wedge g)$ is lsc at a. $\square$

Exercise 10.4

Let $\{f_\alpha\}_{\alpha \in A}$, $f_\alpha : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, be a family of functions where the *index set* A is possibly infinite. Its *upper envelope* is defined as

$$f(x) \stackrel{\text{def}}{=} \sup_{\alpha \in A} f_\alpha(x),\ x \in \mathbb{R}^n.$$

If, for each $\alpha \in A$, f_α is lsc at $x_0 \in \mathbb{R}^n$, prove that f is lsc at x_0.

Answer. For each $h \in \mathbb{R}$ such that

$$h < f(a) = \sup_{\alpha \in A} f_\alpha(a),$$

we know, by definition of a supremum, that there exists $\alpha_1 \in A$ such that

$$h < f_{\alpha_1}(a) \leq \sup_{\alpha \in A} f_\alpha(a).$$

But f_{α_1} is lsc at a and there exists a neighborhood V of a such that

$$\forall x \in V, \;\; h < f_{\alpha_1}(x) \leq \sup_{\alpha \in A} f_\alpha(x) = f(x) \quad \Rightarrow f \text{ is lsc at } a. \qquad \square$$

Exercise 10.5

Consider the following function:

$$f(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } x \text{ is rational} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

Prove that f is lsc and not usc at each rational point; prove that f is usc and not lsc at each irrational point.

Answer. Let $\mathbb{Q}$ be the set of rational numbers in $\mathbb{R}$.

(i) At each $a \in \mathbb{Q}$, $f(a) = 0$. Let $h < f(a) = 0$. For all $y \in \mathbb{R}$, $f(y) \geq 0 > h$ and we can choose $V(a) = \mathbb{R}$ as a neighborhood of a. So f is lsc at each point of $\mathbb{Q}$. Since the irrational numbers are dense in $\mathbb{R}$, for each $\delta > 0$, $B_\delta(a) \cap (\mathbb{R} \setminus \mathbb{Q}) \neq \varnothing$ and hence for $b \in B_\delta(a) \cap (\mathbb{R} \setminus \mathbb{Q})$, we have $|f(b) - f(a)| = 1$, where $B_\delta(a) =]a - \delta, a + \delta[$. Therefore, the function f is not continuous at a and as f is lsc, it is not usc.

(ii) At each $a \in \mathbb{R} \setminus \mathbb{Q}$, $f(a) = 1$. Let $k > f(a) = 1$. At each $y \in \mathbb{R}$, $f(y) \leq 1 < k$ and we can choose $V(a) = \mathbb{R}$ as a neighborhood of a. So f is usc at each point of $\mathbb{R} \setminus \mathbb{Q}$. As the rationals are dense in $\mathbb{R}$, for all $\delta > 0$, $B_\delta(a) \cap \; Q \neq \varnothing$ and hence for $b \in B_\delta(a) \cap \mathbb{Q}$, we have $|f(b) - f(a)| = 1$, where $B_\delta(a) =]a - \delta, a + \delta[$. Therefore, the function f is not continuous at a and as f is usc, it is not lsc. $\square$

Exercise 10.6

Consider an lsc function $f : \mathbb{R}^n \to \mathbb{R}$. Prove or give a counterexample in each of the following cases.

(i) Is the function $f(x)^2$ lsc on $\mathbb{R}^n$?

(ii) Is the function $e^{f(x)}$ lsc on $\mathbb{R}^n$?

(iii) Is the function $e^{-f(x)}$ usc on $\mathbb{R}^n$?

(iv) Is the function (here $n = 1$)

$$F(x) \stackrel{\text{def}}{=} \begin{cases} f(x), & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

lsc on $\mathbb{R}$?

Answer. (i) No. In fact, it is sufficient to consider the following lsc function $f : \mathbb{R} \to \mathbb{R}$:

$$f(x) \stackrel{\text{def}}{=} \begin{cases} 0, & x \in \mathbb{R}\backslash\{0\} \\ -1, & x = 0 \end{cases}$$

whose square is usc but not lsc:

$$f(x)^2 \stackrel{\text{def}}{=} \begin{cases} 0, & x \in \mathbb{R}\backslash\{0\} \\ 1, & x = 0. \end{cases}$$

By definition, the function f is constant in the open set $\mathbb{R}\backslash\{0\}$: so it is continuous there. At $x = 0$, for all $h < f(0) = -1$, we have for all $x \in \mathbb{R}$, $h < -1 \leq f(x)$. Choosing the neighborhood $V(0) = \mathbb{R}$, f is lsc at 0. In a similar fashion, $f(x)^2$ is continuous on the open set $\mathbb{R}\backslash\{0\}$. However, it is not lsc at 0 since for $h = 1/2 < 1 = f(0)^2$, we have $h = 1/2 \not< 0 = f(x)$ for all $x \neq 0$.

(ii) Yes. To show that, at each point $x \in \mathbb{R}^n$ and for all $k > e^{-f(x)}$, there exists a neighborhood $V(x)$ of x such that for all $y \in V(x)$, $k > e^{-f(y)}$. As the function log is monotone strictly increasing and as $k > e^{-f(x)} > 0$, we get $\log k > -f(x)$ or $-\log k < f(x)$. As f is lsc at x, there exists a neighborhood $V(x)$ of x such that

$$\forall y \in V(x), \quad -\log k < f(y).$$

As above, $\log k > -f(y)$, and, since the exponential $t \mapsto e^t : \mathbb{R} \to \mathbb{R}$ is monotone strictly increasing, we get

$$\forall y \in V(x), \quad k = e^{\log k} > e^{-f(y)}$$

and, by definition, $e^{-f(x)}$ is usc at x.

(iii) Define $F(x) = e^{-f(x)}$ and note that $F(x) > 0$. At each point $a \in \mathbb{R}^n$,

$$\forall k \text{ such that } k > F(a) \quad \Rightarrow \log k > \log F(a) = -f(a),$$

since the log function is strictly increasing. As $-f$ is usc, there exists a neighborhood $V(a)$ such that

$$\forall x \in V(a), \quad \log k > \log F(x) = -f(x) \quad \Rightarrow k > F(x),$$

since the exponential function $x \mapsto e^x$ is also strictly increasing. Therefore, F is usc at a.

(iv) By definition, the function F is lsc on $\mathbb{R}\backslash\{0\}$. Since $F(0) = f(0)$, it is lsc at $x = 0$ if and only if $f(0) \leq 0$. □

Exercise 10.7

Consider the function

$$f(x) = |x|^{17} - x + 3x^2, \quad x \in \mathbb{R}.$$

(i) Do we have existence of minimizers of f in $\mathbb{R}$? If so, provide the elements of the proof.

(ii) Assuming that there are minimizers of f on $\mathbb{R}$, prove that they are positive or zero ($x \geq 0$).

Answer. (i) Existence of minimizers of f with respect to $\mathbb{R}$. Firstly, f is continuous:

$$\begin{array}{ll} x < 0, & f(x) = \text{sum of polynomials} \\ x > 0, & f(x) = \text{sum of polynomials} \\ \text{at } x = 0, & \lim_{x \searrow 0^+} f(x) = 0 = f(0) = 0 = \lim_{x \nearrow 0^-} f(x). \end{array}$$

Secondly, $\mathbb{R}$ is closed. Finally, f has the growth property at infinity:

$$\frac{f(x)}{|x|} = \frac{|x|^{17} - x + 3x^2}{|x|} \geq |x|^{16} - 1 + 3|x| \to +\infty \text{ as } |x| \to +\infty.$$

Therefore, f has a bounded lower section. By Theorem 5.3(i) of Chapter 2,

$$\exists a \in \mathbb{R} \text{ such that } f(a) = \inf_{r \in \mathbb{R}} f(x).$$

(ii) The minimizers x^* are $x^* \geq 0$. If $x^* < 0$, then

$$\begin{aligned} f(x^*) &= |x^*|^{17} - x^* + 3|x^*|^2 \\ &= |x^*|^{17} + |x^*| + 3|x^*|^2 > 0. \end{aligned}$$

But $f(0) = 0$ and this contradicts the minimality of f at $x = x^*$. Since global minimizers do exist, we necessarily have $x^* \geq 0$.

Alternative proof. For each $x < 0$,

$$\begin{aligned} f(-x) &= |-x|^{17} - (-x) + 3(-x)^2 \\ &= |x|^{17} + x + 3x^2 = f(x) + 2x < f(x). \end{aligned}$$

To each point $x < 0$, associate the point $-x > 0$ such that $f(-x) < f(x)$. Therefore, no point $x < 0$ can be a global minimizer and necessarily $x^* \geq 0$. □

Exercise 10.8

Let A be an $n \times n$ matrix and b a vector in $\mathbb{R}^n$. Define the function

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x, \quad x \in \mathbb{R}^n.$$

(i) Give necessary and sufficient conditions on A and b under which f is convex on $\mathbb{R}^n$.

(ii) Give necessary and sufficient conditions on A and b under which f has a unique minimizer in $\mathbb{R}^n$.

(iii) Give necessary and sufficient conditions on A and b under which f has a minimizer in $\mathbb{R}^n$.

(iv) Are the functions f associated with the following matrices A and vectors b convex on $\mathbb{R}^2$:

$$A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, b = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}?$$

Answer. First notice that the matrix A can be replaced by its symmetrized matrix $(A+A^\top)/2$ without changing the function f. As a result, A can be assumed symmetrical for the purpose of the proof.

(i) By definition, f is convex if for all $\lambda \in [0,1]$ and all x and y in $\mathbb{R}^n$,

$$f(\lambda x+(1-\lambda)y)-\lambda f(x)-(1-\lambda)f(y)\leq 0.$$

Since the term $b\cdot x$ is linear in the variable x, it disappears and the condition on f is equivalent to the following condition on A:

$$\begin{aligned}
&f(\lambda x+(1-\lambda)y)-\lambda f(x)-(1-\lambda)f(y)\\
&=A(\lambda x+(1-\lambda)y)\cdot(\lambda x+(1-\lambda)y)-\lambda(Ax)\cdot x-(1-\lambda)(Ay)\cdot y\\
&=(\lambda^2-\lambda)Ax\cdot x+\Big[(1-\lambda)^2-(1-\lambda)\Big]Ay\cdot y+2\lambda(1-\lambda)Ax\cdot y\\
&=-\lambda(1-\lambda)[Ax\cdot x+Ay\cdot y-2Ax\cdot y]\\
&=-\lambda(1-\lambda)A(x-y)\cdot(x-y)\leq 0.
\end{aligned}$$

As the inequality

$$-\lambda(1-\lambda)A(x-y)\cdot(x-y)\leq 0$$

must be verified for all $\lambda\in[0,1]$ and all $x,y\in\mathbb{R}^n$, for $\lambda=1/2$ and $y=0$, we get

$$\forall x\in\mathbb{R}^n,\quad Ax\cdot x\geq 0$$

and $A\geq 0$ is positive semidefinite. There are no condition on the linear term $b\cdot x$. Conversely, if $A\geq 0$ is positive semidefinite, then, from the previous computation,

$$f(\lambda x+(1-\lambda)y)-\lambda f(x)-(1-\lambda)f(y)=-\lambda(1-\lambda)A(x-y)\cdot(x-y)\leq 0$$

and f is convex.

(ii) If there exists a unique minimizer $\hat{x}$, then

$$\begin{aligned}
\forall x\neq\hat{x},\quad 0&<f(x)-f(\hat{x})\\
&=\frac{1}{2}\big[Ax\cdot x-A\hat{x}\cdot\hat{x}\big]+b\cdot(x-\hat{x})\\
&=A\frac{x+\hat{x}}{2}\cdot(x-\hat{x})+b\cdot(x-\hat{x}).
\end{aligned}$$

Associate with $h\neq 0$ the points $x=\hat{x}\pm 2h\neq\hat{x}$:

$$\begin{aligned}
0&<A(\hat{x}+h)\cdot(+2h)+b\cdot(+2h)=2Ah\cdot h+2\big[A\hat{x}\cdot h+b\cdot h\big]\\
0&<A(\hat{x}-h)\cdot(-2h)+b\cdot(-2h)=2Ah\cdot h-2\big[A\hat{x}\cdot h+b\cdot h\big].
\end{aligned}$$

By adding the two inequalities, we get

$$\forall h\neq 0,\quad Ah\cdot h>0$$

and A is positive definite. Conversely, f is continuous since it is polynomial, $\mathbb{R}^n$ is closed. In addition, f has the growth property at infinity since by Lemma 5.1 of Chapter 2, there exists $\alpha > 0$ such that $Ax \cdot x \geq \alpha \|x\|^2$:

$$\frac{f(x)}{\|x\|} \geq \frac{\alpha}{2}\|x\| - \|b\| \to \infty \text{ as } \|x\| \to \infty.$$

So there exist minimizers by Theorem 5.3(i) of Chapter 2. As for the uniqueness, it is sufficient to observe that from part (i), f is not only convex but also strictly convex since, from the previous computation, for all $\lambda \in]0,1[$ and $x \neq y$,

$$f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda)f(y) = -\lambda(1-\lambda)A(x-y)\cdot(x-y) < 0.$$

(iii) If there exists a minimizer $\hat{x}$, then

$$\forall x, \quad 0 \leq f(x) - f(\hat{x}) = A\frac{x+\hat{x}}{2}\cdot(x-\hat{x}) + b\cdot(x-\hat{x}).$$

Associate with $h \in \mathbb{R}^n$ the points $x = \hat{x} \pm 2h$:

$$\begin{aligned} 0 &\leq 2Ah\cdot h + 2(A\hat{x}\cdot h + b\cdot h) \\ 0 &\leq 2Ah\cdot h - 2(A\hat{x}\cdot h + b\cdot h). \end{aligned}$$

By adding the two inequalities, we get

$$\forall h, \quad Ah\cdot h \geq 0$$

and A is positive semidefinite. Again with $x = \hat{x} \pm 2t\,h$ and $t > 0$, divide the two expressions by $2t$:

$$\begin{aligned} 0 &< A(\hat{x} + th)\cdot(+h) + b\cdot(+h) \\ 0 &< A(\hat{x} - th)\cdot(-h) + b\cdot(-h). \end{aligned}$$

As t goes to 0,

$$\forall h \in \mathbb{R}^n, \quad 0 < \pm(A\hat{x} + b)\cdot h \quad \Rightarrow A\hat{x} + b = 0.$$

Conversely, assume that $A \geq 0$ and that there exists $\hat{x}$ such that $A\hat{x} + b = 0$. From the computation of part (ii), for all x,

$$\begin{aligned} f(x) - f(\hat{x}) &= A\frac{x+\hat{x}}{2}\cdot(x-\hat{x}) + b\cdot(x-\hat{x}) \\ &= (A\hat{x} + b)\cdot(x-\hat{x}) + \frac{1}{2}A(x-\hat{x})\cdot(x-\hat{x}) \geq 0 \end{aligned}$$

and $\hat{x}$ is a minimizer of f.

(iv) To prove the convexity of f, use the symmetrized matrix of A:

$$A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, \qquad \frac{A + A^\top}{2} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} > 0.$$

Then f is convex and even strictly convex. As for the second case, as the symmetrized matrix

$$A = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, \qquad \frac{A + A^\top}{2} = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}$$

is not a diagonal matrix, consider the quadratic form

$$\begin{aligned} \frac{A + A^\top}{2} v \cdot v &= 2v_1^2 + 8v_1 v_2 + v_2^2 \\ &= (v_2 + 4v_1)^2 - 16v_1^2 + 2v_1^2 \\ &= (v_2 + 4v_1)^2 - 14v_1^2. \end{aligned}$$

For $v_2 + 4v_1 = 0$ and $v_1 \neq 0$, $\frac{A+A^\top}{2} \not\geq 0$, (it is sufficient to choose $v_1 = 1$ and $v_2 = -4$). Therefore, the associated function f is not convex. □

Exercise 10.9

Let A be a symmetric $n \times n$ positive definite matrix and b a vector of $\mathbb{R}^n$. Prove that the function

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x, \quad x \in \mathbb{R}^n$$

is strongly convex on $\mathbb{R}^n$.

Answer. Since A is positive definite, by Lemma 5.1 of Chapter 2, there exists $\alpha > 0$ such that for all $v \in \mathbb{R}^n$, $Av \cdot v \geq \alpha \|v\|^2$. The function f is convex from part (i) of the previous exercise. Moreover, from the previous computation with $\lambda = 1/2$,

$$f\left(\frac{x+y}{2}\right) - \frac{1}{2} f(x) - \frac{1}{2} f(y) = -\frac{1}{4} A(x-y) \cdot (x-y) \leq -\frac{\alpha}{4} \|x-y\|^2$$

and f is strongly convex. □

Exercise 10.10

Find and justify the Fenchel–Legendre transforms of the following functions: (i) $f(x) = a \cdot x + b$,

$$f^*(x^*) = \begin{cases} -b, & \text{if } x^* = a \\ +\infty, & \text{if } x^* \neq a \end{cases} \quad \text{and} \quad f^{**}(x) = a \cdot x + b = f(x);$$

and (ii) $f(x) = \|x\|$,

$$f^*(x^*) = \begin{cases} 0, & \text{if } \|x^*\| \leq 1 \\ +\infty, & \text{if } \|x^*\| > 1 \end{cases} \quad \text{and} \quad f^{**}(x) = \|x\| = f(x).$$

Answer. (i) Given x^* and x,

$$x^* \cdot x - (a \cdot x + b) = (x^* - a) \cdot x - b$$
$$\Rightarrow\ f^*(x^*) = \sup_{x \in \mathbb{R}^n} x^* \cdot x - (a \cdot x + b) = \begin{cases} -b, & \text{if } x^* = a \\ +\infty, & \text{if } x^* \neq a. \end{cases}$$

Then

$$x \cdot x^* - f^*(x^*) = \begin{cases} x \cdot a + b, & \text{if } x^* = a \\ -\infty, & \text{if } x^* \neq a. \end{cases}$$
$$\Rightarrow\ f^{**}(x) = \sup_{x^* \in \mathbb{R}^n} x \cdot x^* - f^*(x^*) = x \cdot a + b.$$

(ii) For $f(x) = \|x\|$ and x^*, consider the expression $x^* \cdot x - \|x\|$. For $\|x^*\| \leq 1$,

$$x^* \cdot x - \|x\| \leq \|x^*\|\,\|x\| - \|x\| \leq \|x\| - \|x\| = 0 \quad \Rightarrow\ \sup_{x \in \mathbb{R}^n} x^* \cdot x - \|x\| \leq 0$$

and $x^* = x/\|x\|$ achieves the maximum 0. If $\|x^*\| > 1$, choose the sequence $x_n = n\, x^*/\|x^*\|$:

$$x^* \cdot x_n - \|x_n\| = x^* \cdot n \frac{x^*}{\|x^*\|} - n = n\left(\|x^*\| - 1\right) \to +\infty$$
$$\Rightarrow\ f^*(x^*) = \sup_{x \in \mathbb{R}^n} x^* \cdot x - \|x\| = +\infty$$
$$\Rightarrow\ f^*(x^*) = \begin{cases} 0, & \text{if } \|x^*\| \leq 1 \\ +\infty, & \text{if } \|x^*\| > 1. \end{cases}$$

Then,

$$f^{**}(x) = x \cdot x^* - f^*(x^*) = \begin{cases} x \cdot x^*, & \text{if } \|x^*\| \leq 1 \\ -\infty, & \text{if } \|x^*\| > 1 \end{cases}$$
$$f^{**}(x^*) = \sup_{x^* \in \mathbb{R}^n} x \cdot x^* - f^*(x^*) = \sup_{\substack{x^* \in \mathbb{R}^n \\ \|x^*\| \leq 1}} x \cdot x^* = \|x\| = f(x).$$

$\square$

Exercise 10.11

Consider the function

$$f(x) \overset{\text{def}}{=} \frac{1}{2} Ax \cdot x + b \cdot x,$$

where A is an $n \times n$ symmetric matrix and $b \in \mathbb{R}^n$. Find and justify the Fenchel–Legendre transform when A is (i) positive definite and (ii) positive semidefinite.

Answer. (i) By definition,

$$f^*(x^*) = \sup_{x \in \mathbb{R}^n} x^* \cdot x - \left(\frac{1}{2} Ax \cdot x + b \cdot x\right) = -\inf_{x \in \mathbb{R}^n} \left(\frac{1}{2} Ax \cdot x + (b - x^*) \cdot x\right).$$

For $A > 0$, we have shown in Exercise 10.8(ii) and (iii) of Chapter 2 that there exists a unique solution $\hat{x}$ that satisfies the equation $A\hat{x} + b - x^* = 0$. Since A is invertible, $\hat{x} = -A^{-1}(b - x^*)$ and

$$f^*(x^*) = \frac{1}{2} A^{-1}(x^* - b) \cdot (x^* - b).$$

(ii) If $A \geq 0$, we have seen in Exercise 10.8(iii) that the necessary and sufficient condition that $\hat{x}$ be a minimizer is that $\hat{x}$ satisfies the equation $A\hat{x} + b - x^* = 0$, that is, the points that belong to the set

$$A^{-1}\{x^* - b\} \stackrel{\text{def}}{=} \{z \in \mathbb{R}^n : Az + b - x^* = 0\}.$$

Thus, for all x^* of the form $x^* = Ay + b$,

$$\inf_{x \in \mathbb{R}^n} \left(\frac{1}{2} Ax \cdot x + (b - x^*) \cdot x \right) = -\frac{1}{2}(Ay) \cdot y, \quad \forall y \in \mathbb{R}^n$$
$$\Rightarrow f^*(Ay + b) = \frac{1}{2}(Ay) \cdot y, \quad \forall y \in \mathbb{R}^n.$$

It is readily seen that this function is well defined. In fact, if there exist y_1 and y_2 such that $Ay_1 + b = x^* = Ay_2 + b$, then $Ay_1 = Ay_2$ and, by symmetry of A,

$$f^*(Ay_1 + b) = \frac{1}{2}(Ay_1) \cdot y_1 = \frac{1}{2}(Ay_2) \cdot y_1 = \frac{1}{2} y_2 \cdot Ay_1 = \frac{1}{2} y_2 \cdot Ay_2 = f^*(Ay_2 + b).$$

If $A^{-1}\{x^* - b\} = \varnothing$, that is, $b - x^* \notin \operatorname{Im} A = [\operatorname{Ker} A^\top]^\perp = [\operatorname{Ker} A]^\perp$, let p be the projection of $b - x^*$ onto the linear subspace $\operatorname{Ker} A$ that is a solution of the minimization problem

$$\inf_{z \in \operatorname{Ker} A} \|z - (b - x^*)\|^2.$$

There exists a solution since the function is continuous and has the growth property at infinity, and $U = \operatorname{Ker} A$ is closed. The minimizer $p \in \operatorname{Ker} A$ is unique since the function is strictly convex on the convex $\operatorname{Ker} A$. Moreover, for all $z \in \operatorname{Ker} A$,

$$0 \leq \|z - (b - x^*)\|^2 - \|p - (b - x^*)\|^2 = \big(z + p - 2(b - x^*)\big) \cdot (z - p).$$

Substitute $z = p \pm th$ for $t > 0$ and $h \in \operatorname{Ker} A$ in the above inequality. Divide by t,

$$0 \leq 2\big(p \pm th - (b - x^*)\big) \cdot (\pm h).$$

As t goes to 0,

$$0 \leq \pm\big(p - (b - x^*)\big) \cdot h \quad \Rightarrow \forall h \in \operatorname{Ker} A, \quad \big(p - (b - x^*)\big) \cdot h = 0. \tag{1.1}$$

Since $b - x^* \notin [\operatorname{Ker} A]^\perp$, this last inequality yields $p \neq 0$. Now, choose the sequence $x_n = -np$ for which $Ax_n = -nAp = 0$. From identity (1.1) with $h = p$, we get

$$\frac{1}{2} Ax_n \cdot x_n + (b - x^*) \cdot x_n = -n\,(b - x^*) \cdot p = -n\,\|p\|^2 \to -\infty$$
$$\Rightarrow \inf_{x \in \mathbb{R}^n} \left(\frac{1}{2} Ax \cdot x + (b - x^*) \cdot x \right) = -\infty.$$

We conclude that

$$f^*(x^*) = -\inf_{x\in\mathbb{R}^n}\left(\frac{1}{2}Ax\cdot x + (b-x^*)\cdot x\right) = +\infty$$

for all x^* such that $A^{-1}\{x^*-b\} = \varnothing$. □

To summarize,

$$f^*(Ay+b) = \frac{1}{2}(Ay)\cdot y, \quad \forall y\in\mathbb{R}^n$$
$$f^*(x^*) = +\infty, \quad \forall x^*\notin \operatorname{Im} A + b.$$

For the double transform,

$$f^{**}(x) = \sup_{x^*\in\mathbb{R}^n} x\cdot x^* - f^*(x^*) = \sup_{y\in\mathbb{R}^n} x\cdot(Ay+b) - \frac{1}{2}(Ay)\cdot y = \frac{1}{2}(Ax)\cdot x + b\cdot x.$$

Exercise 10.12

The *infimum convolution* of two proper functions for the infimum $f,g:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is defined as

$$(f\,\square\, g)(x) \stackrel{\text{def}}{=} \inf_{y\in\mathbb{R}^n}[f(x-y)+g(y)]. \tag{1.2}$$

(i) Find the Fenchel–Legendre transform of $f\,\square\, g$.

(ii) Prove that $f\,\square\, g$ is convex if f and g are convex.

(iii) Compute $f\,\square\, g$ for $f(x) = \|x\|_{\mathbb{R}^n}$ and $g = I_U$, the indicator function of the nonempty set $U\subset\mathbb{R}^n$.

Answer. (i) It is a matter of rewriting things. Indeed

$$\begin{aligned}
(f\,\square\, g)^*(x^*) &\stackrel{\text{def}}{=} \sup_x\left(x^*\cdot x - \inf_y[f(x-y)+g(y)]\right)\\
&= \sup_x\sup_y\left[x^*\cdot x - f(x-y)-g(y)\right]\\
&= \sup_x\sup_y\left(\left[x^*\cdot(x-y)-f(x-y)\right]+\left[x^*\cdot y - g(y)\right]\right)\\
&= \sup_y\left(x^*\cdot y - g(y) + \sup_x\left[x^*\cdot(x-y)-f(x-y)\right]\right)\\
&= \sup_y\left(x^*\cdot y - g(y) + \sup_x\left[x^*\cdot x - f(x)\right]\right) = g^*(x^*)+f^*(x^*).
\end{aligned}$$

(ii) For the pair (x, ξ), $\lambda \in]0,1[$, and the pair (y, η),

$$\begin{aligned}
(f \,\Box\, g)(\lambda x + (1-\lambda)\xi) &\le f(\lambda x + (1-\lambda)\xi - (\lambda y + (1-\lambda)\eta)) + g(\lambda y + (1-\lambda)\eta) \\
&\le f(\lambda(x-y) + (1-\lambda)(\xi - \eta)) + \lambda\, g(y) + (1-\lambda)\, g(\eta) \\
&\le \lambda\, f(x-y) + (1-\lambda)\, f(z-y) + \lambda\, g(y) + (1-\lambda)\, g(\eta) \\
&\le \lambda\,[f(x-y) + g(y)] + (1-\lambda)\,[f(\xi - \eta) + g(\eta)] \\
&\le \lambda \inf_y\, [f(x-y) + g(y)] + (1-\lambda) \inf_\eta\, [f(\xi-\eta) + g(\eta)] \\
&= \lambda\, (f \,\Box\, g)(x) + (1-\lambda)\, (f \,\Box\, g)(\xi).
\end{aligned}$$

(iii) Compute

$$\inf_{y \in \mathbb{R}^n} (\|x - y\| - I_U(y)) = \inf_{y \in U} \|x - y\| = d_U(x),$$

where $d_U(x)$ is the distance function from x to U. Now

$$I_U^*(x^*) = \sup_{x \in U} x^* \cdot x$$

and for the norm $n(x) = \|x\|$,

$$\begin{aligned}
n^*(X^*) &= \sup_x x^* \cdot x - \|x\| \le \big(\|x^*\| - 1\big)\, \|x\| \\
\Rightarrow n^*(X^*) &= \left\{\begin{array}{ll} 0, & \text{if } \|x^*\| \le 1 \\ +\infty, & \text{if } \|x^*\| > 1 \end{array}\right\} = I_{\overline{B_1(0)}}(x^*) \\
\Rightarrow (n \,\Box\, I_U)^*(x^*) &= \sup_{x \in U} x^* \cdot x + I_{\overline{B_1(0)}}(x^*).
\end{aligned}$$

□

Exercise 10.13

(i) Let $B : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, and define the *marginal value* function

$$(Bf)(y) \overset{\text{def}}{=} \begin{cases} \inf_{x \in \mathbb{R}^n} \{f(x) : Bx = y\}, & \text{if } y \in \operatorname{Im} B \\ +\infty, & \text{if } y \notin \operatorname{Im} B. \end{cases} \tag{1.3}$$

Show that $(Bf)^* = f^* \circ B^\top$ and that, if f is convex, then Bf and $(Bf)^*$ are convex.

(ii) Let $A : \mathbb{R}^m \to \mathbb{R}^n$ be a linear map and $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Show that

$$(g^{**} \circ A)^* = (A^\top g^*)^{**}$$

and that, for g convex, we get the formula

$$(\text{cl } g \circ A)^* = \text{cl } (A^\top g^*)$$

(see R. T. ROCKAFELLAR [**1**, p. 38, Thm. 16.1, p. 142]).

Answer. (i) By definition, $Bf : \mathbb{R}^m \to \mathbb{R}$ and for $y^* \in \mathbb{R}^m$,

$$
\begin{aligned}
(Bf)^*(y^*) &= \sup_{y\in\mathbb{R}^n} \begin{cases} y^*\cdot y - \inf\limits_{x\in\mathbb{R}^n} \{f(x) : Bx = y\}, & \text{if } y \in \operatorname{Im} B \\ y^*\cdot y - \infty, & \text{if } y \notin \operatorname{Im} B \end{cases} \\
&= \sup_{y\in \operatorname{Im} B} \left(y^*\cdot y - \inf_{\substack{x\in\mathbb{R}^n \\ Bx=y}} f(x) \right) \\
&= \sup_{y\in \operatorname{Im} B} \sup_{\substack{x\in\mathbb{R}^n \\ Bx=y}} \big(y^*\cdot y - f(x)\big) = \sup_{x\in\mathbb{R}^n} \sup_{\substack{y\in\mathbb{R}^m \\ y=Bx}} \big(y^*\cdot Bx - f(x)\big) \\
&= \sup_{x\in\mathbb{R}^n} \Big(B^\top y^*\cdot x - f(x)\Big) = f^*(B^\top y^*) = (f^*\circ B^\top)(y^*).
\end{aligned}
$$

Assume that f is convex. Let $y_1, y_2 \in \operatorname{Im} B$ and $x_1, x_2 \in \mathbb{R}^n$ such that $Bx_1 = y_1$ and $Bx_2 = y_2$. For all $\lambda \in]0,1[[$, $B(\lambda x_1 + (1-\lambda)x_2) = \lambda\, Bx_1 + (1-\lambda)\, Bx_2 = \lambda\, y_1 + (1-\lambda)\, y_2$ and

$$
\begin{aligned}
(Bf)(\lambda y_1 + (1-\lambda)y_2) &\le f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) \\
(Bf)(\lambda y_1 + (1-\lambda)y_2) &\le \lambda \inf_{\substack{x_1\in\mathbb{R}^n \\ Bx_1=y_1}} f(x_1) + (1-\lambda) \inf_{\substack{x_2\in\mathbb{R}^n \\ Bx_2=y_2}} f(x_2)
\end{aligned}
$$

and Bf is convex in the linear subspace $\operatorname{Im} B$. Since Bf is $+\infty$ outside, Bf is convex on $\mathbb{R}^m$.

(iii) For $g : \mathbb{R}^n \to \mathbb{R}$ and $A : \mathbb{R}^m \to \mathbb{R}^n$, $g\circ A : \mathbb{R}^m \to \mathbb{R}^n$ and a quick computation yields

$$
\begin{aligned}
(g\circ A)^*(y^*) &= \sup_{y\in\mathbb{R}^m} \big\{y^*\cdot y - g(Ay)\big\}, \\
(A^\top g^*)(y^*) &= \begin{cases} \inf\limits_{x\in\mathbb{R}^n} \Big\{g^*(x) : A^\top x = y^*\Big\}, & \text{if } y^* \in \operatorname{Im} A^\top \\ +\infty, & \text{if } y^* \notin \operatorname{Im} A^\top. \end{cases}
\end{aligned}
$$

In general, $(g\circ A)^* \neq A^\top g^*$. By substituting $f = g^*$ and $B = A^\top$ in part (i), $A^\top g^*$ is convex and

$$
\begin{aligned}
(g^{**}\circ A &= (g^*)^*\circ (A^\top)^\top = (A^\top g^*)^* \\
\Rightarrow (g^{**}\circ A)^* &= \Big[(g^*)^*\circ (A^\top)^\top\Big]^* = (A^\top g^*)^{**}.
\end{aligned}
$$

If g is convex, then $g^{**} = \operatorname{cl} g$ and $(A^\top g^*)^{**} = \operatorname{cl}\,(A^\top g^*)$ that yields the formula of R. T. ROCKAFELLAR [**1**, p. 38, Thm. 16.1, p. 142]:

$$
(\operatorname{cl}\, g\circ A)^* = \operatorname{cl}\,(A^\top g^*). \qquad \square
$$

2 Exercises of Chapter 3

Exercise 7.1

Prove that the function (see Figure 3.10 of Chapter 3)

$$f(x,y) \stackrel{\text{def}}{=} \begin{cases} \dfrac{xy^2}{x^2+y^4}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

is directionally differentiable at $(x,y) = (0,0)$, but is neither Gateaux differentiable nor continuous at $(x,y) = (0,0)$. Note the following properties:

$$x < 0 \quad \Rightarrow f(x,y) \leq 0 \quad \text{and} \quad x > 0 \quad \Rightarrow f(x,y) \geq 0$$
$$f(-x,y) = -f(x,y) \quad \text{and} \quad f(x,-y) = f(x,y).$$

Answer. By definition,

$$f(x_1,x_2) = \frac{x_1 x_2^2}{x_1^2 + x_2^4} \text{ if } x_1 \neq 0, \;\; f(0,x_2) = 0$$

and for all $t \neq 0$, the differential quotient is given by

$$q(t) = \frac{f(tv) - f(0)}{t} = \begin{cases} 0, & \text{if } v_1 = 0 \\ \dfrac{tv_1(tv_2)^2/t}{(tv_1)^2 + (tv_2)^4}, & \text{if } v_1 \neq 0. \end{cases}$$

Hence,

$$df(0;v) = 0 \text{ if } v_1 = 0.$$

If $v_1 \neq 0$,

$$q(t) = \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4} \to \frac{v_2^2}{v_1} \text{ as } t \to 0 \quad \Rightarrow df(0;v) = \frac{v_2^2}{v_1}.$$

The function f is differentiable at $x = 0$ in all directions. However, it is not Gateaux differentiable at 0 since for $v_1 \neq 0$,

$$(v_1,v_2) \mapsto df(0,v) = v_2^2/v_1$$

is not linear.

To show that f is discontinuous at $x = 0$, it is sufficient to follow the curve $x_1 = x_2^2$, $x_2 \neq 0$, as x_2 goes to 0. In fact

$$f(x_2^2,x_2) = \frac{x_2^4}{x_2^4 + x_2^4} = \frac{1}{2} \neq 0 = f(0,0). \qquad \square$$

Exercise 7.2

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be linear (or an $n \times n$ matrix) and $b \in \mathbb{R}^n$. Define

$$f(x) \stackrel{\text{def}}{=} \frac{1}{2}(Ax)\cdot x + b \cdot x, \quad x \in \mathbb{R}^n.$$

(i) Compute $df(x;v)$ (or $\nabla f(x)$) and $d^2 f(x;v;w)$ (or $Hf(x)$).

(ii) Give necessary and sufficient conditions on A, b for the convexity of f.

(iii) Give necessary and sufficient conditions on A, b for the strict convexity of f.

(iv) Are the following functions f associated with the matrices and vectors convex?

$$\text{(a)} \quad A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, b = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \text{and} \quad \text{(b)} \quad A = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Answer. (i) Compute the directional derivative of f: for $t > 0$,

$$\begin{aligned} q(t) &= \frac{f(x+tv) - f(x)}{t} \\ &= \left(\frac{1}{2}[A(x+tv)\cdot(x+tv) - (Ax)\cdot x] + [b\cdot(x+tv) - b\cdot x] \right) / t \\ &= \frac{1}{2}\left(Ax\cdot x + Ax\cdot(tv) + A(tv)\cdot x + A(tv)\cdot(tv) - Ax\cdot x\right)/t + b\cdot v \\ &= \frac{1}{2}\left(Ax\cdot v + Av\cdot x + tAv\cdot v\right) + b\cdot v \end{aligned}$$

and as t goes to 0,

$$df(x;v) = \frac{1}{2}\left(Ax\cdot v + Av\cdot x\right) + b\cdot v.$$

It is readily seen that the map

$$v \mapsto df(x;v)$$

is linear and, if $A^\top$ denotes the transpose matrix of A,

$$A = (a_{ij}), \quad A^\top = (a_{ij}),$$

then

$$df(x;v) = \left(\frac{1}{2}[A + A^\top]x + b\right)\cdot v.$$

Therefore, the gradient is

$$\nabla f(x) = \frac{1}{2}[A + A^\top]x + b.$$

For the second-order semidifferential, consider the differential quotient

$$\begin{aligned} p(t) &\stackrel{\text{def}}{=} \frac{df(x+tw;v) - df(x;v)}{t} && (2.1) \\ &= \frac{1}{t}\left\{ \left(\frac{1}{2}[A + A^\top](x+tw) + b\right)\cdot v - \left(\frac{1}{2}[A + A^\top]x + b\right)\cdot v \right\} && (2.2) \\ &= \left(\frac{1}{2}[A + A^\top]w\right)\cdot v && (2.3) \end{aligned}$$

and

$$d^2 f(x;v;w) = \left(\frac{1}{2}[A + A^\top]w\right) \cdot v.$$

The Hessian matrix is

$$Hf(x) = \frac{1}{2}[A + A^\top].$$

As a result, f is also of class $C^{(2)}$ in $\mathbb{R}^n$ by Corollary 1 to Theorem 3.11 of Chapter 3.

As the function f is $C^{(2)}$, we can use Theorem 2.6 of Chapter 3. Since Hf is a constant matrix, for all $x, y \in \mathbb{R}^n$,

$$f(y) = f(x) + \nabla f(x) \cdot (y - x) + \frac{1}{4}[A + A^\top] \cdot (y - x).$$

Here, this yields

$$f(y) - f(x) - \nabla f(x) \cdot (y - x) = \frac{1}{4}[A + A^\top](y - x) \cdot (y - x).$$

By Theorem 4.1(ii) of Chapter 3,

$$\begin{aligned}
f \text{ is convex} \quad &\iff \quad \forall x, y \in \mathbb{R}^n, \;\; f(y) - f(x) - \nabla f(x) \cdot (y - x) \geq 0 \\
&\iff \quad \forall x, y \in \mathbb{R}^n, \;\; [A + A^\top](y - x) \cdot (y - x) \geq 0 \\
&\iff \quad \forall y \in \mathbb{R}^n, \;\; [A + A^\top]y \cdot y \geq 0 \\
&\iff \quad [A + A^\top] \geq 0.
\end{aligned}$$

(iii) By Theorem 4.2(ii) of Chapter 3, f is strictly convex if and only if for all $x \neq y$,

$$f(y) - f(x) - \nabla f(x) \cdot (y - x) > 0.$$

By the computation at the end of part (ii), this is equivalent to $[A + A^\top] > 0$.

(iv) It is sufficient to consider $A + A^\top$ to check the convexity of f:

$$A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, \qquad \frac{A + A^\top}{2} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} > 0 \quad \Rightarrow f \text{ convex.}$$

For the second matrix

$$A = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, \qquad \frac{A + A^\top}{2} = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix},$$

the associated bilinear form is

$$\frac{A + A^\top}{2} v \cdot v = 2v_1^2 + 8v_1v_2 + v_2^2 = (v_2 + 4v_1)^2 - 16v_1^2 + 2v_1^2 = (v_2 + 4v_1)^2 - 14v_1^2.$$

Hence, for $v_2 + 4v_1 = 0$ and $v_1 \neq 0$, $\frac{A+A^\top}{2} \not\geq 0$ (it is sufficient to choose $v_1 = 1$ and $v_2 = -4$). So, the associated function f is not convex. $\square$

Exercise 7.3

Let $f(x) = |x|^n$, and let $x \in \mathbb{R}$, $n \geq 1$ be an integer.

(i) Determine the values of n for which f is differentiable on $\mathbb{R}$.

(ii) Give the directional derivatives of f at $x = 0$ as a function of $n \geq 1$, if it exists. Otherwise, give the semidifferentials.

(iii) Determine the values of $n \geq 1$ for which f is convex on $\mathbb{R}$.

Answer. (i) The function f is $C^{(\infty)}$ on $]0,\infty[$, that is, infinitely differentiable, and

$$\forall x > 0, \;\; f(x) = x^n, \;\; \frac{df}{dx}(x) = nx^{n-1}.$$

Similarly, f is $C^{(\infty)}$ on $]-\infty,0[$ and

$$\forall x < 0, \;\; f(x) = (-x)^n, \;\; \frac{df}{dx}(x) = -n(-x)^{n-1}.$$

A problem might arise at $x = 0$: indeed the quotient

$$q(t) = \frac{f(tv) - f(0)}{t} = \frac{|tv|^n}{t} = (\text{sign } t)\,|t|^{n-1}|v|^n.$$

For $n \geq 2$,

$$\lim_{t \to 0} q(t) = 0, \;\; \text{and} \;\; v \mapsto df(0;v) = 0$$

is linear, and f is Gateaux differentiable at 0. For $n = 1$, the limit of the differential quotient does not exist since it goes to different values as t goes to 0 by positive or negative values:

$$\lim_{\substack{t>0 \\ t \to 0}} q(t) = |v|, \;\; \lim_{\substack{t<0 \\ t \to 0}} q(t) = -|v|.$$

Hence, for $n = 1$, f is not Gateaux differentiable at 0, and a fortiori not Fréchet differentiable at 0.

Since in dimension one the derivative and the Gateaux differential coincide, f is differentiable everywhere on $\mathbb{R}$ for $n \geq 2$, and

$$\frac{df}{dx}(x) = nx|x|^{n-2}, \;\; \forall x \in \mathbb{R}.$$

(ii) From part (i), for $n \geq 2$, for all $x \in \mathbb{R}$,

$$\frac{df}{dx}(x) = \left\{ \begin{matrix} nx^{n-1}, & x \geq 0 \\ -n(-x)^{n-1}, & x < 0 \end{matrix} \right\} = nx|x|^{n-2}.$$

For $n = 1$, the semidifferential at $x = 0$ is

$$df(0;v) = \lim_{\substack{t>0 \\ t \to 0}} \frac{f(tv) - f(0)}{t} = \lim_{\substack{t>0 \\ t \to 0}} \frac{|tv|}{t} = |v|$$

and

$$df(x;v) = \begin{cases} \dfrac{x}{|x|}v, & \text{if } x \neq 0 \\ |v|, & \text{if } x = 0. \end{cases}$$

(iii) For $n = 1$, $f(x) = |x|$. As f is not differentiable in 0, use the definition of a convex function:

$$\forall x, y \in \mathbb{R}, \forall \lambda \in [0,1], \quad |\lambda x + (1-\lambda)y| \leq |\lambda x| + |(1-\lambda)y| = \lambda|x| + (1-\lambda)|y|.$$

For $n = 2$,

$$f(x) = x^2, \ \frac{df}{dx}(x) = 2x \ \text{ and } \ \frac{d^2 f}{dx^2}(x) = 2 > 0,$$

f is of class $C^{(2)}$ and the second-order derivative is positive. Hence, it is convex. For $n \geq 3$, we have

$$\begin{aligned} x > 0, &\qquad \frac{d^2 f}{dx^2}(x) = n(n-1)x^{n-2} \\ x < 0, &\qquad \frac{d^2 f}{dx^2}(x) = n(n-1)(-x)^{n-2}. \end{aligned}$$

At $x = 0$ for $t > 0$,

$$\frac{df(0+tw;v) - df(0;v)}{t} = \frac{n\,tw\,|tw|^{n-2}}{t} = |t|^{n-2} n\,w\,|w|^{n-2} \to 0$$

and $(v,w) \mapsto d^2 f(0;v;w) = 0$ is indeed bilinear and $d^2 f/dx^2(0) = 0$. Therefore,

$$\forall n \geq 3, \forall x \in \mathbb{R}, \quad \frac{d^2 f}{dx^2}(x) = n(n-1)|x|^{n-2} \geq 0$$

and, from the previous argument, this formula is also valid for $n = 2$. For $n \geq 2$, the function f is of class $C^{(2)}$ in $\mathbb{R}$ and, as the second-order derivative is positive, f is convex on $\mathbb{R}$.

The computation can be carried out by observing that the derivative $f'(x) = nx|x|^{n-2}$ of f is the product of two semidifferentiable functions: nx and $|x|^{n-2}$. Indeed

$$\begin{aligned} df'(x;v) &= nv\,|x|^{n-2} + nx \begin{cases} (n-2)x^{n-3}v, & x > 0 \\ 0, & x = 0 \\ -(n-2)(-x)^{n-3}v, & x < 0 \end{cases} \\ &= nv\,|x|^{n-2} + n(n-2) \begin{cases} x^{n-2}v, & x > 0 \\ 0, & x = 0 \\ (-x)^{n-2}v, & x < 0 \end{cases} \\ &= nv\,|x|^{n-2} + n(n-2)|x|^{n-2}v = n(n-1)|x|^{n-2}v \end{aligned}$$

and, as it is linear with respect to v, f' is Gateaux differentiable, the second-order derivative exists, and $f''(x) = n(n-1)|x|^{n-2} \geq 0$. $\square$

Exercise 7.4

Given an integer $n \geq 1$, define the function $f : \mathbb{R}^n \to \mathbb{R}$:

$$f(x) \overset{\text{def}}{=} \sum_{i=1}^{n} |x_i|. \tag{2.4}$$

(i) Prove that the function f is convex and Lipschitzian on $\mathbb{R}^n$ and give the expression of its semidifferential $df(x;v)$.

(ii) For $n = 2$, compute $df(x;v)$ at the points $x = (1,1)$, $(0,1)$, $(1,0)$, $(0,0)$ as a function of $v = (v_1, v_2)$. In which cases is f Gateaux differentiable?

Answer. (i) The function f is the sum of n functions $f_i(x) = |x_i|$. Therefore, it is sufficient to prove that each function f_i is convex and Lipschitzian on $\mathbb{R}^n$ and that its semidifferential exists since the finite sum preserves the three properties. For the absolute value,

$$||y_i| - |x_i|| \leq |y_i - x_i| \leq \left\{ \sum_{i=1}^{n} |y_i - x_i|^2 \right\}^{1/2} = \|y - x\|$$

and f_i is Lipschitzian on $\mathbb{R}^n$; for all $\lambda \in [0,1]$ and $x, y \in \mathbb{R}^n$,

$$|(\lambda x + (1-\lambda)y)_i| \leq |\lambda x_i| + |(1-\lambda)y_i| = \lambda\,|x_i| + (1-\lambda)|y_i|$$

and f_i is convex. For the semidifferential, consider the differential quotient with $t > 0$, x and v:

$$\frac{|x_i + tv_i| - |x_i|}{t} \to \begin{cases} \dfrac{x_i}{|x_i|} v_i, & \text{if } x_i \neq 0 \\ |v_i|, & \text{if } x_i = 0 \end{cases}$$

$$\Rightarrow df(x;v) = \sum_{i=1}^{n} \begin{cases} \dfrac{x_i}{|x_i|} v_i, & \text{if } x_i \neq 0 \\ |v_i|, & \text{if } x_i = 0. \end{cases}$$

(ii) For $n = 2$,

$$df(x;v) = \begin{cases} \dfrac{x_1}{|x_1|} v_1, & \text{if } x_1 \neq 0 \\ |v_1|, & \text{if } x_1 = 0 \end{cases} + \begin{cases} \dfrac{x_2}{|x_2|} v_2, & \text{if } x_2 \neq 0 \\ |v_2|, & \text{if } x_2 = 0. \end{cases}$$

In particular,

$$df((1,1);(v_1,v_2)) = v_1 + v_2, \quad df((0,1);(v_1,v_2)) = |v_1| + v_2,$$
$$df((1,0);(v_1,v_2)) = v_1 + |v_2|, \quad df((0,0);(v_1,v_2)) = |v_1| + |v_2|.$$

(iii) From part (i), f is Gateaux differentiable at $x = (x_1, \ldots, x_n)$ if and only if $x_i \neq 0$ for all i. □

Exercise 7.5

Show that the function $f(x) = \sin x + (1+x)^2$ is convex on the interval $[0,1]$.

Answer. This function is of class $C^{(2)}$ on $\mathbb{R}$ and for all $x \in \mathbb{R}$,

$$\frac{df}{dx}(x) = \cos x + 2(1+x) \quad \frac{d^2 f}{dx^2}(x) = -\sin x + 2 \geq 1.$$

So, it is convex on $\mathbb{R}$ and its restriction to $[0,1]$ is also convex. □

3 Exercises of Chapter 4

Exercise 9.1

Show that the matrix (2.9) of Chapter 4 is positive definite.

Answer. The objective function is

$$J(\vec{p}) \stackrel{\text{def}}{=} \frac{1}{2}\int_0^1 \left(\sum_{i=0}^{n} p_i\, \phi_i(x) - g(x)\right)^2 dx.$$

The semidifferential is

$$dJ(\vec{p};\vec{v}) = \int_0^1 \left(\sum_{i=0}^{n} p_i\, \phi_i(x) - g(x)\right)\left(\sum_{j=0}^{n} \phi_j(x)\, v_j\right) dx$$

and the second-order semidifferential is

$$A\vec{w}\cdot\vec{v} = d^2 J(\vec{p};\vec{v};\vec{w}) = \int_0^1 \left(\sum_{i=0}^{n} w_i\, \phi_i(x)\right)\left(\sum_{j=0}^{n} v_j\, \phi_j(x)\right) dx.$$

It is easy to check that

$$\nabla J(\vec{p}) = A\vec{p} - \vec{g} \text{ and } H(\vec{p}) = A,$$

where the matrix A is of dimension $n+1$ and tridiagonal (that is, the elements of A are nonzero on the diagonal and on the two adjacent diagonals)

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & 0 & 0 & 0 & 0 \\ a_{1,0} & a_{1,1} & a_{1,2} & 0 & \ddots & 0 \\ 0 & a_{2,1} & a_{2,2} & a_{2,3} & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \dots & 0 & a_{n,n-1} & a_{n,n} \end{bmatrix} \quad \vec{g} = \begin{bmatrix} \displaystyle\int_{x_0}^{x_1} g(x)\phi_0(x)\,dx \\ \vdots \\ \displaystyle\int_{x_{i-1}}^{x_{i+1}} g(x)\phi_i(x)\,dx \\ \vdots \\ \displaystyle\int_{x_{n-1}}^{x_n} g(x)\phi_n(x)\,dx \end{bmatrix}.$$

Coming back to the second-order semidifferential with $w = v$,

$$A\vec{v}\cdot\vec{v} = d^2J(\vec{p};\vec{v};\vec{v}) = \int_0^1 \left(\sum_{j=0}^{n} v_j\,\phi_j(x)\right)^2 dx \geq 0$$

and the matrix A is positive semidefinite on $\mathbb{R}^n$. If $A\vec{v}\cdot\vec{v} = 0$, then for all $x \in [0,1]$,

$$\sum_{j=0}^{n} v_j\,\phi_j(x) = 0 \quad \text{and} \quad \forall i,\, v_i = \sum_{j=0}^{n} v_j\,\phi_j(x_i) = 0$$

by choice of the basis functions $\{\phi_i\}$, and $\vec{v} = 0$. Therefore, A is positive definite. This approach is much simpler than attempting to prove the positive definiteness of A directly. □

Exercise 9.2

Consider the function

$$f(x_1,x_2) = (x_1+1)^2 + (x_2-2)^2.$$

Find the minimizers of f with respect to the set

$$U = \{(x_1,x_2) : x_1 \geq 0 \text{ and } x_2 \geq 0\}.$$

Justify.

Answer. f is a strictly convex function since its Hessian matrix is positive definite:

$$\frac{1}{2}\nabla f(x_1,x_2) = \begin{bmatrix} x_1+1 \\ x_2-2 \end{bmatrix}, \quad \frac{1}{2}Hf(x_1,x_2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} > 0.$$

The set U is a closed convex cone in 0. Hence, there exists a unique minimizer since f is continuous (polynomial) and has the growth property at infinity and it is strictly convex on the closed convex set U.

In the convex case, we know that the necessary and sufficient condition for the existence of a minimizer is given by Corollary 1(iii) to Theorem 3.2 of Chapter 4:

$$x \in U, \quad \nabla f(x)\cdot x = 0 \text{ and for all } y \in U,\, \nabla f(x)\cdot y \geq 0 \tag{3.1}$$

that yields

$$x_1 \geq 0, \quad x_2 \geq 0, \quad (x_1+1)x_1 + (x_2-2)x_2 = 0 \tag{3.2}$$

$$\forall y_1 \geq 0,\, y_2 \geq 0, \quad (x_1+1)y_1 + (x_2-2)y_2 \geq 0. \tag{3.3}$$

By letting $y_1 = 0$ and $y_2 \geq 0$, and then $y_1 \geq 0$ and $y_2 = 0$, we get

$$\begin{cases} x_1 \geq 0 & x_1+1 \geq 0 \quad (x_1+1)x_1 = 0 \\ x_2 \geq 0 & x_2-2 \geq 0 \quad (x_2-2)x_2 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x_1 \geq 0 & x_1+1 \geq 0 \quad x_1 = -1 \text{ or } x_1 = 0 \\ x_2 \geq 0 & x_2-2 \geq 0 \quad x_2 = 2 \text{ or } x_2 = 0. \end{cases}$$

The solutions $x_1 = -1$ and $x_2 = 0$ violate one of the two inequalities. The only candidate left is $(x_1,x_2) = (0,2)$ that satisfies all inequalities. □

Exercise 9.3

Consider the function

$$f(x_1,x_2) \stackrel{\text{def}}{=} \frac{1}{2}\left[(x_1-1)^2+(x_2+2)^2\right]+\frac{1}{2}x_1x_2 \tag{3.4}$$

to minimize with respect to the set

$$U \stackrel{\text{def}}{=} \{(x_1,x_2) : x_1 \geq 0 \text{ and } x_2 \geq 0\}. \tag{3.5}$$

Find the minimizers. Justify.

Answer. f is a strictly convex function since its Hessian matrix is positive definite:

$$\frac{1}{2}\nabla f(x_1,x_2) = \begin{bmatrix} x_1-1+\frac{1}{2}x_2 \\ x_2+2+\frac{1}{2}x_1 \end{bmatrix} \quad \text{and} \quad \frac{1}{2}Hf(x_1,x_2) = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} > 0.$$

The set U is a closed convex cone in 0. As a result, there exists a unique minimizer since f is continuous (polynomial), has the growth property at infinity, and is strictly convex on the closed cone U. In the convex case, we know that the necessary and sufficient condition for the existence of a minimizer is (see Corollary 1(iii) to Theorem 3.2 in Chapter 4)

$$\exists x \in U, \quad \nabla f(x)\cdot x = 0 \text{ and for all } y \in U,\, \nabla f(x)\cdot y \geq 0 \tag{3.6}$$

that yields

$$x_1 \geq 0, \quad x_2 \geq 0, \quad \left(x_1-1+\frac{1}{2}x_2\right)x_1+\left(x_2+2+\frac{1}{2}x_1\right)x_2 = 0 \tag{3.7}$$

$$\forall y_1 \geq 0,\, y_2 \geq 0, \quad \left(x_1-1+\frac{1}{2}x_2\right)y_1+\left(x_2+2+\frac{1}{2}x_1\right)y_2 \geq 0. \tag{3.8}$$

By choosing $y_1 = 0$ and $y_2 \geq 0$, and then $y_1 \geq 0$ and $y_2 = 0$, we get

$$\begin{cases} x_1 \geq 0 & x_1-1+\frac{1}{2}x_2 \geq 0 & \left(x_1-1+\frac{1}{2}x_2\right)x_1 = 0 \\ x_2 \geq 0 & x_2+2+\frac{1}{2}x_1 \geq 0 & \left(x_2+2+\frac{1}{2}x_1\right)x_2 = 0 \end{cases} \tag{3.9}$$

$$\Rightarrow \begin{cases} x_1 \geq 0 & x_1-1+\frac{1}{2}x_2 \geq 0 & x_1+\frac{1}{2}x_2 = 1 \text{ or } x_1 = 0 \\ x_2 \geq 0 & x_2+2+\frac{1}{2}x_1 \geq 0 & x_2+\frac{1}{2}x_1 = -2 \text{ or } x_2 = 0. \end{cases} \tag{3.10}$$

The solution $(x_1,x_2) = (0,0)$ violates the second inequality of the first line. The solution $(x_1,x_2) = (0,-2)$ also violates the second inequality of the first line. Therefore, $x_1 > 0$ and $x_1+\frac{1}{2}x_2 = 1$. If $x_2 > 0$, then we get the system of equations of the unconstrained case

$$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{8}{3} \\ -2-\frac{4}{3} \end{bmatrix},$$

where x_2 violates the condition $x_2 \geq 0$. It only remains

$$x_1 - 1 + \frac{1}{2}x_2 = 0 \text{ and } x_2 = 0 \quad \Rightarrow (x_1, x_2) = (1, 0)$$

that satifies all inequalities. $\square$

Exercise 9.4

Let $P^1[0,1]$ be the set of polynomials from $[0,1]$ to $\mathbb{R}$ of degree less than or equal to 1; that is,

$$p(x) = ax + b, \quad \forall a \in \mathbb{R}, \quad \forall b \in \mathbb{R}.$$

Consider the function

$$f(p) \stackrel{\text{def}}{=} \int_0^1 \left(p(x) - (2x-1)^3\right)^2 dx.$$

Characterize and find the polynomials that minimize f with respect to the set

$$U \stackrel{\text{def}}{=} \{p \in P^1[0,1] : \forall x \in [0,1],\ p(x) \geq 0\}.$$

Hint: prove that $b_0(x) = 1 - x$ and $b_1(x) = x$ are basis functions for $P^1[0,1]$ and that

$$U = \{p_0 b_0 + p_1 b_1 : \forall p_0 \geq 0 \text{ and } \forall p_1 \geq 0\}.$$

Then prove that U is a convex cone in 0.

Answer. (i) Consider the mapping

$$(p_0, p_1) \mapsto \Lambda(p_0, p_1) \stackrel{\text{def}}{=} p_0 b_0 + p_1 b_1 : \mathbb{R}^2 \to P^1[0,1].$$

It is well defined since $(p_0 b_0 + p_1 b_1)(x) = p_0(1-x) + p_1 x$ is an element of $P^1[0,1]$. Conversely, the function $p \mapsto (p(0), p(1)) : P^1[0,1] \to \mathbb{R}^2$ is well defined. Each $p \in P^1[0,1]$ is of the form $p(x) = a + bx$, $p(0) = a$, and $p(1) = a + b$. Hence, $p(x) = p(0) + (p(1) - p(0))x = p(0)(1-x) + p(1)x$ and

$$\Lambda(p(0), p(1)) \stackrel{\text{def}}{=} p(0) b_0 + p(1) b_1 = p.$$

The mapping Λ is surjective. Finally, it is also injective. Since Λ is linear, it is sufficient to prove that $\Lambda p = 0 \Rightarrow p = 0$. In fact, this implies $(p(0), p(1)) = (0,0)$ and, since $p(x) = p(0)(1-x) + p(1)x$, we get $p = 0$.

Define the new function and the new set

$$(p_0, p_1) \mapsto \vec{f}(p_0, p_1) \stackrel{\text{def}}{=} f(\Lambda(p_0, p_1)) : \mathbb{R}^2 \to \mathbb{R} \quad \text{and} \quad \hat{U} \stackrel{\text{def}}{=} \Lambda^{-1} U.$$

It is readily seen that $\hat{U} = C \stackrel{\text{def}}{=} \{(p_0, p_1) : p_0 \geq 0,\ p_1 \geq 0\}$. In fact, by definition of Λ, on a $\Lambda^{-1} U \subset C$. Conversely, for $(p_0, p_1) \in C$, $\Lambda(p_0, p_1)(x) = p_0(1-x) + p_1 x \geq 0$ and we have $\Lambda C \subset U$.

The objective function becomes

$$\begin{aligned}
\hat{f}(p_0,p_1) &= f(p_0\, b_0 + p_1\, b_1) \\
&= \int_0^1 [p_0\, b_0(x) + p_1\, b_1(x) - (2x-1)^3]^2\, dx \\
&= \int_0^1 \{p_0^2\, b_0(x)^2 + p_1^2\, b_1(x)^2 + 2\, p_0\, p_1\, b_0(x)\, b_1(x) \\
&\qquad - (2x-1)^6 - 2(2x-1)^3(p_0\, b_0(x) + p_1\, b_1(x))\}\, dx
\end{aligned}$$

$$\begin{aligned}
\hat{f}(p_0,p_1) = p_0^2 \int_0^1 b_0(x)^2\, dx + p_1^2 \int_0^1 b_1(x)^2\, dx + 2p_0\, p_1 \int_0^1 b_0(x)\, b_1(x)\, dx \\
- 2p_0 \int_0^1 (2x-1)^3 b_0(x)\, dx - 2\, p_1 \int_0^1 (2x-1)^3 b_1(x)\, dx \\
- \int_0^1 (2x-1)^6\, dx.
\end{aligned}$$

It is a quadratic function in the variable (a,b) that can be written in terms of a symmetrical matrix A, a vector $d \in \mathbb{R}^2$, and a constant c:

$$\hat{f}(p_0,p_1) = \left(A \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} + d\right) \cdot \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} + c,$$

where

$$A = \begin{bmatrix} \displaystyle\int_0^1 b_0(x)^2\, dx, & \displaystyle\int_0^1 b_0(x) b_1(x)\, dx, \\ \displaystyle\int_0^1 b_0(x) b_1(x)\, dx, & \displaystyle\int_0^1 b_1(x)^2\, dx, \end{bmatrix}, \quad d = \begin{bmatrix} \displaystyle -2 \int_0^1 (2x-1)^3 b_0(x)\, dx \\ \displaystyle -2 \int_0^1 (2x-1)^3 b_1(x)\, dx \end{bmatrix}$$

$$c = -\int_0^1 (2x-1)^6\, dx.$$

The computation of A and d yields

$$A = \frac{1}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad d = \frac{1}{5}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Hence, the problem reduces to

$$\inf_{(p_0,p_1)\in \hat{U}} \hat{f}(p_0,p_1), \quad \hat{f}(p_0,p_1) = \left(A \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} + d\right) \cdot \begin{bmatrix} p_0 \\ p_1 \end{bmatrix},$$

where the constant c can be dropped.

The function $\hat{f}$ is of class $C^{(2)}$ in $\mathbb{R}^2$ and

$$\begin{aligned}
\nabla \hat{f}(p_0,p_1) &= 2A(p_0,p_1) + d = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} + \frac{1}{5}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
H\hat{f}(p_0,p_1) &= \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.
\end{aligned}$$

It is readily seen that the matrix $H\hat{f}$ is positive definite on $\mathbb{R}^2$.

(ii) The initial problem reduces to

$$\inf_{(a,b)\in C}\left(A\begin{bmatrix}a\\b\end{bmatrix}+d\right)\cdot\begin{bmatrix}a\\b\end{bmatrix}.$$

The function f is continuous and has the growth property at infinity and C is closed. Hence, there exists a minimizer in C and since $\hat{f}$ is strictly convex, the minimizer is unique.

As we are in the convex differentiable case, a minimizer is completely characterized by the following system:

$$a\geq 0,\quad b\geq 0,\quad \left(2A\begin{bmatrix}a\\b\end{bmatrix}+d\right)\cdot\begin{bmatrix}a\\b\end{bmatrix}=0 \tag{3.11}$$

$$\forall a'\geq 0, b'\geq 0,\quad \left(2A\begin{bmatrix}a\\b\end{bmatrix}+d\right)\cdot\begin{bmatrix}a'\\b'\end{bmatrix}\geq 0. \tag{3.12}$$

By substituting for A and b,

$$a\geq 0,\quad b\geq 0,\quad \left(\frac{1}{3}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}-\frac{1}{5}\begin{bmatrix}1\\-1\end{bmatrix}\right)\cdot\begin{bmatrix}a\\b\end{bmatrix}=0 \tag{3.13}$$

$$\forall a'\geq 0, b'\geq 0,\quad \left(\frac{1}{3}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}-\frac{1}{5}\begin{bmatrix}1\\-1\end{bmatrix}\right)\cdot\begin{bmatrix}a'\\b'\end{bmatrix}\geq 0. \tag{3.14}$$

Finally,

$$a\geq 0,\quad \left[\frac{1}{3}(2a+b)+\frac{1}{5}\right]a=0,\quad \frac{1}{3}(2a+b)+\frac{1}{5}\geq 0 \tag{3.15}$$

$$b\geq 0,\quad \left[\frac{1}{3}(a+2b)-\frac{1}{5}\right]b=0,\quad \frac{1}{3}(a+2b)-\frac{1}{5}\geq 0. \tag{3.16}$$

There are four possibilities. The case $a>0$ and $b>0$ is not acceptable since we have already seen that $a=-3/5$ and $b=3/5$. If $b=0$, then from the inequality

$$\frac{1}{3}(a+2b)-\frac{1}{5}\geq 0 \quad\Rightarrow a\geq\frac{3}{5}>0 \quad\Rightarrow \frac{1}{3}(2a+b)+\frac{1}{5}=0 \quad\Rightarrow a=-\frac{3}{10}\not\geq 0.$$

What is left is $b>0$ and $a=0$, for which

$$\frac{1}{3}(a+2b)=\frac{1}{3}(2b)=\frac{1}{5}\quad\Rightarrow b=\frac{3}{10}.$$

So, the minimizing polynomial is

$$p(x)=\frac{3}{10}x.$$

□

Exercise 9.5

Compute the dual cones for the following subsets of $\mathbb{R}^2$.

(i) $U_1 = \{(x,y) \in \mathbb{R}^2 : \; x = 0 \text{ and } y \geq 0\}$

(ii) $U_2 = \{(x,y) \in \mathbb{R}^2 : \; x + 5y \geq 0 \text{ and } 2x + 3y \leq 0\}$

(iii) $U_3 = \{(x,y) \in \mathbb{R}^2 : \; 2x + y \geq 0 \text{ and } 10x + 5y \leq 0\}$

(iv) $U_4 = \{(x,y) \in \mathbb{R}^2 : \; x + y = 0 \text{ and } 2x + 3y \geq 0\}$

(v) $U_5 = \{(1,2), (3,1)\}$.

Answer. (i) It is useful to put U_1 in the following generic form:

$$\begin{aligned} U_1 &= \{(x,y) \in \mathbb{R}^2 \, : x = 0 \text{ and } \; y \geq 0\} \\ &= \left\{(x,y) \in \mathbb{R}^2 : \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0 \text{ and } \begin{bmatrix} 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \leq 0 \right\}. \end{aligned}$$

By Theorem 7.2 of Chapter 4,

$$U_1^* = \left\{\lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} : \lambda_1 \in \mathbb{R} \text{ and } \lambda_2 \geq 0 \right\}.$$

To complete the exercise, express U_1^* in the form of equality and inequality constraints by introducing the variable

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \overset{\text{def}}{=} \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \quad \Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \begin{matrix} \in \mathbb{R} \\ \geq 0 \end{matrix}$$

that yields

$$U_1^* = \{(y_1, y_2) : y_1 \in \mathbb{R}, \; y_2 \geq 0\}.$$

(ii) Express U_2 in the form

$$U_2 = \left\{(x,y) \in \mathbb{R}^2 \, : \begin{bmatrix} 1 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \text{ and } \begin{bmatrix} -2 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \right\}.$$

By Theorem 7.2 of Chapter 4,

$$U_2^* = \left\{\lambda_1 \begin{bmatrix} 1 \\ 5 \end{bmatrix} + \lambda_2 \begin{bmatrix} -2 \\ -3 \end{bmatrix} : \forall \lambda_1 \geq 0 \text{ and } \lambda_2 \geq 0 \right\}.$$

This characterization can be expressed in terms of the following new variables:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} 1 & -2 \\ 5 & -3 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \quad \Rightarrow \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} -3 & 2 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -3y_1 + 2y_2 \\ -5y_1 + y_2 \end{bmatrix} \begin{matrix} \geq 0 \\ \geq 0 \end{matrix}$$

$$U_2^* = \{(y_1, y_2) : -3y_1 + 2y_2 \geq 0 \text{ and } -5y_1 + y_2 \geq 0\}.$$

(iii) $U_3 = \{(x,y) \in \mathbb{R}^2 : 2x+y \geq 0 \text{ and } 10x+5y \leq 0\}$

$$\begin{gathered}(1) \quad 2x+y \geq 0 \text{ and } 10x+5y \leq 0 \implies (2) \quad 2x+y \leq 0\\ (1) \text{ and } (2) \qquad 2x+y=0\\ \Rightarrow U_3 = \left\{(x,y) \in \mathbb{R}^2 : \begin{bmatrix}2\\1\end{bmatrix} \cdot \begin{bmatrix}x\\y\end{bmatrix} = 0\right\}.\end{gathered}$$

By Theorem 7.2 of Chapter 4,

$$U_3^* = \left\{\lambda \begin{bmatrix}2\\1\end{bmatrix} : \forall\lambda \in \mathbb{R}\right\} = \{(y_1,y_2) : y_1 - 2y_2 = 0\}.$$

(iv) By Theorem 7.2 of Chapter 4,

$$\begin{aligned}U_4 &= \left\{(x,y) \in \mathbb{R}^2 : x+y=0 \text{ and } 2x+3y \geq 0\right\}\\ &= \left\{(x,y) \in \mathbb{R}^2 : \begin{bmatrix}1\\1\end{bmatrix} \cdot \begin{bmatrix}x\\y\end{bmatrix} = 0 \text{ and } \begin{bmatrix}2\\3\end{bmatrix} \cdot \begin{bmatrix}x\\y\end{bmatrix} \geq 0\right\}\\ U_4^* &= \left\{\lambda_1 \begin{bmatrix}1\\1\end{bmatrix} + \lambda_2 \begin{bmatrix}2\\3\end{bmatrix} : \lambda_1 \in \mathbb{R},\ \lambda_2 \geq 0\right\} = \left\{\begin{bmatrix}1 & 2\\1 & 3\end{bmatrix}\begin{bmatrix}\lambda_1\\\lambda_2\end{bmatrix} : \lambda_1 \in \mathbb{R},\ \lambda_2 \geq 0\right\}.\end{aligned}$$

Again introduce the new variables

$$\begin{bmatrix}y_1\\y_2\end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix}1 & 2\\1 & 3\end{bmatrix}\begin{bmatrix}\lambda_1\\\lambda_2\end{bmatrix} \quad \Rightarrow \begin{matrix}\mathbb{R} \ni\\ 0 \leq\end{matrix} \begin{bmatrix}\lambda_1\\\lambda_2\end{bmatrix} = \begin{bmatrix}3 & -2\\-1 & 1\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix} = \begin{bmatrix}3y_1 - 2y_2\\-y_1+y_2\end{bmatrix}.$$

As $3y_1 - 2y_2 = \lambda_1 \in \mathbb{R}$, this constraint is inactive. The only remaining constraint is $-y_1 + y_2 = \lambda_2 \geq 0$. As a result,

$$U_4^* = \{(y_1,y_2) : y_2 - y_1 \geq 0\}.$$

(v) $U_5 = \{(1,2),(3,1)\}$. By definition of U_5^*,

$$U_5^* = \{(y_1,y_2) : y_1 + 2y_2 \geq 0,\ \ 3y_1 + y_2 \geq 0\}. \qquad \square$$

Exercise 9.6

Prove the equivalence of the standard, canonical, and inequality forms of a linear program in Definition 7.1 of Chapter 4 by using slack and/or extra variables and by transforming inequalities into equalities and vice versa.

Answer. (i) (SLP) $\Rightarrow$ (CLP) The linear programming problem (LP) in *standard form* (SLP) is given as

$$\begin{aligned}&\text{minimize } -b \cdot x\\ &\text{subject to } Ax \leq c, \quad x \geq 0.\end{aligned} \tag{3.17}$$

Introduce the slack variable $y \in \mathbb{R}^m_+$ to transform the inequality $Ax \le c$ into an equality

$$\begin{aligned} &\text{minimize } -\begin{bmatrix} b \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \\ &\text{subject to } \begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = c, \quad \begin{bmatrix} x \\ y \end{bmatrix} \ge \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{3.18}$$

(ii) (CLP) $\Rightarrow$ (LPI) The linear programming problem (LP) in *canonical form* (CLP) is given as

$$\begin{aligned} &\text{minimize } -b \cdot x \\ &\text{subject to } Ax = c, \quad x \ge 0. \end{aligned} \tag{3.19}$$

Split $Ax = c$ into two inequalities $Ax \ge c$ and $-Ax \ge -c$ and add the inequalities $-x \le 0$ to the above two to get

$$\begin{aligned} &\text{minimize } -b \cdot x \\ &\text{subject to } \begin{bmatrix} A \\ -A \\ -I \end{bmatrix} x \le \begin{bmatrix} c \\ -c \\ 0 \end{bmatrix}. \end{aligned} \tag{3.20}$$

(iii) (LPI) $\Rightarrow$ (SLP) The linear programming problem (LP) in *inequality form* (LPI) is given as

$$\begin{aligned} &\text{minimize } -b \cdot x \\ &\text{subject to } Ax \le c. \end{aligned} \tag{3.21}$$

Any real number can be written as the difference of two positive numbers. So introduce the vectors $x^+ \in \mathbb{R}^n_+$ and $x^- \in \mathbb{R}^n_+$ and replace x by $x^+ - x^-$ to get

$$\begin{aligned} &\text{minimize } -\begin{bmatrix} b \\ -b \end{bmatrix} \cdot \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \\ &\text{subject to } \begin{bmatrix} A & -A \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \le c, \quad \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \ge \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{3.22}$$

□

Exercise 9.7

Let Q be an $n \times n$ symmetric matrix and B an $m \times n$ ($m \le n$) matrix. Let $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ be two vectors. Consider the function

$$f(x) = \frac{1}{2} Qx \cdot x - a \cdot x, \tag{3.23}$$

the set of constraints

$$U = \{x \in \mathbb{R}^n : Bx = b\}, \tag{3.24}$$

and the minimization problem

$$\inf \{f(x) : x \in U\}. \tag{3.25}$$

(i) Prove that the vector $\hat{x}$ in $\mathbb{R}^n$ is the unique solution of problem (3.25) if and only if there exists a vector $\hat{\lambda}$ in $\mathbb{R}^m$ such that

$$\begin{cases} Q\hat{x} + B^\top \hat{\lambda} = a \\ B\hat{x} = b \end{cases} \tag{3.26}$$

$$\forall x \in \operatorname{Ker} B,\ x \neq 0, \quad Qx \cdot x > 0. \tag{3.27}$$

(ii) Prove that if the rank of the matrix B is m and the matrix A is positive definite, then there always exists a unique solution to problem (3.25) for all a and b.

(iii) Assume that the rank of the matrix B is m and that Q is positive definite. Give the expression of the solution $\hat{x}$ as a function of the data Q, a, B, and b [show that $BQ^{-1}B^\top$ is invertible and eliminate $\hat{\lambda}$].

(iv) Assume that the rank of the matrix B is m and that Q is positive definite. Denote by $\hat{x}(b)$ the solution of (3.25) that corresponds to b and define the new function

$$g(b) = f(\hat{x}(b)).$$

Prove that

$$\nabla g(b) = -\hat{\lambda}(b),$$

where $\hat{\lambda}(b)$ is the vector $\hat{\lambda}$ of part (i) [compute $(\hat{x}(b), \hat{\lambda}(b))$ as a function of B, Q, b, and a and use the chain rule to get the semidifferential of the composition $f(\hat{x}(b))$].

Answer. (i) If $\hat{x} \in U$ is the unique minimizer of f in U, then

$$\forall x \in U, x \neq \hat{x} \quad f(x) > f(\hat{x}). \tag{3.28}$$

Since A is symmetrical,

$$\begin{aligned} \forall x \in U, \quad f(x) - f(\hat{x}) &= Q\frac{x+\hat{x}}{2} \cdot (x - \hat{x}) - a(x - \hat{x}) \\ &= (Q\hat{x} - a) \cdot (x - \hat{x}) + \frac{1}{2} Q(x - \hat{x}) \cdot (x - \hat{x}) \geq 0. \end{aligned}$$

Given $h \in \operatorname{Ker} B$ and $t > 0$,

$$B(\hat{x} \pm th) = B\hat{x} \pm t\,Bh = b \quad \Rightarrow \hat{x} \pm th \in U$$

and for all $t > 0$,

$$0 \leq (Q\hat{x} - a) \cdot (\pm th) + \frac{1}{2} Q(\pm th) \cdot (\pm th) = \pm t(Q\hat{x} - a) \cdot h + \frac{1}{2} t^2\, Qh \cdot h. \tag{3.29}$$

Dividing by $t > 0$,

$$0 \leq \pm(Q\hat{x} - a) \cdot h + \frac{1}{2} t\, Qh \cdot h \tag{3.30}$$

and letting t go to 0, we get the following orthogonality condition:

$$0 \leq \pm(Q\hat{x}-a)\cdot h \quad \Rightarrow \boxed{(Q\hat{x}-a)\cdot h = 0, \quad \forall h \in \operatorname{Ker} B.} \tag{3.31}$$

For $h \in \operatorname{Ker} B$, $h \neq 0$, and $t > 0$, $x = \hat{x} + th \neq \hat{x}$. Substituting back into (3.28),

$$\forall h \in \operatorname{Ker} B, h \neq 0, \quad 0 < f(x) - f(\hat{x}) = (Q\hat{x}-a)\cdot h + \frac{1}{2} Qh\cdot h \tag{3.32}$$

$$\forall h \in \operatorname{Ker} B, h \neq 0, \quad 0 < \frac{1}{2} Qh\cdot h \tag{3.33}$$

that yields the strict positivity condition of A on $\operatorname{Ker} B$. Finally, condition (3.31) means that $Q\hat{x} - a \in [\operatorname{Ker} B]^{\perp} = \operatorname{Im} B^{\top}$. As a result, there exists $\lambda \in \mathbb{R}^m$ such that

$$Q\hat{x} - a + B^{\top}\lambda = 0 \quad \text{and} \quad B\hat{x} = b,$$

since $\hat{x} \in U$.

Conversely, let $\hat{x}$ and $\hat{\lambda}$ be solutions of system (3.26). From the second equation, $\hat{x} \in U$. We get by the previous computation: for all $x \in U$,

$$f(x) - f(\hat{x}) = (Q\hat{x}-a)\cdot(x-\hat{x}) + \frac{1}{2} Q(x-\hat{x})\cdot(x-\hat{x}).$$

But $x - \hat{x} \in \operatorname{Ker} B$ since $B(x-\hat{x}) = b - b = 0$. By the first condition (3.26), $Q\hat{x} - a = -B^{\top}\lambda$, we get $A(x-\hat{x})\cdot(x-\hat{x}) = -B^{\top}\lambda\cdot(x-\hat{x}) = -\lambda\cdot B(x-\hat{x}) = 0$ and

$$\begin{aligned} f(x) - f(\hat{x}) &= (Q\hat{x}-a)\cdot(x-\hat{x}) + \frac{1}{2} Q(x-\hat{x})\cdot(x-\hat{x}) \\ &= \frac{1}{2} Q(x-\hat{x})\cdot(x-\hat{x}). \end{aligned}$$

By condition (3.27), for $x \neq \hat{x}$, the second term is strictly positive and

$$f(x) - f(\hat{x}) = \frac{1}{2} Q(x-\hat{x})\cdot(x-\hat{x}) > 0.$$

Thus, the point $\hat{x} \in U$ is the unique minimizer of f in U.

(ii) The matrix $(n+m)\times(n+m)$

$$\mathcal{A} \overset{\text{def}}{=} \begin{bmatrix} Q & B^{\top} \\ B & 0 \end{bmatrix}$$

is symmetrical. It is inverstible if and only if it is injective, that is, if

$$\begin{bmatrix} Q & B^{\top} \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow x = 0 \text{ and } \lambda = 0.$$

Then, since $Bx = 0$, $x \in \operatorname{Ker} B$, and $Qx + B^{\top}\lambda = 0$,

$$0 = (Qx + B^{\top}\lambda)\cdot x = Qx\cdot x + B^{\top}\lambda\cdot x = Qx\cdot x + \lambda\cdot Bx = Qx\cdot x.$$

Since A is positive definite, there exists $\beta > 0$ such that

$$\forall x \in \mathbb{R}^n, \quad Qx \cdot x \geq \beta \|x\|^2,$$

and $x = 0$. Finally, the last remaining equation is $B^\top \lambda = 0$. Since B is surjective, its transpose is injective and $\lambda = 0$. This means that the minimization problem always has a unique solution in U whatever the vectors a and b are.

(iii) Start with the first equation and solve for $\hat{x}$. Then, substitute the expression of $\hat{x}$ in the second equation to obtain an equation for $\hat{\lambda}$:

$$\begin{aligned} Q\hat{x} + B^\top \hat{\lambda} = a \quad &\Rightarrow \hat{x} = Q^{-1}[a - B^\top \hat{\lambda}] \\ B\left[Q^{-1}[a - B^\top \hat{\lambda}]\right] = b \quad &\Rightarrow \left[B\,Q^{-1}B^\top\right]\hat{\lambda} = B\,Q^{-1}a - b. \end{aligned}$$

The new matrix $B\,Q^{-1}B^\top$ of dimension $m \times m$ is invertible if and only if it is injective, that is, $[B\,Q^{-1}B^\top]\hat{x} = 0$ implies $\hat{x} = 0$. It is readily checked that

$$[B\,Q^{-1}B^\top]\hat{x} = 0 \quad \Rightarrow [B\,Q^{-1}B^\top]\hat{x} \cdot \hat{x} = 0 \quad \Rightarrow Q^{-1}B^\top \hat{x} \cdot B^\top \hat{x} = 0.$$

As A is positive definite,

$$\begin{aligned} 0 = Q^{-1}B^\top \hat{x} \cdot B^\top \hat{x} = (Q^{-1}B^\top \hat{x}) \cdot Q(Q^{-1}B^\top \hat{x}) \geq \beta \|Q^{-1}B^\top \hat{x}\|^2 \\ \Rightarrow Q^{-1}B^\top \hat{x} = 0 \quad \Rightarrow B^\top \hat{x} = 0 \quad \Rightarrow \hat{x} = 0, \end{aligned}$$

since Q^{-1} is bijective and, since B is surjective, $B^\top$ is injective. So, finally

$$\begin{aligned} \hat{\lambda} &= \left[B\,Q^{-1}B^\top\right]^{-1}(B\,Q^{-1}a - b) \\ Q\hat{x} &= a - B^\top \hat{\lambda} = a - B^\top \left[B\,Q^{-1}B^\top\right]^{-1}(B\,Q^{-1}a - b) \\ \hat{x} &= Q^{-1}a - Q^{-1}B^\top \left[B\,Q^{-1}B^\top\right]^{-1}(B\,Q^{-1}a - b). \end{aligned}$$

(iv) We have seen in part (iv) that the minimizer is given by the expression

$$\begin{aligned} \hat{x}(b) &= Q^{-1}a - Q^{-1}B^\top \left[B\,Q^{-1}B^\top\right]^{-1}(B\,Q^{-1}a - b) \\ \hat{\lambda}(b) &= \left[B\,Q^{-1}B^\top\right]^{-1}(B\,Q^{-1}a - b), \end{aligned}$$

where the notation indicates that $(\hat{x}(b), \hat{\lambda}(b))$ is a function of b. Note that as $\hat{x}(b)$ is an affine function of b, it is Fréchet differentiable and its semidifferential at b in the direction b' is given by

$$d_H\hat{x}(b; b') = Q^{-1}B^\top \left[B\,Q^{-1}B^\top\right]^{-1} b'.$$

The inf $f(U)$ as a function of b is then given by the function

$$g(b) = f(\hat{x}(b)) = \frac{1}{2} Q\hat{x}(b) \cdot \hat{x}(b) - a \cdot \hat{x}(b).$$

The function f is quadratic and hence Fréchet differentiable. Its semidifferential at x in the direction x' is

$$d_H f(x;x') = [Qx - a] \cdot x'.$$

The semidifferential of the composition $g = f \circ \hat{x}$ at b in the direction b' is given by the chain rule applied to the semidifferentials of f and $\hat{x}$:

$$\begin{aligned} d_H g(b;b') &= d_H f(\hat{x}(b); d_H \hat{x}(b;b')) \\ &= \left[Q\hat{x}(b) - a\right] \cdot \left(Q^{-1}B^\top \left[B\,Q^{-1}B^\top\right]^{-1} b'\right) \\ &= \left(\left[B\,Q^{-1}B^\top\right]^{-1} B\,Q^{-1}\left[Q\hat{x}(b) - a\right]\right) \cdot b'. \end{aligned}$$

But, recall that, from the first equation of the optimality system, $Q\hat{x}(b) + B^\top\hat{\lambda}(b) = a$. By substituting, this gives

$$\begin{aligned} d_H g(b;b') &= \left(\left[B\,Q^{-1}B^\top\right]^{-1} B\,Q^{-1}\left[Q\hat{x}(b) - a\right]\right) \cdot b' \\ &= \left(\left[B\,Q^{-1}B^\top\right]^{-1} \left[-B\,Q^{-1}B^\top\hat{\lambda}(b)\right]\right) \cdot b' = -\hat{\lambda}(b) \cdot b' \\ &\Rightarrow \nabla g(b) = -\hat{\lambda}(b) \end{aligned}$$

since the semidifferential is linear with respect to the direction and hence Fréchet (Hadamard) differentiable. □

4 Exercises of Chapter 5

Exercise 5.1

Find the global minima of

$$f(x) = x_1 + (x_2 - 1)^2$$

with respect to the set

$$U = \{x = (x_1, x_2) \in \mathbb{R}^2 : g(x) = x_1^2 = 0\}$$

by using the Lagrange multipliers theorem. If this is not possible or if it provides no information, use the method of your choice.

Answer. The functions f and g are $C^{(1)}$ in $\mathbb{R}^2$ and

$$\nabla f(x) = (1, 2(x_2 - 1)), \quad \nabla g(x) = (2x_1, 0).$$

None of the points of U is a regular point of g:

$$\forall x = (0, x_2), \ \ \nabla g(0, x_2) = 0.$$

The Lagrange multipliers theorem yields no information.

However, substituting $x_1 = 0$ in f,

$$f(x) = (x_2 - 1)^2 \geq 0 \quad \Rightarrow \inf_{x \in U} f(x) = \inf_{x_2 \in \mathbb{R}} f(0, x_2) = 0$$

and $x_2 = 0$ or $x = (0, 1)$ is a global minimizer.

Since the set U is a linear subspace and the objective function is convex, we can use the necessary and sufficient condition of Corollary 1 to Theorem 3.2 of Chapter 4 for the existence of a minimizer:

$$\begin{aligned} &x \in U \text{ and } \forall y \in U,\ \nabla f(x) \cdot y = 0 \\ \Rightarrow\ &x = (0, x_2) \text{ and } \forall y_2 \in \mathbb{R},\ (x_2 - 1)\, y_2 = 0, \quad \Rightarrow x_2 = 1. \end{aligned}$$

□

Exercise 5.2

Find the maximum and the minimum of the function

$$f(x, y, z) = -4xy - z^2$$

with respect to the set

$$U = \left\{(x, y, z) : x^2 + y^2 + z^2 = 1\right\},$$

if they exist.

Answer. (i) This is a generic problem. Given a symmetrical $n \times n$ matrix A, consider the constrained minimization problem

$$\inf_{x \in U} f(x), \quad f(x) \stackrel{\text{def}}{=} \frac{1}{2} Ax \cdot x,\ \ U \stackrel{\text{def}}{=} \{x : \|x\|^2 = 1\}.$$

The function f is continuous on the compact unit sphere U. So, there exist global minimizers and maximizers by the Weierstrass theorem (Theorem 2.1) of Chapter 2.

Given the function $g(x) = \|x\|^2 - 1$, the constraint becomes $g(x) = 0$, that is, an equality constraint. Both f and g are of class $C^{(1)}$ and

$$\nabla g(x) = 2x \neq 0 \text{ for } x \in U$$

since $\|x\| = 1$. All points of U are regular points of g. The Lagrange multipliers theorem (Theorem 2.2 of Chapter 5) can be applied with $\lambda_0 = 1$: there exists $\lambda \in \mathbb{R}$ such that

$$\nabla f(x) + \lambda \nabla g(x) = 0 \quad \text{and} \quad g(x) = 0.$$

Upon substitution, the condition becomes

$$[A + 2\lambda I]x = Ax + 2\lambda x = 0 \quad \text{and} \quad \|x\|^2 = 1. \tag{4.1}$$

Condition (4.1) says that the multiplier λ is a solution of

$$\det [A + 2\lambda I] = 0.$$

This means that -2λ is one of the eigenvalues of the matrix A. Moreover, it is readily seen that

$$Ax + 2\lambda x = 0 \quad \Rightarrow f(x) = \frac{1}{2} Ax \cdot x = -\lambda x \cdot x = -\lambda \|x\|^2 = -\lambda \tag{4.2}$$

and the local minima are equal to $-\lambda$. So, it is not necessary to compute the eigenvectors $\vec{x}$ associated with the eigenvalues of the matrix $A + 2\lambda I$ to find the local extrema of f.

In our example, we get (beware of the multiplying factor 2 in specifying the matrix A!!)

$$\frac{1}{2} A = \begin{bmatrix} 0 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \text{ and } A + 2\lambda I = 2 \begin{bmatrix} \lambda & -2 & 0 \\ -2 & \lambda & 0 \\ 0 & 0 & \lambda - 1 \end{bmatrix}.$$

Dropping the 2, the condition becomes

$$0 = \det \begin{bmatrix} \lambda & -2 & 0 \\ -2 & \lambda & 0 \\ 0 & 0 & \lambda - 1 \end{bmatrix} = \lambda^2(\lambda - 1) - 4(\lambda - 1) = (\lambda - 1)(\lambda^2 - 4)$$
$$\Rightarrow \lambda = -2, 1, 2 \quad \Rightarrow f(\vec{x}) = 2, -1, \text{or } -2.$$

So, the minimum is -2 and the maximum $+2$ corresponding to $\lambda = 2$ and $\lambda = -2$ for which there exists at least one vector $\vec{x}$ with norm one. □

Exercise 5.3

Consider the parabola P and the line L:

$$P \stackrel{\text{def}}{=} \left\{ (x_1, y_1) : y_1 = 2x_1^2 \right\}, \quad L \stackrel{\text{def}}{=} \{(x_2, y_2) : y_2 = -1 + x_2\}.$$

(i) First prove the existence of a solution to the problem of finding the shortest distance between P and L.

(ii) Find the points of P and L that achieve that distance and compute that distance.

Answer. (i) The set of admissible points is

$$U \stackrel{\text{def}}{=} \left\{ (x_1, y_1, x_2, y_2) \in \mathbb{R}^4 : (x_1, y_1) \in P \text{ and } (x_2, y_2) \in L \right\} \tag{4.3}$$

and the objective function

$$f(x_1, y_1, x_2, y_2) \stackrel{\text{def}}{=} (x_2 - x_1)^2 + (y_2 - y_1)^2 \tag{4.4}$$

which is equal to the square of the distance between a point $(x_1, y_1) \in P$ and a point $(x_2, y_2) \in L$.

The function f is of class $C^{(\infty)}$. The difficulty in proving the existence of a minimizer is that the closed set U is not bounded and that the function f does not have the growth property at infinity. So, we prove that f has a bounded lower section on U:

$$\exists c \in \mathbb{R} \text{ such that } S \stackrel{\text{def}}{=} \{u \in U : f(u) \le c\} \tag{4.5}$$

be nonempty and bounded. Then, the existence of a minimizer follows from Theorem 5.3 of Chapter 2. Start with the point $(x_1, y_1, x_2, y_2) = (0,0,0,-1) \in U$ for which $f(0,0,0,-1) = 1$ and the lower section

$$S = \{(x_1, y_1, x_2, y_2) \in U : f(x_1, y_1, x_2, y_2) \leq 1\}. \tag{4.6}$$

By construction, $(0,0,0,-1) \in S$ and $S \neq \varnothing$. It remains to prove the boundedness. For any $(x_1, y_1, x_2, y_2 \in S)$,

$$\begin{aligned}
&(x_2 - x_1)^2 + (y_2 - y_1)^2 \leq 1, \quad y_1 = 2x_1^2 \text{ and } y_2 = x_2 - 1 \\
&\Rightarrow |x_1 + 2x_1^2 + 1| \leq |x_1 - x_2| + |2x_1^2 + 1 - x_2| = |x_1 - x_2| + |y_1 - y_2| \leq \sqrt{2} \\
&\Rightarrow 2\left[\left(x_1 + \frac{1}{4}\right)^2 + \frac{7}{16}\right] = |x_1 + 2x_1^2 + 1| =\leq \sqrt{2} \quad \Rightarrow \left(x_1 + \frac{1}{4}\right)^2 + \frac{7}{16} \leq \frac{\sqrt{2}}{2} \\
&\Rightarrow \left(x_1 + \frac{1}{4}\right)^2 \leq \frac{9}{16} \quad \Rightarrow \left|x_1 + \frac{1}{4}\right| \leq \frac{3}{4} \quad \Rightarrow |x_1| \leq 1 \text{ and } |y_1| = 2x_1^2 \leq 2.
\end{aligned}$$

Using the fact that the pairs (x_1, y_1) are bounded, we get

$$\left[x_2^2 + y_2^2\right]^{1/2} \leq \left[(x_2 - x_1)^2 + (y_2 - y_1)^2\right]^{1/2} + \left[x_1^2 + y_1^2\right]^{1/2} \leq 1 + \sqrt{1^2 + 2^2}$$

and the pairs of points (x_2, y_2) are also bounded. Thence, the lower section S is bounded and we have the existence of minimizers.

(ii) Introducing the constraint functions

$$g_1(x_1, y_1, x_2, y_2) \stackrel{\text{def}}{=} 2x_1^2 - y_1 \quad \text{and} \quad g_2(x_1, y_1, x_2, y_2) \stackrel{\text{def}}{=} y_2 - x_2 + 1, \tag{4.7}$$

the conditions of application of the Lagrange multipliers theorem are met: existence of minimizers and functions f, g_1, and g_2 of class $C^{(\infty)}$.

It is always a good idea to check if the points of U are regular points of the vector function $g = (g_1, g_2)$. The Jacobian matrix is

$$Dg(x) = \begin{bmatrix} 4x_1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

For any $y = (y_1, y_2) \in \mathbb{R}^2$, choose $z = (0, -y_1, 0, y_2)$ for which $Dg(x)z = (y_1, y_2)$. So all points of $\mathbb{R}^4$ are regular points of g.

Introduce the Lagrangian with $\lambda_0 = 1$:

$$L(u, \lambda) \stackrel{\text{def}}{=} f(u) + \lambda_1 g_1(u) + \lambda_2 g_2(u). \tag{4.8}$$

If $u \in U$ is a local minimizer, then there exists (λ_1, λ_2) such that

$$\begin{gathered}
\nabla f(u) + \lambda_1 \nabla g_1(u) + \lambda_2 \nabla g_2(u) = 0 \\
g_1(u) = 0 \text{ and } g_2(u) = 0.
\end{gathered}$$

This yields the following system of equations:

$$2(x_1 - x_2) + 4\lambda_1 x_1 = 0 \tag{4.9}$$
$$2(y_1 - y_2) - \lambda_1 = 0 \tag{4.10}$$
$$2(x_2 - x_1) - \lambda_2 = 0 \tag{4.11}$$
$$2(y_2 - y_1) + \lambda_2 = 0 \tag{4.12}$$
$$2x_1^2 - y_1 = 0 \tag{4.13}$$
$$y_2 - x_2 + 1 = 0. \tag{4.14}$$

From (4.9) and (4.10), and from (4.10)–(4.12), we get

$$2\lambda_1 x_1 = x_2 - x_1 = \frac{\lambda_2}{2} \tag{4.15}$$
$$\frac{\lambda_1}{2} = y_1 - y_2 = \frac{\lambda_2}{2} = x_2 - x_1 \quad \Rightarrow \boxed{\lambda_1 = \lambda_2.} \tag{4.16}$$

If $\lambda_1 = \lambda_2 = 0$, then from (4.15)–(4.16), $x_2 = x_1$ and $y_2 = y_1$. From (4.13) and (4.14), we get

$$2x_1^2 = y_1 \text{ and } x_1 - y_1 - 1 = 0 \quad \Rightarrow 2x_1^2 - x_1 + 1 = 0.$$

But this case cannot occur since the last equation has no real solution. Therefore, $\lambda_1 = \lambda_2 \neq 0$, and from (4.15),

$$2\lambda_1 x_1 = \frac{\lambda_2}{2} \quad \Rightarrow \boxed{x_1 = \frac{1}{4}} \text{ and } y_1 = 2x_1^2 = \frac{1}{8}.$$

From (4.16),

$$y_1 - y_2 = x_2 - x_1 \quad \Rightarrow \frac{1}{8} - y_2 = x_2 - \frac{1}{4} \quad \Rightarrow \boxed{y_2 + x_2 - \frac{3}{8} = 0.}$$

Finally, two equations remain

$$y_2 - x_2 + 1 = 0 \text{ and } y_2 + x_2 - \frac{3}{8} = 0 \quad \Rightarrow y_2 = -\frac{5}{16} \text{ and } x_2 = y_2 + 1 = \frac{11}{16}.$$

So, the minimizer is unique:

$$(x_1, y_1) = \left(\frac{1}{4}, \frac{1}{8}\right) \text{ and } (x_2, y_2) = \left(\frac{11}{16}, -\frac{5}{16}\right).$$

Compute the infimum

$$\begin{aligned} f(x_1, y_1, x_2, y_2) = (x_2 - x_1)^2 + (y_2 - y_1)^2 &= 2\,(x_2 - x_1)^2 \\ &= 2\left|\frac{11}{16} - \frac{4}{16}\right|^2 = 2\left|\frac{7}{16}\right|^2 \end{aligned}$$
$$\text{minimum distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \frac{7}{16}\sqrt{2}. \qquad \square$$

Exercise 5.4

(i) Find the local minima of the function

$$f(x) = (x_1 - 3)^2 + (x_2 - 4)^2$$

with respect to the constraint

$$U_0 = \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \le 1 \right\}.$$

(ii) Introduce the extra variable x_3 and transform the constraint U_0 into

$$U_1 = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1 \right\}.$$

Find the local minima of the function

$$f(x_1, x_2, x_3) = (x_1 - 3)^2 + (x_2 - 4)^2$$

with respect to U_1.

(iii) What can you conclude by comparing the results of (i) and (ii)?

Answer. (i) The function f is continuous since it is polynomial on $\mathbb{R}^2$. The set of constraints is closed and bounded. So, by the Weierstrass theorem (Theorem 2.1 of Chapter 2), there exist minimizers and maximizers of f in U_0.

Since f is Fréchet differentiable and the constraint function

$$g(x_1, x_2) \stackrel{\text{def}}{=} x_1^2 + x_2^2 - 1$$

is of class $C^{(1)}$, the minimizers satisfy the conditions of the Karush–Kuhn–Tucker theorem (Theorem 3.1 of Chapter 5). The constraint function is convex and Slater's condition of Theorem 3.2(iii) of Chapter 5 is verified at point $(0,0)$:

$$g(0,0) = -1 < 0.$$

We can take $\lambda_0 = 1$. This yields the following system of equalities and inequalities: there exists λ such that

$$\begin{aligned} &\lambda \ge 0, \quad x_1^2 + x_2^2 - 1 \le 0, \quad \lambda (x_1^2 + x_2^2 - 1) = 0 \\ &2(x_1 - 3) + 2\lambda x_1 = 0 \\ &2(x_2 - 4) + 2\lambda x_2 = 0. \end{aligned} \tag{4.17}$$

For $\lambda = 0$, the pair $(x_1, x_2) = (3,4)$ violates the constraint $x_1^2 + x_2^2 - 1 \le 0$. Therefore, $\lambda > 0$ and $x_1^2 + x_2^2 = 1$. We get

$$\begin{gathered} (1+\lambda)x_1 = 3 \quad \text{and} \quad (1+\lambda)x_2 = 4 \\ \Rightarrow (1+\lambda)^2 [x_1^2 + x_2^2] = 3^2 + 4^2 = 5^2 \quad \Rightarrow |1+\lambda| = 5. \end{gathered}$$

Since $\lambda > 0$, we have $1 + \lambda > 0$ and $\lambda = 4$. Hence, $(x_1, x_2) = (3/5, 4/5)$.

(ii) The function $f_1(x_1,x_2,x_3) = (x_1-3)^2 + (x_2-4)^2$ is continuous as a polynomial function on $\mathbb{R}^3$. The set U is the compact unit sphere in $\mathbb{R}^3$. By the Weierstrass theorem, there exist global minimizers and maximizers of f in U.

Since f_1 is Fréchet differentiable and the constraint function

$$g_1(x_1,x_2) \stackrel{\text{def}}{=} x_1^2 + x_2^2 + x_3^2 - 1$$

is of class $C^{(1)}$, the minimizers satisfy the condition of the Lagrange multipliers theorem. All points of U are regular points of g since

$$\forall x_1,x_2,x_3 \in U_1, \quad \nabla g_1(x_1,x_2,x_3) = 2\,(x_1,x_2,x_3) \neq (0,0,0).$$

So we can choose $\lambda_0 = 1$. This yields the following system of equations: there exists $\lambda \in \mathbb{R}$ such that

$$\begin{aligned} &x_1^2 + x_2^2 + x_3^2 - 1 = 1 \\ &2\,(x_1-3) + 2\,\lambda\,x_1 = 0 \\ &2\,(x_2-4) + 2\,\lambda\,x_2 = 0 \\ &2\,\lambda\,x_3 = 0. \end{aligned} \tag{4.18}$$

If we choose $\lambda = 0$, then $(x_1,x_2,x_3) = (3,4,x_3)$ violates the constraint. We conclude from this that $\lambda \neq 0$ and $x_3 = 0$. We get

$$\begin{gathered} (1+\lambda)\,x_1 = 3 \text{ and } (1+\lambda)\,x_2 = 4 \\ \Rightarrow\ (1+\lambda)^2\,[x_1^2 + x_2^2 + x_3^2] = 3^2 + 4^2 = 5^2 \quad \Rightarrow\ |1+\lambda| = 5. \end{gathered}$$

If $\lambda > 0$, we get $1+\lambda > 0$ and $\lambda = -6$. Hence $(x_1,x_2,x_3) = (3/5,4/5,0)$ and

$$\begin{aligned} f_1(3/5,4/5,0) &= (3/5-3)^2 + (4/5-4)^2 \\ &= (12/5)^2 + (16/5)^2 = (4/5)^2\,[3^2+4^2] = 16. \end{aligned}$$

If $\lambda < 0$, we get $1+\lambda > 0$ and $\lambda = 4$. Hence, $(x_1,x_2,x_3) = (-3/5,-4/5,0)$ and

$$\begin{aligned} f_1(3/5,4/5,0) &= (-3/5-3)^2 + (-4/5-4)^2 \\ &= (18/5)^2 + (24/5)^2 = (6/5)^2\,[3^2+4^2] = 36. \end{aligned}$$

Since we have the existence of global minimizers and maximizers, the first point is a minimizer and the second one a maximizer.

(iii) Observe that both problems give the same minimizer in $\mathbb{R}^2$. The variable x_3 is called a *slack variable*. When the constraint function is positive, its introduction changes the inequality constraint that requires the Karush–Kuhn–Tucker theorem into an equality constraint that requires the Lagrange multipliers theorem, but it introduces local maxima among the solutions. $\square$

Exercise 5.5

Find the local minima of the function

$$f(x) = 2x_1^2 + x_2^2 - 2x_1 + 4x_2, \ \ x = (x_1,x_2) \in \mathbb{R}^2$$

with respect to

$$U_0 = \mathbb{R}^2,$$
$$U_1 = \left\{(x_1,x_2) \in \mathbb{R}^2 \,:\, g(x_1,x_2) = 1 - x_1^2 + x_2 \le 0\right\},$$
$$U_2 = \{(x_1,x_2) \,|\, x_1 \ge 0,\ x_2 \ge 0\}.$$

Answer. (i) The function f is continuous since it is polynomial. It has the growth property at infinity:

$$f(x) = 2x_1^2 + x_2^2 - 2x_1 + 4x_2 \ge x_1^2 + x_2^2 - \left[\sqrt{2^2+4^2}\right]^{1/2} \left[\sqrt{x_1^2+x_2^2}\right]^{1/2}$$
$$\ge \|x\|^2 - 5\,\|x\| \to \infty$$

as $\|x\| \to \infty$. By Theorem 5.4 of Chapter 2, f has a bounde lower section. Since U_0, U_1, and U_2 are closed, there exist global minimizers by Theorem 5.3(i) of Chapter 2.

The function f is quadratic and infinitely differentiable:

$$\nabla f(x) = \begin{bmatrix} 4x_1 - 2 \\ 2x_2 + 4 \end{bmatrix} \quad \text{and} \quad Hf(x) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} > 0.$$

f is strictly convex in $\mathbb{R}^2$ since its Hessian matrix is positive definite.

By Theorem 2.2 of Chapter 4, we have the existence of a unique minimizer in $U_0 = \mathbb{R}^2$ that is a solution of

$$\nabla f(x) = 0 \quad \Rightarrow \begin{bmatrix} 4x_1 - 2 \\ 2x_2 + 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow x_1 = \frac{1}{2} \quad \text{and} \quad x_2 = -2.$$

For U_1, we have an inequality constraint and we apply the Karush–Kuhn–Tucker theorem (Theorem 3.1 of Chapter 5). There exist $\lambda_0 \ge 0$, and $\lambda \ge 0$, not both zero, such that

$$\lambda_0 \nabla f(x) + \lambda\, \nabla g(x) = 0, \quad g(x) \le 0, \quad \lambda\, g(x) = 0$$
$$\lambda_0 \begin{bmatrix} 4x_1 - 2 \\ 2x_2 + 4 \end{bmatrix} + \lambda \begin{bmatrix} -2x_1 \\ 1 \end{bmatrix} = 0, \quad 1 - x_1^2 + x_2 \le 0, \quad \lambda(1 - x_1^2 + x_2) = 0.$$

If $\lambda_0 = 0$, we get from the second line the contradiction $\lambda = 0$. So we can take $\lambda_0 = 1$. As a result,

$$x_2 = -\frac{1}{2}(\lambda + 4) \quad \text{and} \quad (2 - \lambda)x_1 = 1$$

that implies $x_2 \le -2$. The inequality

$$1 - x_1^2 + x_2 \le 1 - x_1^2 - 2 = -x_1^2 - 1 \le -1 \quad \text{and hence} \quad \lambda = 0$$

from the last equality. We get $x_1 = 1/2$ and $x_2 = -2$. So, the constraint is inactive.

As for the last case, U_2 is a convex cone at 0 and f is convex. So, we can apply Corollary 1(iii) to Theorem 2.2 of Chapter 4:

$$x \in U_2, \quad \nabla f(x) \cdot x = 0, \quad \forall y \in U_2, \nabla f(x) \cdot y \ge 0.$$

This translates into

$$x_1 \ge 0, x_2 \ge 0, \quad \begin{bmatrix} 4x_1 - 2 \\ 2x_2 + 4 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\forall y_1 \ge 0 \quad \text{and} \quad \forall y_2 \ge 0, \quad \begin{bmatrix} 4x_1 - 2 \\ 2x_2 + 4 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \ge 0.$$

Set $y_2 = 0$ and then $y_1 = 0$ to get

$$\begin{aligned} x_1 \ge 0, \quad 4x_1 - 2 \ge 0, \quad x_1(4x_1 - 2) = 0 \\ x_2 \ge 0, \quad 2x_2 + 4 \ge 0, \quad x_2(2x_2 + 4) = 0. \end{aligned}$$

If $x_1 = 0$, the second inequality in the first line is violated. Therefore, $4x_1 - 2 = 0$ and $x_1 = 2$. If $2x_2 + 4 = 0$, $x_2 < 0$, and the inequality $x_2 \ge 0$ is violated. Finally $x_2 = 0$. □

Exercise 5.6

Find the minimizers of the function

$$f(x) = |x|^5 - 5x$$

with respect to the set

$$U = \left\{ x \in \mathbb{R} : g(x) = x^2 - 8x + 15 \le 0 \right\}.$$

Answer. The function f is continuous and the set U is closed. Moreover,

$$x^2 - 8x + 15 = (x-4)^2 - 16 + 15 = (x-4)^2 - 1.$$

It is readily seen that $4 \in U$ and that U is bounded since for all $x \in U$,

$$(x-4)^2 - 1 = x^2 - 8x + 15 \le 0 \quad \Rightarrow |x-4|^2 \le 1 \quad \Rightarrow |x| \le 5.$$

U is compact and, by the Weierstrass theorem (Theorem 2.1 of Chapter 2), there exist global minimizers and maximizers.

We have seen in Exercise 7.3 of Chapter 3 that f is at least twice differentiable and that for $n = 5$,

$$\forall x \in \mathbb{R}, \quad \frac{df}{dx}(x) = 5x|x|^3, \quad \frac{d^2 f}{dx^2}(x) = 20|x|^2 \ge 0.$$

So, the function f is convex and of class $C^{(2)}$ on $\mathbb{R}$. The function g is convex and of class $C^{(\infty)}$ on $\mathbb{R}$. Therefore, the set U specified by the function g is convex and compact.

The Slater condition of Theorem 3.2(iii) of Chapter 5 is verified at the point $x = 4$ where $g(4) = -1 < 0$. So, we can take $\lambda_0 = 1$ in the Karush–Kuhn–Tucker theorem (Theorem 3.1 of Chapter 5).

We obtain the following system of equations: there exists λ such that

$$\lambda \ge 0, \quad (x-4)^2 - 1 \le 0, \quad \lambda \left[(x-4)^2 - 1 \right] = 0, \quad 5x|x|^3 + \lambda 2(x-4) = 0.$$

If $\lambda = 0$, then $5x|x|^3 = 0$, $x = 0$, and the condition $(x-4)^2 - 1 \le 0$ is violated. Therefore, $\lambda > 0$ and $(x-4)^2 - 1 = 0$ that yields either $x = 3$ or $x = 5$. For $x = 5$, the condition $5x|x|^3 + \lambda 2(x-4) = 0$ is violated since $\lambda > 0$ and $5x|x|^3 + \lambda 2 \ge 2\lambda > 0$. Thus, the minimizer is $x = 3$. □

Exercise 5.7

Let A be a positive definite symmetric 2×2 matrix, b a vector of $\mathbb{R}^2$ of norm 1, and $c \in \mathbb{R}$ a scalar. Define the ellipse E and the line L:

$$E \stackrel{\text{def}}{=} \left\{ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : Ax\cdot x \le 1 \right\}, \quad L \stackrel{\text{def}}{=} \left\{ y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : b\cdot y = c \right\}. \tag{4.19}$$

(i) Prove that there exist points $(x,y)\in E\times L$ that minimize the distance $\|x-y\|$ between E and L.

(ii) Prove that, if $E\cap L = \varnothing$, the minimizing points are given by the expressions

$$x = \frac{c}{|c|}\frac{1}{\sqrt{A^{-1}b\cdot b}}A^{-1}b \in E \quad \text{and} \quad y = x + \frac{c}{|c|}\left(|c| - \sqrt{A^{-1}b\cdot b}\right) b \in L,$$

that $|c| > \sqrt{A^{-1}b\cdot b}$, and that the minimum distance is given by

$$\|x-y\| = |c| - \sqrt{A^{-1}b\cdot b}.$$

Answer. (i) Choose as objective function the square of the distance between a point x of the ellipse E and a point y of the line L:

$$f(x,y) \stackrel{\text{def}}{=} \|x-y\|^2 = |x_1-y_1|^2 + |x_2-y_2|^2. \tag{4.20}$$

It is continuous. The set of constraints given by

$$U \stackrel{\text{def}}{=} \left\{ (x,y)\in\mathbb{R}^2\times\mathbb{R}^2 : g_1(x,y) \stackrel{\text{def}}{=} Ax\cdot x - 1 \le 0 \text{ and } g_2(x,y) \stackrel{\text{def}}{=} b\cdot y - c = 0 \right\}$$

is closed since it is defined in terms of continuous functions, but it is not bounded.

Fortunately, the function f has a bounded lower section with respect to U. Indeed, choose $f(0,cb) = c^2$ and the lower section $U_c = \{(x,y)\in U : f(x,y)\le c^2\} \neq \varnothing$. As A is positive definite, there exists $\alpha > 0$ such that $\alpha\,\|x\|^2 \le Ax\cdot x \le 1$. Thence, $\|y\| \le \|y-x\| + \|x\| \le |c| + 1/\sqrt{\alpha}$ and U_c is bounded. In conclusion, there are minimizers of f in U.

(ii) We now use Theorem 4.1 of Chapter 5 for the mixed case. There exist $\lambda_0 \ge 0$, $\lambda \ge 0$, and $\mu \in \mathbb{R}$, not all zero, such that

$$\begin{aligned} &2\lambda_0(x-y) + 2\lambda\, Ax = 0 \quad \text{and} \quad 2\lambda_0(y-x) + \mu b = 0 \\ &\lambda \ge 0,\ Ax\cdot x \le 1,\ \lambda(Ax\cdot x - 1) = 0 \quad \text{and} \quad b\cdot y = c. \end{aligned} \tag{4.21}$$

If $\lambda_0 = 0$, then

$$\begin{aligned} \lambda\, Ax = 0 \quad &\Rightarrow 0 = \lambda\, Ax = 0\cdot x = \lambda \\ \mu b = 0 \text{ and } \|b\| = 1 \quad &\Rightarrow \mu = 0 \end{aligned}$$

and we get a contradiction with the fact that at least one of the multipliers is not zero. So set $\lambda_0 = 1$.

If $\mu = 0$, then $x = y$ and we have an intersection $E \cap L \neq \varnothing$. If $\mu \neq 0$, then $x \neq y$, we have no intersection, and $2\lambda\, Ax + \mu\, b = 0$ implies that $\lambda > 0$ and

$$x = -\frac{\mu}{2\lambda} A^{-1}b, \quad Ax \cdot x = 1, \quad b \cdot y = c \quad \Rightarrow \left|\frac{\mu}{2\lambda}\right| = \frac{1}{\sqrt{A^{-1}b \cdot b}}.$$

We come back to (4.21)

$$\begin{array}{l} [2(x-y)+2\lambda\, Ax]\cdot x = 0 \\ [2(y-x)+\mu\, b]\cdot y = 0 \end{array} \quad \Rightarrow \begin{cases} 2\,\|x-y\|^2 + 2\lambda + \mu\, c = 0 \\ \dfrac{\|x-y\|^2}{2\lambda} + 1 + \dfrac{\mu}{2\lambda} c = 0 \\ \Rightarrow 1 < -\dfrac{\mu}{2\lambda} c \quad \Rightarrow c \neq 0 \end{cases}$$

since $\|x - y\| > 0$. In particular, we get the key condition

$$1 < -\frac{\mu}{2\lambda} c = \left|\frac{\mu}{2\lambda}\right| |c| \quad \Rightarrow \boxed{|c| > \sqrt{A^{-1}b \cdot b}}$$

$$c\,x = -c\frac{\mu}{2\lambda} A^{-1}b = |c| \left|\frac{\mu}{2\lambda}\right| A^{-1}b \quad \Rightarrow \boxed{x = \frac{c}{|c|} \frac{1}{\sqrt{A^{-1}b \cdot b}} A^{-1}b.}$$

Hence, as $|c| > \sqrt{A^{-1}b \cdot b}$, we get

$$\begin{aligned} y = x - \frac{\mu}{2} \quad &\Rightarrow c = b \cdot y = b \cdot x - \frac{\mu}{2} = \frac{c}{|c|}\sqrt{A^{-1}b \cdot b} - \frac{\mu}{2} \\ &\Rightarrow -\frac{\mu}{2} = c - \frac{c}{|c|}\sqrt{A^{-1}b \cdot b} = \frac{c}{|c|}\left[|c| - \sqrt{A^{-1}b \cdot b}\right] \\ &\Rightarrow \|y - x\| = \left|\frac{\mu}{2}\right| = |c| - \sqrt{A^{-1}b \cdot b} > 0 \\ y = \frac{c}{|c|}&\left[\frac{1}{\sqrt{A^{-1}b \cdot b}} A^{-1}b + \left(|c| - \sqrt{A^{-1}b \cdot b}\right) b\right]. \end{aligned}$$

If $|c| \leq \sqrt{A^{-1}b \cdot b}$, then $\mu = 0$, we have intersection, $E \cap L \neq \varnothing$, and $\|y - x\| = 0$. $\square$

Elements of Bibliography

I. ADLER

[1], *On the Equivalence of Linear Programming Problems and Zero-Sum Games*, http://www.optimization-online.org, June 2010.

P. ALART, O. MAISONNEUVE, and R. T. ROCKAFELLAR

[1], *Nonsmooth mechanics and analysis: Theoretical and numerical advances*, Springer-Verlag, Berlin, 2006.

J.-P. AUBIN and A. CELLINA

[1], *Differential inclusions*, Springer-Verlag, Berlin, 1984.

J.-P. AUBIN and H. FRANKOWSKA

[1], *Set-Valued Analysis*, Birkhäuser, Boston, 1990.

M. BARBUT, B. LOCKER, L. MAZILAK, and P. PRIOURET

[1], *Cinquante ans de correspondance mathématique en 107 lettres, Paul Lévy - Maurice Fréchet*, Birkhäuser, Boston, 1990.

A. BASTIANI

[1], *Applications différentiables et variétes diflérentiables de dimension infinie*, J. Analyse Math. **13** (1964), 1–114.

L. D. BERKOVITZ

[1], *Convexity and optimization in R^n*, Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2002.

G. BEVERIDGE and R. SCHECHTER

[1], *Optimization: Theory and practice*, McGraw–Hill Book Co., New York, 1970.

E. BLUM and W. OETTLI

[1], *Direct proof of the existence theorem for quadratic programming*, Operations Research **20**, No. 1 (1972), 165–167.

F. BONNANS

[1], *Optimisation continue*, Dunod, Paris, 2006.

F. BONNANS and A. SHAPIRO

[1], *Perturbation analysis of optimization problems*, Springer Series in Operations Research, Springer-Verlag, New York, 2000.

É. BOREL

[1], *Traité du calcul des probabilités et de ses applications. Applications aux jeux de hasard*, rédigé avec la collab. de Jean Ville, Gauthier-Villars, Paris, 1938.

J. M. BORWEIN and A. S. LEWIS

[1], *Convex analysis and nonlinear optimization. Theory and examples*, CMS (Canadian Mathematical Society) Books in Mathematics, Volume 3, Springer-Verlag, New York, 2000.

G. BOULIGAND

[1], *Sur les surfaces dépourvues de points hyperlimités*, Ann. Soc. Polon. Math. **9** (1930), 32–41.

[2], *Introduction à la géométrie infinitésimale directe*, Gauthier-Villars, Paris, 1932.

S. BOYD and L. VANDENBERGHE

[1], *Convex optimization*, Cambridge University Press, Cambridge, 2004.

P. CANNARSA and C. SINESTRARI

[1], *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Progress in Nonlinear Differential Equations and Their Applications, Volume 58, Birkhäuser Boston, Boston, 2004.

J. CARISTI

[1], *Fixed point theorems for mappings satisfying inwardness conditions*, Trans. Amer. Math. Soc. **215** (1976), 241–251.

J. CARISTI and W. A. KIRK

[1], *Geometric fixed point theory and inwardness conditions*, in "The geometry of metric and linear spaces (Proc. Conf., Michigan State Univ., East Lansing, Mich., 1974)", pp. 74–83, Lecture Notes in Math., Volume 490, Springer, Berlin, 1975.

F. H. CLARKE

[1], *Necessary Conditions for Nonsmooth Problems in Optimal Control and the Calculus of Variations*, Ph.D. thesis, Univ. of Washington, 1973.

[2], *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1983.

T. CLAUSEN

[1], *Über die Form Architektonischer Säulen*, in "Mélanges mathématiques et astronaumiques I" (1849–1853), 279–284.

L. COLLATZ and W. WETTERLING

[1], *Optimierungsaufgaben*, (German) Heidelberger Taschenbücher, Band 15, Springer-Verlag, Berlin-New York 1966 (second edition 1971), English translation by P. Wadsack. *Optimization problems*, Applied Mathematical Sciences, Vol. 17, Springer-Verlag, New York-Heidelberg, 1975.

S. J. COX

[1], *The shape of the ideal column*, Math. Intelligencer (1) **14** (1992), 16–24.

[2], *The generalized gradient at a multiple eigenvalue*, J. Funct. Anal. (1) **133** (1995), 30–40.

S. J. COX, B. KAWOHL and P. X. UHLIG

[1], *On the optimal insulation of conductors*, J. Optim. Theory Appl. (2) **100** (1999), 253–263.

S. J. COX and C. M. MCCARTHY

[1], *The shape of the tallest column*, SIAM J. Math. Anal. (3) **29** (1998), 547–554.

S. J. COX and M. L. OVERTON

[1], *On the optimal design of columns against buckling*, SIAM J. Math. Anal. (2) **23** (1992), 287–325.

G. B. DANTZIG

[1], *Programming in a linear structure*, U. S. Air Force Comptroller, USAF, Washington, D.C., 1948.

[2], *A Proof of the Equivalence of the Programming Problem and the Game Problem*, in "Activity Analysis of Production and Allocation," T. C. Koopmans (ed.), John Wiley & Sons, New York, 1951.

[3], *Maximization of a linear function subject to linear inequalities*, in "Activity Analysis of Production and Allocation," T. C. Koopmans (ed.), John Wiley & Sons, New York, 1951, pp. 339–347.

[4], *Linear Programming and Extensions*, Princeton University Press, Princeton, 1963.

M. C. DELFOUR and J.-P. ZOLÉSIO

[1], *Shapes and geometries: Metrics, analysis, differential calculus and optimization*, SIAM Series on Advances in Design and Control, SIAM, Philadelphia, PA, 2011, second edition.

V. F. DEMYANOV and A. M. RUBINOV

[1], *Constructive Nonsmooth Analysis*, Peter Lang Verlag, New York, 1995.

U. DINI

[1], *Fondamenti per la teorica delle funzioni di variabili reali*, Pisa T. Nistri, 1878 [Foundations of the theory of functions of a real variable] (German translation: *Grundlagen für eine Theorie der Funktionen einer veränderlichen reellen Grösse*, Teubner, 1892).

A. DONTCHEV and R. T. ROCKAFELLAR

[1], *Implicit functions and solution mappings. A view from variational analysis*, Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

A. JA. DUBOVITSKIĬ and A. A. MILJUTIN

[1], *Extremal problems with constraints*, (Russian) Ž. Vyčisl. Mat. i Mat. Fiz. **5** (1965), 395–453.

M. DUREA

[1], *On the existence and stability of approximate solutions of perturbed vector equilibrium problems*, Journal of Mathematical Analysis and Applications **333**, no. 2 (2007), 1165–1179.

I. EKELAND

[1], *On the variational principle*, Journal of Mathematical Analysis and Applications **47**, no. 2 (1974), 324–353.

[2], *Nonconvex minimization problems*, Bulletin of the American Mathematical Society **1** (1979), 443–474.

I. EKELAND and R. TEMAM

[1], *Analyse convexe et problèmes variationnels*, Dunod Gauthier-Villars, Paris, 1974. English translation: *Convex analysis and variational problems*, North–Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., New York, 1976 (Corrected reprinting, SIAM, Philadelphia, 1999).

L. EULER

[1], *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti*, 1744 [A method for discovering curved lines that enjoy a maximum or minimum property, or the solution of the isoperimetric problem taken in the widest sense], Opera Omnia: Series 1, Volume 24 (1952), 1–308.

J. FARKAS

[1], *Über die Theorie der Einfachen Ungleichungen*, Journal für die Reine und Angewandte Mathematik **124** (1902), 1–27.

W. FENCHEL

[1], *On conjugate convex functions*, Canad. J. Math. **1** (1949), 73–77.

[2], *Convex cones, sets and functions*, Mimeographed Lecture Notes, Princeton University, Princeton, NJ, 1951.

[3], *A remark on convex sets and polarity*, Medd. Lunds Univ. Mat. Sem. (Supplementband) (1952), 82–89.

[4], *Über konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten*, Math. Z. **63** (1956), 496–506.

W. FENCHEL and T. BONNESEN

[1], *Theorie der konvexen Körper*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Berlin: 1. Verlag von Julius Springer, 1931.

[2], *Theorie der konvexen Körper*, Chelsea Publishing Co., Bronx, New York, 1971.

[3], *Theorie der konvexen Körper*, Springer-Verlag, Berlin, 1974.

[4], *Theory of convex bodies*, L. Boron, C. Christenson and B. Smith (eds.), BCS Associates, Moscow, Idaho, 1987.

J. FERGUSON

[1], *A. Brief Survey of the History of the Calculus of Variations and Its Applications*, University of Victoria, Canada, 2004 (arXiv:math/0402357).

P. DE FERMAT

[1], *Methodus ad disquirendam Maximam et Minimam*, Varia opera mathematica, 1679. (First written in a letter to Mersenne in 1638, the first printed publication of the method is in the fifth volume of the *Supplementum Cursus Mathematici* (1642) written by Herigone, and only in 1679 it appears in Fermat's *Varia opera mathematica* as Methodus ad disquirendam maximam et minima, followed by De tangentibus linearum curvarum.) Electronic version in latin: http://fr.wikisource.org/wiki/Œuvres de Fermat - Livre I - Maxima et Minima.

[2], *De tangentibus linearum curvarum*, Varia opera mathematica, 1679.

W. H. FLEMING

[1], *Functions of several variables*, Addison–Wesley, Reading, Massachusetts, 1965.

M. FORTIN and Z. MGHAZLI

[1], *Analyse d'un élément mixte pour le problème de Stokes. I. Résultats généraux* (French) [Analysis of a mixed element for the Stokes problem. I. General results], Numer. Math. 62 (1992), 149–160.

[2], *Analyse d'un élément mixte pour le problème de Stokes. II. Construction et estimations d'erreur* (French) [Analysis of a mixed element for the Stokes problem. II. Construction and error estimates], Numer. Math. 62 (1992), no. 2, 161–188.

M. FRANK and P. WOLFE

[1], *An Algorithm for Quadratic Programming*, Naval Res. Log. Quart. **3** (1956), 95–110.

M. FRÉCHET

[1], *Sur la notion de différentielle totale*, Nouv. Ann. Math. **12** (1912), 385–403 and 433–449.

[2], *La notion de différentielle dans l'Analyse générale*, Ann. Čc. Norm. **XLII** (1925), 293–323.

[3], *Sur la notion de différentielle*, Journal de Mathématiques Pures et Appliquées **16** (1937), 233–250.

R. GATEAUX

[1], *Sur la représentation des fonctionnelles continues*, Rend. Acc. Linc. **22**-2 (1913), 646–648.

[2], *Sur les fonctionnelles continues et les fonctionnelles analytiques*, Comptes rendus de l'académie des sciences (Paris) **157** (1913), 325–327.

[3], *Sur la représentation des fonctionnelles continues*, Rend. Acc. Linc. **23**-1 (1914), 310–315.

[4], *Sur les fonctionnelles d'ordre entier d'approximation*, Rend. Acc. Linc. **23**-1 (1914), 405–408.

[5], *Représentation d'une fonctionnelle continue, satisfaisant à la condition du cycle fermé*, Rend. Acc. Linc. **23**-1 (1914), 481–486.

[6], *Sur la notion d'intégrale dans le domaine fonctionnel et sur la théorie du potentiel*, Bulletin de la Société Mathématique de France **47** (1919), 47–67 (followed by a note of P. Lévy 67–70).

[7], *Fonctions d'une infinité de variables indépendantes*, Bulletin de la Société Mathématique de France **47** (1919), 70–96.

[8], *Sur diverses questions du calcul fonctionnel*, Bulletin de la Société Mathématique de France **50** (1922), 1–37.

I. V. GIRSANOV

[1], *Lectures on mathematical theory of extremum problems*, Springer-Verlag, New York, 1974.

J. HADAMARD

[1], *Leçons sur le calcul des variations*, vol. I, Hermann, Paris, 1910.

[2], *La notion de différentielle dans l'enseignement*, Scripta Univ. Ab. Bib., Hierosolymitanarum, Jerusalem, 1923. Reprinted in the Mathematical Gazette **19**, no. 236 (1935), 341–342.

G. HADLEY

[1], *Nonlinear and dynamic programming*, Addison–Wesley Publishing Co., Reading, Massachusetts, 1964.

M. R. HESTENES

[1], *Optimization theory, the finite dimensional case*, J. Wiley-Interscience, New York, 1975.

M. R. HESTENES and E. STIEFEL

[1], *Methods of conjugate gradients for solving linear systems*, J. Res. Natl. Bur. Stand. **49** (1952), 409–436.

M. INTRILIGATOR

[1], *Mathematical optimization and economic theory*, Prentice–Hall, Englewood Cliffs, NJ, 1971.

A. D. IOFFE

[1], *Metric regularity and subdifferential calculus*, (Russian) Uspekhi Mat. Nauk **55** (2000), no. 3 (333), 103–162; translation in Russian Math. Surveys **55** (2000), no. 3, 501–558.

A. D. IOFFE and V. M. TIKHOMIROV

[1], *Theory of extremal problems*, (Russian) Series in Nonlinear Analysis and Its Applications, Izdat. "Nauka", Moscow, 1974. English translation by Karol Makowski: *Theory of extremal problems*, Studies in Mathematics and its Applications, 6. North–Holland Publishing Co., Amsterdam-New York, 1979.

G .G. JOSEPH

[1], *The Crest of the Peacock*, Princeton University Press, Princeton, NJ, 2000, pp. 298–300.

L .V. KANTOROVICH

[1], *Mathematical Methods of Organizing and Planning Production*, Leningrad State University Press, Leningrad, Russia, 1939.

[2], *A new method of solving some classes of extremal problems*, Doklady Akad Sci USSR **28** (1940), 211–214.

W. KARUSH

[1], *Minima of functions of several variables with inequalities as side conditions*, Master's thesis, University of Chicago, 1939.

J. KOWALIK and M. OSBORNE

[1], *Methods for unconstrained optimization problems*, American Elsevier Publishing Co., New York, 1968.

H. W. KUHN and A. W. TUCKER

[1], *Nonlinear programming*, in "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability", J. Neyman, ed., University of California Press, Berkeley, CA, 1951, pp. 481–492.

[2], *John von Neumann's work in the theory of games and mathematical economics*, Bull. Amer. Math. Soc. **64**, Number 3, Part 2 (1958), 100–122.

J. L. LAGRANGE

[1], *Sur la figure des colonnes*, in "Miscellanea Taurinensia V," **123**, 1770–1773.

[2], *Mécanique analytique*, Courcier, Paris, 1788 (reissued by Cambridge University Press, 2009).

S. R. LAY

[1], *Convex sets and their applications*, J. Wiley & Sons, New York, 1982 (book dedicated to F. A. Valentine).

A.-M. LEGENDRE

[1], *Mémoire sur la manière de distinguer les Maxima des Minima dans le Calcul des Variations*, [On the method of distinguishing maxima from minima in the calculus of variations], in "Histoire de l'Académie Royale des Sciences", Année 1786, Mémoires de Mathématiques et de Physique, Paris, 1786, pp. 7–37.

G. W. LEIBNIZ

[1], *Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractas nec irrationales quantitates moratur, et singulare pro illis calculi genus*, Acta Eruditorum, 1684 [A new method for maxima and minima and their tangents, that are not limited to fractional or irrational expressions, and a remarkable type of calculus for these].

B. LEMAIRE

[1], *Problèmes min-max et applications au contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles linéaires*, Thèse de doctorat d'état, Université de Montpellier, Montpellier, France, 1970.

D. G. LUENBERGER

[1], *Introduction to linear and nonlinear programming* Addison–Wesley, Reading, Mass., Don Mills, Ont., 1973; *Linear and nonlinear programming*, 2nd ed. Addison–Wesley, Reading, Mass., Don Mills, Ont., 1984.

[2], *Optimization by vector space methods*, Wiley, New York, 1969.

O. MANGASARIAN

[1], *Nonlinear programming*, McGraw–Hill Book Co., New York, 1969.

L. MAZILAK

[1], *The ghosts of the Ecole Normale. Life, death and destiny of René Gateaux*, submitted to Historia Mathematica (*Les fantômes de l'École Normale. Vie, mort et destin de René Gateaux*, Université Pierre et Marie Curie Paris, France, 2007).

[2], Private communication.

L. MAZILAK and R. TAZZIOLI

[1], *Mathematicians at war. Volterra and his French colleagues in World War I*, Archimedes: New Studies in the History and Philosophy of Science and Technology, 22. Springer, New York, 2009.

A. D. MICHAL

[1], *Differential calculus in linear topological spaces*, PYOC. Nat. Acad. Sci. U.S.A. **24** (1938), 340–342.

A. A. MILJUTIN

[1], *General schemes for obtaining necessary conditions for an extremum, and problems of optimal control*, (Russian) Uspehi Mat. Nauk **25**, no. 5 (155) (1970), 110–116.

B. S. MORDUKHOVICH

[1], *Variational analysis and generalized differentiation. I. Basic theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Volume 330, Springer-Verlag, Berlin, 2006.

[2], *Variational analysis and generalized differentiation. II. Applications*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Volume 331, Springer-Verlag, Berlin, 2006.

M. NAGUMO

[1], *Über die Loge der Integralkurven gewöhnlicher Differentialgleichungen*, Proc. Phys. Math. Soc. Japan **24** (1942), 551–559.

J. F. NASH

[1], *Non-Cooperative Games*, The Annals of Mathematics (2) **54** (1951), 286–295.

I. NEWTON

[1], *Method of fluxions and infinite series*, Heney Woodfall, London, 1736 (written in 1671).

[2], *Philosophiae Naturalis Principia Mathematica*, 1687 and 1726 (third edition) [Mathematical Principles of Natural Philosophy]. *Definitiones* [Definitions], pp. 1–11; *Axiomata sive leges motus* [Axioms or laws of motion], pp. 12–25); *De motu corporum: liber primus* [On the motion of bodies: book one], pp. 26–235; *De motu corporum: liber secundus* [On the motion of bodies: book two], pp. 236–400; *De mundi systemate: liber tertius* [On the system of the world: book three], pp. 401–510.

J. M. ORTEGA and W. C. RHEINBOLDT

[1], *Iterative solution of nonlinear equations in several variables*, Classics in Applied Mathematics, Volume 30, SIAM, Philadelphia, 2000.

J.-P. PENOT

[1], *Calcul sous-différentiel et optimisation*, Journal of Functional Analysis **27** (1978), 248–276.

J. PIERPONT

[1], *The theory of functions of real variables*, Vol. I, Ginn and Company, Boston, New York, Chicago, London, Atlanta, Dallas, Columbus, San Francisco, 1905.

[2], *The theory of functions of real variables*, Vol. II, Ginn and Company, Boston, New York, Chicago, London, 1912.

J. A. PLATEAU

[1], *Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires*, Gauthier-Villars, Paris, 1873.

E. POLAK

[1], *Computational methods in optimization*, Academic Press, New York, 1972.

[2], *Optimization: Algorithms and consistent approximations*, Applied Mathematical Sciences, Springer, New York,1997.

B T. POLYAK

[1], *Introduction to optimization*, Optimization Software, Inc., Publications Division, New York, 1987.

R. T. ROCKAFELLAR

[1], *Convex analysis*, Princeton University Press, Princeton, NJ, 1972.

R. T. ROCKAFELLAR and R. J.-B. WETS

[1], *Variational analysis*, Fundamental Principles of Mathematical Sciences, Volume 317, Springer, Berlin, 1998.

H. H. ROSENBROCK

[1], *An automatic method for finding the greatest or least value of a function*, Brit. Computer Journal **3**, (1960), 175–184.

W. RUDIN

[1], *Principles of mathematical analysis*, McGraw–Hill, New York, 1964.

D. L. RUSSELL

[1], *Optimization theory*, W.A. Benjamin, New York, 1970.

L. SCHWARTZ

[1], *Cours d'analyse*, Hermann, Paris, 1967.

F. SEVERI

[1], *Su alcune questioni di topologia infinitesimale*, Ann. Polon. Soc. Math. **9** (1930), 97–108.

J. STOER and C. WITZGALL

[1], *Convexity and optimization in finite dimensions*, Volume 1, Springer-Verlag, New York, 1970.

O. STOLZ

[1], *Grundzüge der Differential und Integralrechnung, I*, B. G. Teubner, Leipzig, 1893 (The Cornell University Library Collections).

W. TAKAHASHI

[1], *Existence theorems generalizing fixed point theorems for multivalued mappings*, in "Fixed Point Theory and Applications (Marseille, 1989)", M. Thra and J. Baillon, eds., pp. 397–406, Pitman Res. Notes Math. Ser., Volume 252, Longman Sci. Tech., Harlow, 1991.

V. M. TIHOMIROV

[1], *Elements of the theory of extrema*, Econometric Institute, Erasmus University Rotterdam, 1997.

H. TUY

[1], *Convex analysis and global optimization*, Kluwer Academic Publishers, Dordrecht, 1998.

F. A. VALENTINE

[1], *Convex sets*, McGraw–Hill, New York, 1964.

B. VAN BRUNT

[1], *The Calculus of Variations*, Springer-Verlag, New York, 2004.

J. VON NEUMANN

[1], *Zur Theorie der Gesellschaftsspiele*, Mathematische Annalen **100** (1) (1928), 295–320. English translation: "On the Theory of Games of Strategy," in "Contributions to the Theory of Games," v. 4, A. W. Tucker and R. D. Luce, eds. (1959), pp. 13–42.

J. VON NEUMANN and O. MORGENSTERN

[1], *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944.

K. WEIERSTRASS

[1], *Mathematische Werke*. Rester Band. *Abhandlungen I*. Mayer & Müller, Berlin 1894; Zweiter Band. *Abhandlungen II*. Mayer & Müller, Berlin 1895; Dritter Band. *Abhandlungen III*. Mayer & Müller, Berlin 1903.

[2], *Mathematische Werke*. Vierter Band. *Vorlesungen über die Theorie der Abelschen Transcendenten*. Mayer & Müller, Berlin 1902.

[3], *Mathematische Werke*. Siebenter Band. *Vorlesungen über Variationsrechnung*. Akademische Verlagsgesellschaft M. B. H., Leipzig, 1927.

J. D. WESTON

[1], *A characterization of metric completeness*, Proc. Amer. Math. Soc. **64** (1977), no. 1, 186–188.

H. WHITNEY

[1], *A function that is not constant on a connected set of critical points*, Duke Math. J. **1**, no. 4 (1935), 514–517.

W. H. YOUNG

[1], *On differentials*, Proc. London Mathematical Society, series 2, **7** (1909), 157.

[2], *The Fundamental Theorems of Differential Calculus*, Cambridge University Press, Cambridge, 1910.

W. ZANGWILL

[1], *Nonlinear Programming—A Unified Approach*, Prentice–Hall, Englewood Cliffs, NJ, 1969.

J.-P. ZOLÉSIO

[1], *Identification de domaines par déformation*, Thèse de doctorat d'état, Université de Nice, France, 1979.

[2], *Semiderivatives of repeated eigenvalues*, in "Optimization of Distributed Parameter Structures," Vol. II (Iowa City, IA, 1980), E. J. Haug and J. Céa, eds., 1457–1473, NATO Adv. Sci. Inst. Ser. E: Appl. Sci., 50, Sijhofff and Nordhoff, Alphen aan den Rijn, 1981 (Nijhoff, The Hague).

Index of Notation

Index