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