# Springer Undergraduate Texts in Mathematics and Technology

**Springer Undergraduate Texts in Mathematics and Technology** (SUMAT) publishes textbooks aimed primarily at the undergraduate. Each text is designed principally for students who are considering careers either in the mathematical sciences or in technology-based areas such as engineering, finance, information technology and computer science, bioscience and medicine, optimization or industry. Texts aim to be accessible introductions to a wide range of core mathematical disciplines and their practical, real-world applications; and are fashioned both for course use and for independent study.

More information about this series at http://www.springer.com/series/7438

Francisco J. Aragón · Miguel A. Goberna ·
Marco A. López · Margarita M. L. Rodríguez

# Nonlinear Optimization

Springer

Francisco J. Aragón
Department of Mathematics
University of Alicante
Alicante, Spain

Miguel A. Goberna
Department of Mathematics
University of Alicante
Alicante, Spain

Marco A. López
Department of Mathematics
University of Alicante
Alicante, Spain

Margarita M. L. Rodríguez
Department of Mathematics
University of Alicante
Alicante, Spain

ISSN 1867-5506 ISSN 1867-5514 (electronic)
Springer Undergraduate Texts in Mathematics and Technology
ISBN 978-3-030-11183-0 ISBN 978-3-030-11184-7 (eBook)
https://doi.org/10.1007/978-3-030-11184-7

Library of Congress Control Number: 2018966836

Mathematics Subject Classification (2010): 65kxx, 97uxx, 47N10



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

This book is aimed at upper-level undergraduate students of mathematics and statistics, and graduate students of industrial engineering. We assume that the readers are familiar with real analysis and linear algebra (subjects that are briefly revisited in Chapter 1), and we decline to use sophisticated tools from convex analysis (such as the Rockafellar subdifferential or the Fenchel–Moreau conjugate function) and nonsmooth analysis (such as the Clarke and Mordukhovich subdifferentials). It is not easy to select, among the vast amount of knowledge on the subject, the suitable contents for a textbook on nonlinear optimization. We have included in this book some basic theoretical results (in particular, optimality and duality theorems for different problems) and a representative selection of the many available numerical methods for problems with or without constraints, some of them not using derivatives. Both theoretical results and numerical methods are illustrated with applications to a diverse range of fields.

The contents are organized in two parts that can be used as textbooks for undergraduate courses on convex and nonconvex optimization, respectively. In fact, this book is based on the contents of two undergraduate courses taught at the University of Alicante. These two courses are preceded by a course on linear programming, which is based on the book [42]. Nevertheless, this knowledge is not assumed in the present book. With the exception of Chapter 1, which is preliminary for both parts, Part I and Part II are essentially independent. The sections and subsections marked with an asterisk (*) are mathematically harder than the rest and can be skipped when either the students have a limited background or when the available time is insufficient.

Part I, composed of Chapters 2–4, focuses on the analytic calculation of global minima (required in engineering, operations research, statistics, and economy) and local minima (required in natural sciences, where the equilibrium situations correspond to local minima—or even critical points—of certain functions). Chapters 1 and 2 contain the basic ingredients for the calculus of local minima (optimality conditions for differentiable functions) and global minima (coercivity and convexity) for unconstrained optimization problems. Chapter 3 provides closed formulas for unconstrained optimization problems arising in different fields, while

Chapter 4 deals with unconstrained and constrained convex optimization problems for which local and global minima coincide. For the sake of simplicity, the sections devoted to constrained continuous optimization preclude linear equations and make use of a unique constraint qualification (the Slater condition).

Part II is focused on the numerical calculation of local optima in problems whose solutions cannot be analytically obtained, and it consists of two chapters. Chapter 5 deals with the standard algorithms for unconstrained optimization problems, such as the steepest descent method, Newton's method and variants (trust regions, Gauss–Newton, Levenberg–Marquardt), and other gradient-based methods such as those using conjugate directions (conjugate gradient and quasi-Newton). The local or global convergence of such methods is discussed in detail, as well as the respective rates of convergence (linear, superlinear, and quadratic). This chapter also includes an introduction to methods that do not use derivatives. Finally, Chapter 6 presents, in the first part, an introduction to the so-called penalty and barrier methods. These procedures address the numerical solution of constrained optimization problems by applying the algorithms described in Chapter 5, and they are based on the idea of converting a constrained problem in a sequence of unconstrained ones. The second part of this final chapter is devoted to the optimality conditions for constrained optimization problems, with equality and/or inequality restrictions. The relevance of the so-called constraint qualifications is emphasized as they yield necessary optimality conditions, which can be used as stopping rules for the algorithms. Finally, we show how the optimality conditions give rise to the sequential quadratic programming methods, when Newton's method is applied for solving the associated system of equations.

Definitions, proofs, and numerical methods are illustrated with figures. Moreover, all chapters contain collections of exercises, and the solutions of those exercises that our students usually find harder are provided in a separate appendix at the end of the book.

This is a book on nonlinear optimization for undergraduate students, so it is not intended to provide an exhaustive overview of the available optimality theory and optimization methods. Readers wishing to get a deeper knowledge of the topics included in this book are kindly invited to consult the list of references provided at the end, especially the books [5, 8, 11, 28, 42, 52, 54, 67, 70, 75], whose treatment of different subjects has inspired the approach adopted in some sections. Readers interested in the history of optimization, or in its key role in operations research, may find sources of information in the books [39, 41, 47, 62].

The top five reasons to use this textbook are:

- It only assumes basic knowledge of differential and matrix calculus. All the main concepts, algorithms, and proofs are illustrated with highly explanatory and appealing figures.
- It pays special attention to model building and validation of real problems, and emphasizes the practical advantages of obtaining good reformulations of optimization problems. It shows important applications of optimization models to natural and social sciences, engineering, and data science.

- It provides rigorous optimality conditions. The existence and uniqueness of optimal solutions is analyzed via coercivity and convexity. Dual problems are introduced in order to get lower bounds, sensitivity information ("what if" questions), and stopping rules for primal-dual algorithms.
- It provides an accurate description of the main numerical approaches to solve nonlinear optimization problems. These numerical methods have been chosen with the aim of covering an ample range of techniques. Intentionally, the algorithms described are not implemented in a specific software, as experience shows that this type of choice is highly ephemeral and mainly depends on personal preferences. Nonetheless, five assignments for laboratory sessions of two hours have been included in Section 5.8.
- It has been thoroughly tested. Its preliminary versions have been used for many years as class notes for various undergraduate courses on optimization theory and methods. The exercises have been carefully selected to push the students to a deeper understanding of the main topics in each chapter. Further, a detailed solution to those exercises that our students consider harder is provided, to allow independent study.

We would like to thank our families for their patience and constant support. We are grateful to the anonymous referees and to our editors at Springer, Elizabeth Loew and Razia Amzad, whose suggestions helped us improve this book. We are indebted to Bernardo Cascales, who proposed us to write the initial version of this book in Spanish, and to Jonathan Borwein, who encouraged us to translate the manuscript into English. Our gratitude also goes to all the students of the Mathematics Degree at the University of Alicante who gave us critical remarks on the preliminary versions of this book. Finally, we are thankful to the funding agencies MINECO (Spain) and ERDF (EU) for their financial support along the years. In particular, this book was written as part of the outreach activities of the grant MTM2014-59179-C2-1-P and the research contract RYC-2013-13327 of the first author.

Alicante, Spain
November 2018

Francisco J. Aragón
Miguel A. Goberna
Marco A. López
Margarita M. L. Rodríguez

The original version of this book was inadvertently published with incorrect reference numbers in the frontmatter. The corrections to this book can be found at https://doi.org/10.1007/978-3-030-11184-7_7

# Contents

# Nomenclature

| | |
|---|---|
| $]a,b[$ | Open interval between the real numbers $a$ and $b$ |
| $[a,b]$ | Closed interval between the real numbers $a$ and $b$ |
| $\mathbb{R}_+$ | Set of nonnegative real numbers |
| $\mathbb{R}_{++}$ | Set of positive real numbers |
| $\overline{\mathbb{R}}$ | Extended real line, page 154 |
| $\min X$ | Minimum of $X \subset \mathbb{R}$ |
| $\max X$ | Maximum of $X \subset \mathbb{R}$ |
| $\inf X$ | Infimum of $X \subset \mathbb{R}$ |
| $\sup X$ | Supremum of $X \subset \mathbb{R}$ |
| $I_n$ | $n \times n$ identity matrix |
| $A_k$ | $k \times k$ submatrix of $A$ which results of taking the first $k$ rows and columns |
| $\Delta_k$ | $k$th director principal minor of $A$ |
| $A^T$ | Transpose matrix of $A$ |
| $A^{-1}$ | Inverse matrix of $A$ |
| $\det A$ | Determinant of the matrix $A$ |
| $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ | Diagonal matrix with diagonal elements $\lambda_1, \ldots, \lambda_n$ |
| $\rho(A)$ | Spectral radius of the matrix $A$, page 31 |
| $\mathrm{cond}(A)$ | Condition number of the matrix $A$, page 31 |
| $\mathcal{S}_n$ | Family of symmetric $n \times n$ real matrices |
| $\mathcal{S}_q^+$ | Cone of positive semidefinite symmetric matrices |
| $0_{n\times n}$ | $n \times n$ null matrix |
| $0_n$ | Zero vector of $\mathbb{R}^n$ |
| $e_i$ | $i$th vector of the standard basis |
| $1_n$ | Vector of all ones in $\mathbb{R}^n$ |
| $\|x\|_1$ | Manhattan (or $\ell_1$) norm of $x$, page 15 |
| $\|x\|_\infty$ | Chebyshev (or $\ell_\infty$) norm of $x$, page 15 |
| $\|x\|$ | Euclidean (or $\ell_2$) norm of $x$, page 15 |
| $\mathbb{B}$ | Euclidean closed unit ball |
| $x^T y$ | Scalar product of the vectors $x$ and $y$ |

| | |
|---|---|
| $\{x_k\}$ | Sequence in $\mathbb{R}^n$ |
| lim sup | Upper limit of a sequence |
| $\lvert X\rvert$ | Cardinality of $X$ |
| $\operatorname{int} X$ | Topological interior of $X$ |
| $\operatorname{cl} X$ | Closure of $X$ |
| $\operatorname{bd} X$ | Boundary of $X$ |
| $\operatorname{span} X$ | Linear hull of $X$ |
| $\operatorname{conv} X$ | Convex hull of $X$, page 55 |
| $\operatorname{cone} X$ | Conical (convex) hull of $X$, page 58 |
| $K^\circ$ | (Negative) polar cone of the cone $K$, page 135 |
| $K_p^m$ | Second-order cone, page 172 |
| $f'(x)$ | Derivative of $f$ at $x$ |
| $\frac{\partial f(x)}{\partial x_i}$ | Partial derivative of $f$ with respect to $x_i$ at $x$ |
| $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ | Second-order partial derivative of $f$ with respect to $x_i$ and to $x_j$ at $x$ |
| $f'(x;u)$ | Directional derivative of $f$ at $x$ in the direction of $u$ |
| $f'_+(a)$ | Right derivative of $f$ at $a$, page 62 |
| $f'_-(a)$ | Left derivative of $f$ at $a$, page 62 |
| $\nabla f(x)$ | Gradient of $f$ at $x$, page 19<br>Gradient matrix of $f$ at $x$, page 26 |
| $\nabla^2 f(x)$ | Hessian matrix of $f$ at $x$, page 24 |
| $J_f(x)$ | Jacobian matrix of $f$ at $x$, page 26 |
| $\operatorname{dom} f$ | Domain of the function $f$, pages 17 and 154 |
| $\operatorname{gph} f$ | Graph of the function $f$, page 17 |
| $\operatorname{epi} f$ | Epigraph of the function $f$, page 72 |
| $S_\lambda(f)$ | Sublevel set $\lambda$ of $f$, page 38 |
| $\operatorname{argmin} f$ | Set of points in which $f$ reaches its minimum value |
| $o(g(x))$ | Landau notation, page 24 |
| $v_+$ | Positive part of the vector $v$, page 120 |
| $p_+$ | Positive part function, page 121 |
| $\mathcal{F}$ | Feasible set multifunction, page 154 |
| $\vartheta$ | Value function, page 154 |
| $\mathcal{C}(X)$ | Linear space of continuous real-valued functions on $X$ |
| $\mathcal{C}^m(X)$ | Linear space of real functions on $X$ with continuous partial derivatives of order $m$ |
| $F$ | Feasible set of the problem $P$, page 8 |
| $F^*$ | Optimal set of the problem $P$, page 8 |
| $v(P)$ | Optimal value of the problem $P$, page 8 |
| $I(\overline{x})$ | Set of active indices at $\overline{x}$, page 134 |
| $A(\overline{x})$ | Active cone at $\overline{x}$, page 135 |
| $D(\overline{x})$ | Feasible direction cone at $\overline{x}$, page 134 |