Convexity and its Applications in Discrete and Continuous Optimization

Using a pedagogical, unified approach, this book presents both the analytic and combinatorial aspects of convexity and its applications in optimization. On the structural side, this is done via an exposition of classical convex analysis and geometry, along with polyhedral theory and geometry of numbers. On the algorithmic/optimization side, this is done by the first ever exposition of the theory of general mixed-integer convex optimization in a textbook setting. Classical continuous convex optimization and pure integer convex optimization are presented as special cases, without compromising on the depth of either of these areas. For this purpose, several new developments from the past decade are presented for the first time outside technical research articles: discrete Helly numbers, new insights into sublinear functions, and best known bounds on the information and algorithmic complexity of mixed-integer convex optimization. Pedagogical explanations and more than 300 exercises make this book ideal for students and researchers.

Amitabh Basu is Professor of Applied Mathematics and Statistics at Johns Hopkins University. He has received the NSF CAREER award and the Egon Balas Prize from the INFORMS Optimization Society. He serves on the editorial boards of *Mathematics of Operations Research, Mathematical Programming, SIAM Journal on Optimization*, and the *MOS-SIAM Series on Optimization*.

"Written by one of the most brilliant researchers in the field, this book provides an elegant, rigorous, and original presentation of the theory of convexity, describing in a unified way its use in continuous and discrete optimization, and also covering some very recent advancements in these areas."

– **Marco Di Summa**, *University of Padua*

"Convexity is central to most optimization algorithms. This book brings together classical and new developments at the interface between these two vibrant areas of mathematics. It is an essential reference for scholars in optimization. The numerous exercises make it an ideal textbook at the graduate and upper undergraduate levels."

– **Gérard Cornuéjols**, *Carnegie Mellon University*

Convexity and its Applications in Discrete and Continuous Optimization

AMITABH BASU
Johns Hopkins University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108837590

DOI: 10.1017/9781108946650

When citing this work, please include a reference to the DOI 10.1017/9781108946650

First published 2025

Printed in the United Kingdom by CPI Group Ltd, Croydon CR0 4YY

A catalogue record for this publication is available from the British Library

A Cataloging-in-Publication data record for this book is available from the Library of Congress

ISBN 978-1-108-83759-0 Hardback

Dedicated to Ma and Baba

Contents

Preface *page* xiii
Suggestions for classroom use xvi
Prerequisites xvii
Acknowledgments xvii

1 Preliminaries 1
1.1 Euclidean Spaces 2
1.2 Linear Algebra 4
1.2.1 Singular Value and Eigen Decompositions 7
1.2.2 Matrix Norms 8
1.2.3 Exercises 9
1.3 Real Analysis 10
1.3.1 Exercises 13
1.4 Models of Computation 13
1.4.1 Arithmetic Model of Computation 14
1.4.2 Turing Machine Model of Computation 14
1.4.3 Oracle-Based Computation 16
1.4.3.1 Turing Machines Augmented with Real Number Oracles 16
1.4.3.2 General Oracles 17
1.5 Notes and Related Literature 18

PART I STRUCTURAL ASPECTS 21

2 Convex Sets 23
2.1 Definitions and Basic Properties 23
2.1.1 Exercises 26
2.2 Convex Cones, Affine Sets, and Dimension 26
2.2.1 Coordinatization of the Affine Hull and Volumes 30

2.2.2 Carathéodory's Theorem 32
2.2.3 Exercises 34
2.3 The Projection Map 36
2.3.1 Exercises 38
2.4 Representations of Convex Sets 39
2.4.1 Extrinsic Description: Separating Hyperplanes 39
2.4.1.1 How to Represent General Convex Sets: Separation Oracles 40
2.4.1.2 Duality/Polarity 41
2.4.2 Intrinsic Description: Faces, Recession Cone, Lineality Space 42
2.4.3 A Remark about Extrinsic and Intrinsic Descriptions 50
2.4.4 Exercises 51
2.5 Polyhedra 55
2.5.1 The Minkowski–Weyl Theorem 58
2.5.2 Farkas' Lemma 60
2.5.3 Valid Inequalities and Feasibility 61
2.5.4 Faces of Polyhedra 63
2.5.5 Implicit Equalities, Dimension of Polyhedra, and Facets 65
2.5.6 Exercises 68
2.6 Helly Numbers 71
An application to learning theory: VC dimension of halfspaces 73
2.6.1 Exercises 77
2.7 Ellipsoids 78
2.7.1 Ellipsoidal Approximations 81
2.7.2 Exercises 85
2.8 Notes and Related Literature 86

3 Convex Functions 89
3.1 General Properties, Epigraphs, Subgradients 89
3.1.1 Exercises 97
3.2 Analytic Properties 98
3.2.1 Continuity 98
3.2.2 First-Order Derivative Properties 101
3.2.3 Second-Order Derivative Properties 104
3.2.4 Exercises 106
3.3 Sublinear Functions 107
3.3.1 Nonnegative Sublinear Functions: Gauges 109
3.3.2 Support Functions 115

Generalized Cauchy–Schwarz/Hölder's inequality 117
3.3.3 Correspondence between Closed, Convex Sets and Closed, Sublinear Functions 118
3.3.4 Representing Closed, Convex Sets as 1-Sublevel Sets 122
3.3.5 Exercises 126
3.4 Directional Derivatives, Subgradients, and Subdifferential Calculus 128
3.4.1 Exercises 132
3.5 Volume Relations 132
3.5.1 The Brunn–Minkowski Theorem 133
3.5.2 Sectional Volumes and Symmetrization 134
3.5.3 The Rogers–Shephard Inequality 135
3.5.4 Exercises 136
3.6 Notes and Related Literature 137

4 Geometry of Numbers 139
4.1 Lattices in Euclidean Space 139
4.1.1 Exercises 144
4.2 Minkowski's Convex Body Theorem 147
4.2.1 Exercises 150
4.3 Packing, Covering, SVP, and CVP 151
4.3.1 Bounds for Packing and Covering Radii 155
4.3.2 Dirichlet–Voronoi Cells 157
4.3.3 Exercises 160
4.4 Lattice-Free Convex Sets 162
4.4.1 Maximal Lattice-Free Convex Sets 162
4.4.2 Khinchine's Flatness Theorem 166
4.4.3 Exercises 167
4.5 Notes and Related Literature 168

PART II OPTIMIZATION 171

5 Ingredients of Mathematical Optimization 173
5.1 The General Optimization Problem 174
5.1.1 Solution Operators and Loss Functions 177
5.1.2 Exercises 177
5.2 Algorithmic Optimization 178
5.2.1 Direct Encodings and Oracle-Based Access 181
5.3 Bounding the Complexity 183
5.3.1 Oracle Ambiguity and Lower Bounds on Complexity 183

5.3.2 Exercises 186
5.4 Parameterized Complexity 186
5.5 Notes and Related Literature 188

6 Complexity of Convex Optimization with Integer Variables 190
6.1 Problem Setup 190
6.1.1 First-Order Oracles 191
6.1.2 Parameterization of the Instances 192
6.1.3 Approximate Solutions 193
6.2 Information Complexity 194
6.2.1 Proof of the Lower Bounds in Theorem 6.2.1 195
The case with $d \geq 1$ 196
The case with $d = 0$ 201
6.2.2 A Geometric Detour: Centerpoints 202
The case with $n \geq 1$ 203
The case with $n = 0$ 203
6.2.3 Proof of the Upper Bounds in Theorem 6.2.1 204
The mixed-integer case with $n, d \geq 1$ 204
The pure integer case with $d = 0$ 207
The pure continuous case with $n = 0$ 207
6.2.4 Exercises 208
6.3 Algorithmic Complexity 210
6.3.1 The Unconstrained Quadratic Case 211
6.3.1.1 Inner Products, Positive Definite Matrices, and Quadratic Functions 212
6.3.1.2 The Pure Continuous Case: Conjugate Gradients 214
6.3.1.3 The Pure Discrete Case: Closest Lattice Vector Problem (CVP) 218
Reduced bases 218
The Lenstra–Lenstra–Lovasz (LLL) algorithm 221
Computing CVP and Voronoi cells 223
6.3.1.4 The Mixed-Integer Case 229
6.3.1.5 Shortest Lattice Vector Problem (SVP) from CVP 232
6.3.2 General Mixed-Integer Convex Optimization 233
6.3.3 Exercises 238
6.4 The Branch-and-Cut Method 240
6.4.1 The Relationship between Branching and Cutting Planes 244
6.4.2 Exercises 247
6.5 The Special Case of Continuous Convex Optimization 247

6.5.1 Cutting-Plane Methods 248
Information complexity 248
Algorithmic complexity 249
6.5.2 Subgradient Descent 250
6.5.2.1 Smooth Gradients 253
6.5.2.2 Strong Convexity 255
6.5.3 Exercises 256
6.6 Notes and Related Literature 257
Arithmetic versus Turing machine models of computation 261

7 Certificates and Duality 263
7.1 Continuous Optimization: The Case with No Integer Variables 263
7.1.1 Exact Optimality Certificates 264
7.1.2 Approximate Optimality Certificates 267
7.1.2.1 Volume-Based Certificates 267
7.1.2.2 Lagrangian Duality with Generalized Constraints 268
Explicit examples of the Lagrangian dual 273
ϵ-optimality certificates from the Lagrangian dual 275
Strong duality, complementary slackness, and the KKT conditions 275
Solving the Lagrangian dual problem 280
Saddle point interpretation of the Lagrangian dual 280
7.1.3 Exercises 282
7.2 The General Mixed-Integer Case 283
7.3 Notes and Related Literature 284

Hints to Selected Exercises 287
References 291
Index 305

Preface

This book is about convex analysis and geometry, as well as their use in mathematical optimization. Mathematical optimization plays a central role in almost all mathematical and computational disciplines, and its importance in the age of data science, machine learning, and artificial intelligence cannot be overemphasized. Almost every problem in these areas has a mathematical optimization problem at its core. This book focuses primarily on convex optimization (both discrete and continuous) and its mathematical foundations. The reach of convex optimization methods and, more generally, the underlying concepts from convex analysis and geometry in business, engineering, and scientific applications is also significant. The simple idea behind convexity – that the "value" of an average of multiple potential solutions is at least as good as the average "value" of these solutions – has the ability to model an amazing variety of phenomena.

There are enough textbooks on convexity and convex optimization to occupy an entire section of an academic (virtual) library. There are two reasons why I believe this book provides something new and valuable, particularly from the perspective of designing a course that introduces these ideas rigorously. First, I feel that previous textbooks and monographs focus either on the continuous/analytic aspects of convexity [42, 44, 45, 57, 59, 114, 131, 140, 141, 170, 183, 199, 210, 220] or on the discrete/combinatorial aspects of convexity [24, 116, 124, 128, 234], and no one single book surveys in a unified way the main ideas in both of these aspects (theory and applications in optimization).[1] Second, there have been several recent developments, especially in the use of convexity in discrete optimization, that appear only in specialized research papers or surveys. These developments

[1] An exception to this is the classic text [218], but we would argue that this book is somewhat limited and outdated in its scope.

have now reached a level of maturity that can be presented in a textbook, with clean and elegant proofs available for the major results. I am convinced that the time is right for a textbook that gives a unified perspective on continuous and discrete optimization via the lens of convexity, and presents some of the recent advances in optimization that have not yet made it to the texts on convexity.

The book is divided into two parts. The first part focuses on the mathematical foundations of convex analysis and geometry, starting with a study of convex sets, followed by an exposition of convex functions, and ending with an introduction to the geometry of numbers. The second part focuses on optimization, presenting discrete and continuous ideas in a common framework. Here are some topics covered in this textbook that, to the best of my knowledge, are not contained in any previous textbook on convex analysis, geometry, or optimization.

1. A careful exposition of the conceptual underpinnings of algorithmic or computational optimization. This topic is approached by continuous and discrete optimizers in related but distinct ways. At the risk of making a sweeping generalization, one could say that computation in continuous optimization has its origins in the traditions of scientific computing and numerical analysis, whereas discrete optimization broadly views computation via the Turing machine model. The different views lead to some friction when trying to cross the boundaries. In the continuous world, one often designs algorithms assuming one can perform exact operations with real numbers (consider, for example, Newton's method), which is impossible in the Turing machine model. In the discrete world, the "input" to a Turing machine becomes a tricky question when dealing with general nonlinear functions and sets. The question of "complexity" of an optimization algorithm is also treated in somewhat different ways in the two communities. For example, it is not clear what the "size" of an optimization problem is when one has nonlinear objective functions and constraints, and "complexity" without a notion of "size" is hard to formulate precisely in the Turing machine model. We believe Section 1.4 and Chapter 5 show how all these issues can be handled in a unified, coherent way, making no distinction whatsoever between "continuous" and "discrete" optimization.
2. Mixed-integer convex optimization, i.e., minimizing a convex function subject to convex constraints where some of the decision variables have to take integer values, is a natural model for presenting continuous and discrete optimization under one umbrella. Continuous convex optimization is the special case where no variable is integer constrained. On the other hand, the use of integer variables to model combinatorial optimization

problems is well known. Chapter 6 of this book presents state-of-the-art results on information and algorithmic complexity of mixed-integer convex optimization.

Information complexity of classical continuous optimization has been well understood since the 1970s due to seminal work by Nemirovski and Yudin. The information complexity in the presence of integer variables was not well developed until research work done in the past decade, and I believe this fundamental topic should become part of the early education in optimization.

On the algorithmic side, decades of work in integer optimization has focused on understanding the best possible dependence of algorithmic complexity on the number n of integer variables. This book presents the best known upper bound of $2^{n\log(n)}$ on the complexity of *deterministic* algorithms for convex integer optimization, which does not appear outside specialized, technical research articles. Moreover, this book gives a general mixed-integer complexity bound allowing for both integer and continuous variables. Such a bound does not explicitly appear anywhere in the literature, to the best of our knowledge; previous work focused on the pure integer case with no continuous variables.

3. Helly's theorem is a classical result in combinatorial convexity that states that if a collection of convex sets in $\mathbb{R}^d$ has empty intersection then there is a subcollection of at most $d+1$ sets whose intersection is already empty. This fundamental result has many applications, and it has been extended in several directions. One such generalization states that if a collection of convex sets in $\mathbb{R}^n \times \mathbb{R}^d$ has empty intersection with $\mathbb{Z}^n \times \mathbb{R}^d$, then a subcollection of at most $2^n(d+1)$ sets already has empty intersection with $\mathbb{Z}^n \times \mathbb{R}^d$. This has major consequences in convex optimization in the presence of integer variables, especially in establishing tight bounds on information complexity as discussed in Chapter 6 of this book. The theory of Helly numbers encompassing the above generalization is presented in Section 2.6. This material appeared previously only in specialized surveys and research articles.
4. In classical convex analysis, it is well known that the gauge function of a convex set C containing the origin is the unique *nonnegative* sublinear function whose 1-sublevel set is C. However, when C is not compact, there are other sublinear functions that also represent C as their 1-sublevel set and it was known that the gauge is the largest of all such functions pointwise. In recent work [33, 73], it was shown that there is a unique function that is the pointwise *smallest* such sublinear function and it can be described quite explicitly as the support function of a certain subset of the polar.

In geometric terms, there exists a subset $C^\star \subseteq C^\circ$ of the polar of C such that for all sets D such that $C = D^\circ$, we must have $C^\star \subseteq D \subseteq C^\circ$. The smallest sublinear function representing C is precisely the support function of $C^\star$. A precise description of $C^\star$ is also available in terms of supporting hyperplanes of C. This result, presented in Section 3.3.4, has not appeared in any previous convex analysis textbook. Its use in discrete optimization is discussed in Section 3.6.

5. The algorithmic theory of mixed-integer convex optimization relies on tools from algorithmic geometry of numbers such as lattice basis reduction, Voronoi cell computations, and closest and shortest lattice vector problems (CVP and SVP). The usual presentation of these algorithms, especially lattice basis reduction algorithms, works under the assumption that the lattice is given by rational vectors. This book presents an analysis of these algorithms for nonrational input (working in the real arithmetic model of computation) and the complexity is given in terms of basic properties of the lattice, or equivalently, in terms of condition numbers and norms, in the spirit of numerical analysis (see Section 6.3). These results reduce to the standard results when restricted to rational input.
6. Recent work on the complexity of branch-and-cut methods for mixed-integer convex optimization is presented in Chapter 6. Some recent work on duality for mixed-integer convex optimization is also summarized in Chapter 7; however, the classical duality theory for continuous optimization receives more attention in this chapter.

Suggestions for classroom use. The book grew out of lecture notes for an upper-level undergraduate course on convexity and continuous convex optimization I designed at Johns Hopkins University. Several sections have also been used in a graduate-level discrete optimization course I teach at Hopkins. Based on my experience with these courses, I give some suggestions below for course material at various levels. I hope the textbook can serve as a reference for courses on continuous optimization as well as discrete optimization. In addition, it should be able to serve courses focusing on structural aspects of convex analysis, convex geometry, and discrete geometry. There are over 300 exercises in the textbook, ranging from routine calculations, technical proofs, and tests of conceptual understanding to covering important results that extend the main material. Hints are provided for exercises marked with an asterisk at the end of the book.

1. An upper-level undergraduate or beginning graduate course on convex analysis and continuous convex optimization can be based on Chapters

2, 3, 6, and 7. Chapter 6 can be restricted to simply Section 6.5, except if a discussion of lower bounds on information complexity is intended for inclusion. For this, the content of Section 6.2.1 should be presented for the $n = 0$ setting. Depending on the emphasis of the course, certain sections of Chapters 2 and 3 can be omitted. For example, in the course I teach at Hopkins, I skip Section 2.7 on ellipsoids and use relevant results from this section without proof in my discussion of cutting-plane methods (ellipsoid or center-of-gravity methods) in Chapter 6. The discussion of Helly numbers from Section 2.6, if included, should be restricted to the continuous case. I also skip Section 3.5 on Brunn–Minkowski theory and appeal to Grunbaum's result on centroids without proof in the discussion of the center-of-gravity method in Chapter 6. Chapter 7 should also be restricted to the classical continuous optimization duality theory.

2. A first- or second-year graduate course on convex analysis and geometry, as well as geometry of numbers, with no optimization, can be based on Chapters 2, 3, and 4. All the material in these chapters can be covered in their entirety in one semester. This can serve as a foundation for advanced graduate courses on continuous convex optimization or integer optimization, discrete geometry, Banach space geometry, or functional analytic aspects of convex geometry.
3. A Ph.D.-level course on algorithmic geometry of numbers and mixed-integer convex optimization can be based on Chapters 4, 5, and 6. Relevant parts of Chapters 2 and 3 should be covered as needed.

Prerequisites. Familiarity with basic linear algebra and real analysis concepts, as reviewed in Sections 1.1, 1.2, and 1.3, will be very helpful. Nevertheless, every year the author has had students in his undergraduate course at Hopkins who did not have this background at the start of the course, but it was possible for the motivated ones to pick up this background through independent reading with Chapter 1 of this book as a guide, as well as discussions with the instructor and fellow students. The only real prerequisites are a familiarity with mathematical arguments and proofs.

Acknowledgments. Several colleagues looked at evolving lecture notes and initial drafts of this book and gave excellent feedback. A special thanks goes to Marco Di Summa, who meticulously went through an initial draft and gave immensely useful mathematical suggestions, including corrections in proofs and exposition, as well as organizational suggestions. I also wish to thank Michele Conforti, Giacomo Zambelli, Santanu Dey, and Marco Molinaro for their feedback and encouragement. I am very grateful to Katie Leach at

Cambridge University Press for suggesting the possibility of converting my lecture notes into a textbook, for her immense patience as the vision of the book evolved, and for providing constant encouragement through all the time I worked on the book. I am also very grateful for support from the National Science Foundation (NSF) and the Air Force Office of Scientific Research (AFOSR) provided during the period that I worked on this book. Finally, I would like to express my deepest gratitude to my wife, Deepthi, and my daughter, Maitreyi. This book ultimately owes its existence to time stolen from them.