Contents

1 Introduction 1

1.1 Motivating examples 1
1.2 Optimization problems 5
1.3 Important classes of optimization problems 10
1.4 History 14

I Linear algebra models 19

2 Vectors and functions 21

2.1 Vector basics 21
2.2 Norms and inner products 28
2.3 Projections onto subspaces 37
2.4 Functions 43
2.5 Exercises 53

3 Matrices 55

3.1 Matrix basics 55
3.2 Matrices as linear maps 61
3.3 Determinants, eigenvalues, and eigenvectors 64
3.4 Matrices with special structure and properties 75
3.5 Matrix factorizations 82

3.6 *Matrix norms* 84
3.7 *Matrix functions* 87
3.8 *Exercises* 91

4 *Symmetric matrices* 97
4.1 *Basics* 97
4.2 *The spectral theorem* 103
4.3 *Spectral decomposition and optimization* 107
4.4 *Positive semidefinite matrices* 110
4.5 *Exercises* 118

5 *Singular value decomposition* 123
5.1 *Singular value decomposition* 123
5.2 *Matrix properties via SVD* 127
5.3 *SVD and optimization* 133
5.4 *Exercises* 145

6 *Linear equations and least squares* 151
6.1 *Motivation and examples* 151
6.2 *The set of solutions of linear equations* 158
6.3 *Least-squares and minimum-norm solutions* 160
6.4 *Solving systems of linear equations and LS problems* 169
6.5 *Sensitivity of solutions* 173
6.6 *Direct and inverse mapping of a unit ball* 177
6.7 *Variants of the least-squares problem* 183
6.8 *Exercises* 193

7 *Matrix algorithms* 199
7.1 *Computing eigenvalues and eigenvectors* 199

7.2 Solving square systems of linear equations 206

7.3 QR factorization 211

7.4 Exercises 215

II Convex optimization models 221

8 Convexity 223

8.1 Convex sets 223

8.2 Convex functions 230

8.3 Convex problems 249

8.4 Optimality conditions 268

8.5 Duality 272

8.6 Exercises 287

9 Linear, quadratic, and geometric models 293

9.1 Unconstrained minimization of quadratic functions 294

9.2 Geometry of linear and convex quadratic inequalities 296

9.3 Linear programs 302

9.4 Quadratic programs 311

9.5 Modeling with LP and QP 320

9.6 LS-related quadratic programs 331

9.7 Geometric programs 335

9.8 Exercises 341

10 Second-order cone and robust models 347

10.1 Second-order cone programs 347

10.2 SOCP-representable problems and examples 353

10.3 Robust optimization models 368

10.4 Exercises 377

11 Semidefinite models 381

11.1 From linear to conic models 381

11.2 Linear matrix inequalities 383

11.3 Semidefinite programs 393

11.4 Examples of SDP models 399

11.5 Exercises 418

12 Introduction to algorithms 425

12.1 Technical preliminaries 427

12.2 Algorithms for smooth unconstrained minimization 432

12.3 Algorithms for smooth convex constrained minimization 452

12.4 Algorithms for non-smooth convex optimization 472

12.5 Coordinate descent methods 484

12.6 Decentralized optimization methods 487

12.7 Exercises 496

III Applications 503

13 Learning from data 505

13.1 Overview of supervised learning 505

13.2 Least-squares prediction via a polynomial model 507

13.3 Binary classification 511

13.4 A generic supervised learning problem 519

13.5 Unsupervised learning 524

13.6 Exercises 533

14 Computational finance 539

14.1 Single-period portfolio optimization 539

14.2 Robust portfolio optimization 546

14.3 Multi-period portfolio allocation 549

14.4 Sparse index tracking 556

14.5 Exercises 558

15 Control problems 567

15.1 Continuous and discrete time models 568

15.2 Optimization-based control synthesis 571

15.3 Optimization for analysis and controller design 579

15.4 Exercises 586

16 Engineering design 591

16.1 Digital filter design 591

16.2 Antenna array design 600

16.3 Digital circuit design 606

16.4 Aircraft design 609

16.5 Supply chain management 613

16.6 Exercises 622

Index 629

Preface

OPTIMIZATION REFERS TO a branch of applied mathematics concerned with the minimization or maximization of a certain function, possibly under constraints. The birth of the field can perhaps be traced back to an astronomy problem solved by the young Gauss. It matured later with advances in physics, notably mechanics, where natural phenomena were described as the result of the minimization of certain "energy" functions. Optimization has evolved towards the study and application of algorithms to solve mathematical problems on computers.

Today, the field is at the intersection of many disciplines, ranging from statistics, to dynamical systems and control, complexity theory, and algorithms. It is applied to a widening array of contexts, including machine learning and information retrieval, engineering design, economics, finance, and management. With the advent of massive data sets, optimization is now viewed as a crucial component of the nascent field of data science.

In the last two decades, there has been a renewed interest in the field of optimization and its applications. One of the most exciting developments involves a special kind of optimization, convex optimization. Convex models provide a reliable, practical platform on which to build the development of reliable problem-solving software. With the help of user-friendly software packages, modelers can now quickly develop extremely efficient code to solve a very rich library of convex problems. We can now address convex problems with almost the same ease as we solve a linear system of equations of similar size. Enlarging the scope of tractable problems allows us in turn to develop more efficient methods for difficult, non-convex problems.

These developments parallel those that have paved the success of numerical linear algebra. After a series of ground-breaking works on computer algorithms in the late 80s, user-friendly platforms such as Matlab or R, and more recently Python, appeared, and allowed generations of users to quickly develop code to solve numerical problems. Today, only a few experts worry about the actual algorithms and techniques for solving numerically linear systems with a few thousands of variables and equations; the rest of us take the solution, and the algorithms underlying it, for granted.

Optimization, more precisely, convex optimization, is at a similar stage now. For these reasons, most of the students in engineering, economics, and science in general, will probably find it useful in their professional life to acquire the ability to recognize, simplify, model, and solve problems arising in their own endeavors, while only few of them will actually need to work on the details of numerical algorithms. With this view in mind, we titled our book *Optimization Models*, to highlight the fact that we focus on the "art" of understanding the nature of practical problems and of modeling them into solvable optimization paradigms (often, by discovering the "hidden convexity" structure in the problem), rather than on the technical details of an ever-growing multitude of specific numerical optimization algorithms. For completeness, we do provide two chapters, one covering basic linear algebra algorithms, and another one extensively dealing with selected optimization algorithms; these chapters, however, can be skipped without hampering the understanding of the other parts of this book.

Several textbooks have appeared in recent years, in response to the growing needs of the scientific community in the area of convex optimization. Most of these textbooks are graduate-level, and indeed contain a good wealth of sophisticated material. Our treatment includes the following distinguishing elements.

- The book can be used both in undergraduate courses on linear algebra and optimization, and in graduate-level introductory courses on convex modeling and optimization.
- The book focuses on *modeling* practical problems in a suitable optimization format, rather than on *algorithms* for solving mathematical optimization problems; algorithms are circumscribed to two chapters, one devoted to basic matrix computations, and the other to convex optimization.
- About a third of the book is devoted to a self-contained treatment of the essential topic of linear algebra and its applications.
- The book includes many real-world examples, and several chapters devoted to practical applications.
- We do not emphasize general non-convex models, but we do illustrate how convex models can be helpful in solving some specific non-convex ones.

We have chosen to start the book with a first part on linear algebra, with two motivations in mind. One is that linear algebra is perhaps the most important building block of convex optimization. A good command of linear algebra and matrix theory is essential for understanding convexity, manipulating convex models, and developing algorithms for convex optimization.

A second motivation is to respond to a perceived gap in the offering in linear algebra at the undergraduate level. Many, if not most, linear algebra textbooks focus on abstract concepts and algorithms, and devote relatively little space to real-life practical examples. These books often leave the students with a good understanding of concepts and problems of linear algebra, but with an incomplete and limited view about *where* and *why* these problems arise. In our experience, few undergraduate students, for instance, are aware that linear algebra forms the backbone of the most widely used machine learning algorithms to date, such as the PageRank algorithm, used by Google's web-search engine.

Another common difficulty is that, in line with the history of the field, most textbooks devote a lot of space to eigenvalues of general matrices and Jordan forms, which do have many relevant applications, for example in the solutions of ordinary differential systems. However, the central concept of singular value is often relegated to the final chapters, if presented at all. As a result, the classical treatment of linear algebra leaves out concepts that are crucial for understanding linear algebra as a building block of practical optimization, which is the focus of this textbook.

Our treatment of linear algebra is, however, necessarily partial, and biased towards models that are instrumental for optimization. Hence, the linear algebra part of this book is not a substitute for a reference textbook on theoretical or numerical linear algebra.

In our joint treatment of linear algebra and optimization, we emphasize tractable models over algorithms, contextual important applications over toy examples. We hope to convey the idea that, in terms of reliability, a certain class of optimization problems should be considered on the same level as linear algebra problems: reliable models that can be confidently used without too much worry about the inner workings.

In writing this book, we strove to strike a balance between mathematical rigor and accessibility of the material. We favored "operative" definitions over abstract or too general mathematical ones, and practical relevance of the results over exhaustiveness. Most proofs of technical statements are detailed in the text, although some results are provided without proof, when the proof itself was deemed not to be particularly instructive, or too involved and distracting from the context.

Prerequisites for this book are kept at a minimum: the material can be essentially accessed with a basic understanding of geometry and calculus (functions, derivatives, sets, etc.), and an elementary knowledge of probability and statistics (about, e.g., probability distributions, expected values, etc.). Some exposure to engineering or economics may help one to better appreciate the applicative parts in the book.

Book outline

The book starts out with an overview and preliminary introduction to optimization models in Chapter 1, exposing some formalism, specific models, contextual examples, and a brief history of the optimization field. The book is then divided into three parts, as seen from Table 1.

Part I is on linear algebra, Part II on optimization models, and Part III discusses selected applications.

The first part on linear algebra starts with an introduction, in Chapter 2, to basic concepts such as vectors, scalar products, projections, and so on. Chapter 3 discusses matrices and their basic properties, also introducing the important concept of factorization. A fuller story on factorization is given in the next two chapters. Symmetric matrices and their special properties are treated in Chapter 4, while Chapter 5 discusses the singular value decomposition of general matrices, and its applications. We then describe how these tools can be used for solving linear equations, and related least-squares problems, in Chapter 6. We close the linear algebra part in Chapter 7, with a short overview of some classical algorithms. Our presentation in Part I seeks to emphasize the optimization aspects that underpin many linear algebra concepts; for example, projections and the solution of systems of linear equations are interpreted as a basic optimization problem and, similarly, eigenvalues of symmetric matrices result from a "variational" (that is, optimization-based) characterization.

The second part contains a core section of the book, dealing with optimization models. Chapter 8 introduces the basic concepts of convex functions, convex sets, and convex problems, and also focuses on some theoretical aspects, such as duality theory. We then proceed with three chapters devoted to specific convex models, from linear, quadratic, and geometric programming (Chapter 9), to second-order

	1	Introduction
I Linear algebra	2	Vectors
	3	Matrices
	4	Symmetric matrices
	5	Singular value decomposition
	6	Linear equations and least squares
	7	Matrix algorithms
II Convex optimization	8	Convexity
	9	Linear, quadratic, and geometric models
	10	Second-order cone and robust models
	11	Semidefinite models
	12	Introduction to algorithms
III Applications	13	Learning from data
	14	Computational finance
	15	Control problems
	16	Engineering design

Table 1: Book outline.

cone (Chapter 10) and semidefinte programming (Chapter 11). Part II closes in Chapter 12, with a detailed description of a selection of important algorithms, including first-order and coordinate descent methods, which are relevant in large-scale optimization contexts.

A third part describes a few relevant applications of optimization. We included machine learning, quantitative finance, control design, as well as a variety of examples arising in general engineering design.

How this book can be used for teaching

This book can be used as a resource in different kinds of courses.

For a senior-level undergraduate course on *linear algebra and applications*, the instructor can focus exclusively on the first part of this textbook. Some parts of Chapter 13 include relevant applications of linear algebra to machine learning, especially the section on principal component analysis.

For a senior-level undergraduate or beginner graduate-level course on *introduction to optimization*, the second part would become the central component. We recommend to begin with a refresher on basic linear algebra; in our experience, linear algebra is more difficult to teach than convex optimization, and is seldom fully mastered by students. For such a course, we would exclude the chapters on algorithms, both Chapter 7, which is on linear algebra algorithms, and Chapter 12, on optimization ones. We would also limit the scope of Chapter 8, in particular, exclude the material on duality in Section 8.5. For a graduate-level course on *convex optimization*, the main material would be the second part again. The instructor may choose to emphasize the material on duality, and Chapter 12, on algorithms. The applications part can serve as a template for project reports.

Bibliographical references and sources

By choice, we have been possibly incomplete in our bibliographical references, opting to not overwhelm the reader, especially in the light of the large span of material covered in this book. With today's online resources, interested readers can easily find relevant material. Our only claim is that we strove to provide the appropriate search terms. We hope that the community of researchers who have contributed to this fascinating field will find solace in the fact that the success of an idea can perhaps be measured by a lack of proper references.

In writing this book, however, we have been inspired by, and we are indebted to, the work of many authors and instructors. We have drawn in particular from the largely influential textbooks listed on the side.[1] We also give credit to the excellent course material of the courses EE364a, EE364b (S. Boyd), EE365 (S. Lall) at Stanford University, and of EE236a, EE236b, EE236c (L. Vandenberghe) at UCLA, as well as the slides that S. Sra developed for the course EE 227A in 2012 at UC Berkeley.

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
D. P. Bertsekas, *Nonlinear Optimization*, Athena Scientific, 1999.
D. P. Bertsekas (with A. Nedic, A. Ozdaglar), *Convex Analysis and Optimization*, Athena Scientific, 2003.
Yu. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Springer, 2004.
A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*, SIAM, 2001.
J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer, 2006.

Acknowledgments

In the last 20 years, we witnessed many exciting developments in both theory and applications of optimization. The prime stimulus for writing this book came to us from the thriving scientific community involved in optimization research, whose members gave us, directly or indirectly, motivation and inspiration. While it would be impossible to mention all of them, we wish to give special thanks to our colleagues Dimitris Bertsimas, Stephen Boyd, Emmanuel Candès, Constantin Caramanis, Vu Duong, Michael Jordan, Jitendra Malik, Arkadi Nemirovksi, Yuri Nesterov, Jorge Nocedal, Kannan Ramchandran, Anant Sahai, Suvrit Sra, Marc Teboulle, Lieven Vandenberghe, and Jean Walrand, for their support, and constructive discussions over the years. We are also thankful to the anonymous reviewers of our initial draft, who encouraged us to proceed. Special thanks go to Daniel Lyons, who reviewed our final draft and helped improve our presentation.

Our gratitude also goes to Phil Meyler and his team at Cambridge University Press, and especially to Elizabeth Horne for her technical support.

This book has been typeset in Latex, using a variant of Edward Tufte's book style.