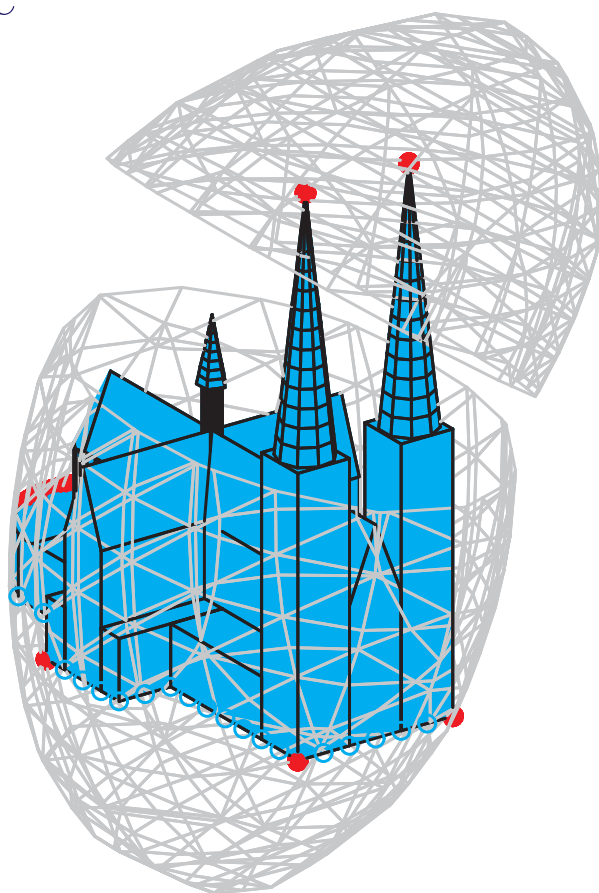


Wilhelm Forst, University of Ulm

Dieter Hoffmann, University of Konstanz

# Optimization

— Theory and Practice



Springer



To my children

*Martin, Christina and Christopher*

WILHELM FORST

To my grandsons

*Léon, Étienne, Gabriel, Nicolas and Luca*

who may some day want to read this book

DIETER HOFFMANN



# Contents

<b>Preface</b> .....	XIII
<b>1 Introduction</b> .....	1
1.1 Examples of Optimization Problems .....	1
Overdetermined Systems of Linear Equations .....	1
Nonlinear Least Squares Problems .....	6
CHEBYSHEV Approximation .....	8
Facility Location Problems .....	10
Standard Form of Optimization Problems .....	14
Penalty Methods .....	16
1.2 Historical Overview .....	20
Optimization Problems in Antiquity .....	21
First Heroic Era: Development of Calculus .....	21
Second Heroic Era: Discovery of Simplex Algorithm .....	26
LEONID KHACHIYAN's Breakthrough .....	27
Exercises .....	29
<b>2 Optimality Conditions</b> .....	35
2.0 Introduction .....	35
2.1 Convex Sets, Inequalities .....	39
2.2 Local First-Order Optimality Conditions .....	44
KARUSH–KUHN–TUCKER Conditions .....	46
Convex Functions .....	52
Constraint Qualifications .....	54
Convex Optimization Problems .....	58

---

2.3	Local Second-Order Optimality Conditions . . . . .	61
2.4	Duality . . . . .	66
	LAGRANGE Dual Problem . . . . .	67
	Geometric Interpretation . . . . .	71
	Saddlepoints and Duality . . . . .	74
	Perturbation and Sensitivity Analysis . . . . .	74
	Economic Interpretation of Duality . . . . .	77
	Strong Duality . . . . .	78
	Exercises . . . . .	80
<b>3</b>	<b>Unconstrained Optimization Problems . . . . .</b>	<b>87</b>
3.0	Logistic Regression . . . . .	88
3.1	Elementary Search and Localization Methods . . . . .	90
	The NELDER and MEAD Polytope Method . . . . .	90
	The SHOR Ellipsoid Method . . . . .	93
3.2	Descent Methods with Line Search . . . . .	98
	Coordinatewise Descent Methods . . . . .	99
	Gradient Method . . . . .	100
	KANTOROVICH's Inequality . . . . .	101
	Requirements on the Step Size Selection . . . . .	104
3.3	Trust Region Methods . . . . .	110
	LEVENBERG–MARQUARDT Trajectory . . . . .	112
	POWELL's Dogleg Trajectory . . . . .	116
	Least Squares Problems . . . . .	118
3.4	Conjugate Gradient Method . . . . .	120
	Generation of $A$ -Conjugate Directions . . . . .	122
	Rate of Convergence . . . . .	125
	The CG-Method in the Nonquadratic Case . . . . .	129
3.5	Quasi-NEWTON Methods . . . . .	130
	Least Change Principle . . . . .	132
	More General Quasi-NEWTON Methods . . . . .	138
	Quasi-NEWTON Methods in the Nonquadratic Case . . . . .	139
	Exercises . . . . .	143

---

<b>4</b>	<b>Linearly Constrained Optimization Problems</b> . . . . .	151
4.1	Linear Optimization . . . . .	152
	The Revised Simplex Method . . . . .	152
	Numerical Realization of the Method . . . . .	155
	Calculation of a Feasible Basis . . . . .	158
	The Active Set Method . . . . .	160
4.2	Quadratic Optimization . . . . .	165
	The BARANKIN–DORFMAN Existence Theorem . . . . .	166
	The Active Set Method . . . . .	168
	Minimization Subject to Linear Equality Constraints . .	175
	The GOLDFARB–IDNANI Method . . . . .	179
4.3	Projection Methods . . . . .	186
	ZOUTENDIJK’s Method . . . . .	188
	Projected Gradient Method . . . . .	189
	Reduced Gradient Method . . . . .	192
	Preview of SQP Methods . . . . .	197
	WILSON’s LAGRANGE–NEWTON Method . . . . .	200
	Exercises . . . . .	204
<b>5</b>	<b>Nonlinearly Constrained Optimization Problems</b> . . . . .	213
5.1	Penalty Methods . . . . .	214
	Classic Penalty Methods (Exterior Penalty Methods) . .	214
	Barrier Methods (Interior Penalty Methods) . . . . .	220
5.2	SQP Methods . . . . .	224
	LAGRANGE–NEWTON Method . . . . .	226
	FLETCHER’s $S\ell_1$ QP Method . . . . .	233
	Exercises . . . . .	237

---

<b>6</b>	<b>Interior-Point Methods for Linear Optimization</b> . . . . .	241
6.1	Linear Optimization II . . . . .	242
	The Duality Theorem . . . . .	243
	The Interior-Point Condition . . . . .	248
6.2	The Central Path . . . . .	252
	Convergence of the Central Path . . . . .	256
6.3	NEWTON's Method for the Primal–Dual System . . . . .	261
6.4	A Primal–Dual Framework . . . . .	263
6.5	Neighborhoods of the Central Path . . . . .	266
	A Closer Look at These Neighborhoods . . . . .	269
6.6	A Short-Step Path-Following Algorithm . . . . .	272
6.7	The MIZUNO–TODD–YE Predictor-Corrector Method . . . . .	277
6.8	The Long-Step Path-Following Algorithm . . . . .	281
6.9	The MEHROTRA Predictor-Corrector Method . . . . .	284
6.10	A Short Look at Interior-Point Methods for Quadratic Optimization . . . . .	289
	Exercises . . . . .	294
<b>7</b>	<b>Semidefinite Optimization</b> . . . . .	299
7.1	Background and Motivation . . . . .	300
	Basics and Notations . . . . .	300
	Primal Problem and Dual Problem . . . . .	304
7.2	Selected Special Cases . . . . .	309
	Linear Optimization and Duality Complications . . . . .	309
	Second-Order Cone Programming . . . . .	313
7.3	The $\mathcal{S}$ -Procedure . . . . .	316
	Minimal Enclosing Ellipsoid of Ellipsoids . . . . .	316
7.4	The Function $\log \circ \det$ . . . . .	323
7.5	Path-Following Methods . . . . .	325
	Primal–Dual System . . . . .	325
	Barrier Functions . . . . .	326



---

Central Path.....	329
7.6 Applying NEWTON's Method .....	330
7.7 How to Solve SDO Problems? .....	332
7.8 Icing on the Cake: Pattern Separation via Ellipsoids .....	334
Exercises .....	337
<b>8 Global Optimization .....</b>	<b>341</b>
8.1 Introduction .....	341
8.2 Branch and Bound Methods .....	347
8.3 Cutting Plane Methods .....	354
Cutting Plane Algorithm by KELLEY .....	354
Concavity Cuts .....	355
Exercises .....	361
<b>Appendices .....</b>	<b>365</b>
A A Second Look at the Constraint Qualifications .....	365
The Linearized Problem .....	365
Correlation to the Constraint Qualifications .....	369
B The FRITZ JOHN Condition.....	370
C Optimization Software Tools for Teaching and Learning .....	374
<i>Matlab</i> <sup>®</sup> Optimization Toolbox.....	374
SeDuMi: An Introduction by Examples.....	377
<i>Maple</i> <sup>®</sup> Optimization Tools.....	381
<b>Index of Symbols .....</b>	<b>383</b>
<b>Subject Index .....</b>	<b>385</b>
<b>Bibliography.....</b>	<b>395</b>



---

## Preface

This self-contained book on optimization is designed to serve a variety of purposes. It will prove useful both as a textbook for undergraduate and first-year graduate-level courses as well as a reference book for mathematicians, engineers and applied scientists interested in a careful exposition of this fascinating and useful branch of mathematics. Students, mathematicians and practitioners alike can profit from an approach which treats central topics of optimization in a concise, comprehensive and modern way. The book is fresh in conception and lucid in style and will appeal to anyone who has a genuine interest in optimization. The mutually beneficial interaction of theory and practice is presented in a stimulating manner. As JOHANN WOLFGANG VON GOETHE states: *“All theory, dear friend, is gray, but the golden tree of life springs ever green.”*

Optimization is not only important in its own right but nowadays forms an integral part of a great number of applied sciences such as operations research, management science, economics and finance, and all branches of math-oriented engineering. Constrained optimization models are used in numerous areas of application and are probably the most widely used mathematical models in operations research and management science.

This book is the outgrowth of many years of teaching optimization in the mathematics departments of the Universities of Konstanz and Ulm (Germany) and W.F.’s teaching experiences with a first English version of parts of this book during his stay as a guest professor at the Universidad Nacional de Trujillo (Peru).

As the title suggests, one major aim of our book is to give a modern and well-balanced treatment of the subject by not only focusing on theory but also including algorithms and instructive examples from different areas of application. We put particular emphasis on presenting the material in a systematic and mathematically convincing way.

We introduce theory and methods at an elementary level but with an accurate and concise formulation of the theoretical tools needed. We phrase ideas and theorems clearly and succinctly, and complement them with detailed proofs and explanations of the ideas behind the proofs. We are convinced that many readers will find our book easy to read and value its accessibility.

Since it is intended as an introduction, the mathematical prerequisites have been kept to a minimum, with only some knowledge of multidimensional calculus, linear algebra and basic numerical methods needed to fully understand the concepts presented in the book. For example, the reader should know a little about the *convergence rate* of iteration methods and have a basic understanding of the *condition number* for matrices, which is a measure of the sensitivity of the error in the solution to a system of linear equations relative to changes in the inputs.

From the wide range of material that could be discussed in optimization lectures, we have selected aspects that every student interested in optimization should know about, as well as some more advanced issues. It is, however, not expected that the whole book will be ‘covered’ in one term. In practice, we have found that a good half is manageable.

This book provides a systematic, thorough and insightful discussion of *optimization of continuous nonlinear problems*. It contains in one volume of reasonable size a very clear presentation of key ideas and basic techniques. The way they are introduced, discussed and illustrated is designed to emphasize the connections between the various concepts.

In addition, we have included several features which we believe will greatly help the readers to get the most out of this book: First of all, every method considered is motivated and explained. Abstract statements are made into working knowledge with the help of detailed explanations on how to proceed in practice.

BENJAMIN FRANKLIN already knew: “*A good example is the best sermon!*” Therefore, the text offers a rich collection of detailed analytical and numerical examples which bring to life central concepts from diverse areas of science and applications. These selected examples have intentionally often been kept simple so that they can still be verified by hand easily. Counterexamples are presented whenever we want to show that certain assumptions cannot simply be dropped.

Additionally, the reader will find many elaborate two-colored graphics which help to facilitate understanding via geometric insight. Often the idea of a result has a simple geometric background — or, as the saying goes: “*A picture is worth a thousand words!*”

Another feature of the book is that it presents the concepts and examples in a way that invites the readers to think for themselves and to critically assess

the observations and methods presented. In short, the book is written for the *active reader*.

Furthermore, we have also included *more than one hundred additional exercises*, which are partly supplemented by hints or *Matlab*<sup>®</sup>/*Maple*<sup>®</sup> code fragments. Here, the student has ample opportunity to practice concepts, statements and procedures, passing from routine problems to demanding extensions of the main text. In these exercises and particularly in appendix C, we will describe relevant features of the *Matlab*<sup>®</sup> *Optimization Toolbox* and demonstrate its use with selected examples.

Today each student uses a notebook computer the way we old guys used slide rules and log tables. Using a computer as a tool in teaching and learning allows one to concentrate on the essential ideas. Nowadays the various opportunities of current computer technology should be part of any lecture on optimization. Nevertheless, we would like to stress that neither knowledge of nor access to *Matlab*<sup>®</sup> or *Maple*<sup>®</sup> is really needed to fully understand this text. On the other hand, readers who follow the proposed way will benefit greatly from our approach.

Since many of the results presented here are nowadays classical — the novelty lies in the arrangement and streamlined presentation — we have not attempted to document the origin of every item. Some of the sources which served as valuable stimuli for our lectures years ago have since fallen into oblivion.

We are, however, aware of the fact that it took many brilliant men numerous years to develop what we teach in one term. In section 1.2, we have therefore compiled a *historical survey* to give our readers an idea of how optimization has developed through the centuries to an important branch of mathematics. Interior-point methods, for example, show that this process is still going on with ever new and exciting results!

We are certain that our selection of topics and their presentation will appeal to students, researchers and practitioners. The emphasis is on insight, not on giving the latest refinements of every method. We introduce the reader to optimization in a gradual and ‘digestible’ manner. Much of the material is ‘classical’ but streamlined proofs are given here in many instances. The systematic organization, structure and clarity of our book together with its insightful illustrations and examples make it an ideal introductory textbook. Written in a concise and straightforward style, this book opens the door to one of the most fascinating and useful branches of mathematics.

The book is divided into eight chapters. Here is a sketch of the *content*:

In *Chapter 1, Examples of Optimization Problems* are given. It is, however, not only intended as a general motivating introduction, but also as a demonstration of the workings and uses of mathematical tools like *Matlab*<sup>®</sup> and

*Maple*<sup>®</sup> and thus serves to encourage our readers not to neglect practice over theory. In particular the graphics features of these tools are important and very helpful in many instances. In addition, the first chapter includes — as already mentioned — a detailed *Historical Overview*.

*Chapter 2* on *Optimality Conditions* studies the key properties of constrained problems and necessary and sufficient optimality conditions for them. As an introduction, we give a summary of the ‘classic’ results for unconstrained and equality constrained problems. In order to formulate optimality conditions for problems with *inequality constraints*, some simple aids from *Convex Analysis* are introduced. The LAGRANGE *multiplier rule* is generalized by the KARUSH–KUHN–TUCKER *conditions*. This core chapter presents the essential tools via a *geometrical point of view* using *cones*.

*Chapter 3* introduces basics on *Unconstrained Optimization Problems*. The corresponding methods seek a local minimum (or maximum) in the absence of restrictions. Optimality criteria are studied and above all algorithmic methods for a wide variety of problems. Even though most optimization problems in ‘real life’ have restrictions to be satisfied, the study of unconstrained problems is useful for two reasons: First, they occur directly in some applications and are thus important in their own right. Second, unconstrained problems often originate as a result of transformations of constrained optimization problems. Some methods, for example, solve a general problem by converting it into a sequence of unconstrained problems.

*Chapter 4* presents *Linearly Constrained Optimization Problems*. Here the problems have general objective functions but linear constraints. Section 4.1 gives a short introduction to *linear optimization* which covers the simplest — yet still highly important — kinds of constrained optimization problems, where the objective function *and* all constraints are linear. We consider two methods as examples of how these types of problems can be solved: the *revised simplex algorithm* and the *active set method*. *Quadratic problems*, which we treat in section 4.2, are linearly constrained with a quadratic objective function. Quadratic optimization is an important field in its own right, since it forms the basis of several algorithms for nonlinearly constrained problems. This section contains, for example, the BARANKIN–DORFMAN existence theorem as well as a lucid description of the GOLDFARB–IDNANI method. In contrast to the active set method this is a *dual method* which has the advantage that it does not need a primally feasible starting point. In section 4.3, we give an outline of selected *projection methods*. Besides classical gradient-based methods we present a sequential quadratic feasible point method.

In the next chapter, *Nonlinearly Constrained Optimization Problems* are treated. In section 5.1, a constrained optimization problem is replaced by an unconstrained one. There are two different approaches to this: In *exterior penalty methods* a term is added to the objective function which ‘penalizes’ a violation of constraints. In *interior penalty methods* a barrier term prevents

points from leaving the interior of the feasible region. In section 5.2, *sequential quadratic programming methods* are introduced. The strategy here is to convert a usually nonlinear problem into a sequence of quadratic optimization problems which are easier to solve.

*Chapter 6* presents *Interior-Point Methods for Linear Optimization*. The development of the last 30 years has been greatly influenced by the aftermath of a “scientific earthquake” triggered in 1979 by the findings of KHACHIYAN (1952–2005) and in 1984 by those of KARMARKAR. Efficient interior-point methods have in the meantime been applied to large classes of nonlinear optimization problems and are still topics of current research.

*Chapter 7* treats basics of *Semidefinite Optimization*. This type of optimization differs from linear optimization in that it deals with problems over the cone of symmetric positive semidefinite matrices  $S_+^n$  instead of nonnegative vectors. It is a branch of *convex optimization* and covers many practically useful problems. The wide range of uses has quickly made semidefinite optimization very popular — besides, of course, the fact that such problems can be solved efficiently via polynomially convergent interior-point methods, which had originally only been developed for *linear* optimization.

The last chapter on *Global Optimization* deals with the computation and characterization of *global* optimizers of — in general — nonlinear functions. It is an important task since many real-world questions lead to global rather than local problems. Although this chapter is relatively short, we are convinced that it will suffice to give our readers a useful and informative introduction to this fascinating topic with all the necessary mathematical precision.

The book concludes with three *Appendices*:

- *A Second Look at the Constraint Qualifications*

The GUIGNARD constraint qualification in section 2.2 seems to somewhat come out of the blue. The correlation between the regularity condition and the corresponding ‘linearized’ problem discussed here makes the matter more transparent.

- *The FRITZ JOHN Condition*

This is — in a certain sense — a weaker predecessor of the KARUSH–KUHN–TUCKER conditions. No regularity condition is required, but in consequence an additional multiplier  $\lambda_0 \geq 0$  must be attached to the objective function. In addition, an arbitrary number of constraints is possible. The striking picture on the title page also falls into this area, the *Minimum Volume Enclosing Ellipsoid of Cologne Cathedral (in an abstract 3D model)*.

- *Optimization Software for Teaching and Learning*

This part of the appendix gives a short overview of the software for the main areas of application in our book. We do not speak about professional software in modeling and solving large optimization problems. In our courses, we use *Matlab*<sup>®</sup> and *Maple*<sup>®</sup> and optimization tools like SeDuMi.

Each chapter starts with a summary, is divided into several sections and ends with numerous exercises which reinforce the material discussed in the chapter. Exercises are carefully chosen both in content and in difficulty to aid understanding and to promote mastery of the subject.

### Acknowledgments

We would like to acknowledge many people who, knowingly or unknowingly, have helped us:

VITA RUTKA and MICHAEL LEHN each held the tutorial accompanying the lecture for one term and provided a number of interesting problems and instructive exercises which we have gladly included in our book.

D.H. would like to thank RAINER JANSSEN for stimulating discussions and an oftentimes helpful look at the working conditions.

We are especially grateful to ANDREAS BORCHERT for his valuable advice concerning our electronic communication between Konstanz and Ulm.

Largely based on [Wri] and [Klerk], JULIA VOGT and CLAUDIA LINDAUER wrote very good master's theses under our supervision which formed the basis of chapters 6 and 7.

The book has also benefited from MARKUS SIGG's careful readings. His comments and suggestions helped to revise a first preliminary version of the text. He also provided valuable assistance to D.H. with some computer problems.

It has been our very good fortune to have had JULIA NEUMANN who translated essential parts of the text and thereby contributed to opening the book to a broader audience. She displayed great skill, courage and intuition in turning — sometimes clumsy — German into good English.

The editorial staff of SPRINGER, especially ANN KOSTANT and ELIZABETH LOEW, deserve thanks for their belief in this special project and the careful attention to our manuscript during the publication process. We would also like to thank the copyeditor who did an excellent job revising our text and providing helpful comments.

Lastly, we would like to say, particularly to experts: We would appreciate personal feedback and welcome any comments, recommendations, spottings of errors and suggestions for improvement. Commendations will, of course, also be accepted!

For additions, updates and *Matlab*<sup>®</sup> and *Maple*<sup>®</sup> sources, please refer to the publisher's website for the book.

Ulm  
Konstanz  
October 2009

WILHELM FORST  
DIETER HOFFMANN