

# Springer Optimization and Its Applications

---

VOLUME 133

---

## *Managing Editor*

Panos M. Pardalos (University of Florida)

## *Editor–Combinatorial Optimization*

Ding-Zhu Du (University of Texas at Dallas)

## *Advisory Board*

J. Birge (University of Chicago)

S. Butenko (Texas A & M University)

S. Rebennack (Karlsruhe Institute of Technology)

F. Giannessi (University of Pisa)

T. Terlaky (Lehigh University)

Y. Ye (Stanford University)

## *Aims and Scope*

Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics and other sciences.

The series *Springer Optimization and Its Applications* publishes undergraduate and graduate textbooks, monographs and state-of-the-art expository works that focus on algorithms for solving optimization problems and also study applications involving such problems. Some of the topics covered include nonlinear optimization (convex and nonconvex), network flow problems, stochastic optimization, optimal control, discrete optimization, multi-objective programming, description of software packages, approximation techniques and heuristic approaches.

More information about this series at <http://www.springer.com/series/7393>

Jan A. Snyman · Daniel N. Wilke

# Practical Mathematical Optimization

Basic Optimization Theory  
and Gradient-Based Algorithms

Second Edition

Jan A. Snyman  
Department of Mechanical  
and Aeronautical Engineering  
University of Pretoria  
Pretoria  
South Africa

Daniel N. Wilke  
Department of Mechanical  
and Aeronautical Engineering  
University of Pretoria  
Pretoria  
South Africa

Additional material to this book can be downloaded from <http://extras.springer.com>.

ISSN 1931-6828                      ISSN 1931-6836 (electronic)  
Springer Optimization and Its Applications  
ISBN 978-3-319-77585-2              ISBN 978-3-319-77586-9 (eBook)  
<https://doi.org/10.1007/978-3-319-77586-9>

Library of Congress Control Number: 2018935221

Mathematics Subject Classification (2010): 65K05, 90C30, 90C26, 90C59

1st edition: © Springer Science+Business Media, Inc. 2005

2nd edition: © Springer International Publishing AG, part of Springer Nature 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG  
part of Springer Nature

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To our wives and friends*  
**Alta and Ella**

# Preface to the second edition

The first edition (2005) of Practical Mathematical Optimization has proved to be a rigorous yet practical introduction to the fundamental principles of mathematical optimization. As stated in the preface to the first edition also included in this new edition, the aim of the text was to equip the reader with the basic theory and algorithms to allow for the solution of practical problems with confidence and in an informed way.

However, since the publication of the first edition more than a decade ago, the complexity of and computing requirements for the solution of mathematical optimization problems have significantly increased. The accessibility and definition of computing platforms have expanded by huge proportions. The fundamental physical limitations on speeding up central processing units have spurred on the advancement of multi-core computing environments that now regularly include graphical processing units. The diversity of software platforms is ever expanding with new domain and computing specific software platforms being released weekly. This edition addresses these recent advancements together with novel ideas already touched on in the first edition that have since matured considerably. They include the handling of noise in objective functions and gradient-only optimization strategies introduced and discussed in the first edition. In order to assist in the coverage of further developments in this area, and in particular of recent work in the application of mainly gradient-only methods to piecewise smooth discontinuous objective functions, it is a pleasure to welcome as co-author for this edition the younger colleague, Daniel N. Wilke.

This second edition of Practical Mathematical Optimization now takes account of the above recent developments and aims to bring this text up to date. Thus, this book now includes a new and separate chapter dedicated to advanced gradient-only formulated solution strategies for optimizing noisy objective functions, specifically piecewise smooth discontinuous objective functions, for which solution formulations and strategies are thoroughly covered. A comprehensive set of alternative solution strategies are presented that include gradient-only line search methods and gradient-only approximations. The application of these strategies is illustrated by application to well-motivated example problems. Also new to this edition is a dedicated chapter on the construction of surrogate models using only zero-order information, zero- and first-order information, and only first-order information. The latter approach being particularly effective in constructing smooth surrogates for discontinuous functions.

A further addition is a chapter dedicated to numerical computation which informs students and practicing scientists and engineers on ways to easily setup and solve problems without delay. In particular, the scientific computing language *Python* is introduced, which is available on almost all computing platforms ranging from dedicated servers and desktops to smartphones. Thus, this book is accompanied by a *Python* module `pmo`, which makes all algorithms presented in this book easily accessible as it follows the well-known `scipy.optimize.minimize` convention. The module is designed to allow and encourage the reader to include their own optimization strategies within a simple, consistent, and systematic framework. The benefit to graduate students and researchers is evident, as various algorithms can be tested and compared with ease and convenience.

To logically accommodate the new material, this edition has been restructured into two parts. The basic optimization theory that covers introductory optimization concepts and definitions, search techniques for unconstrained minimization, and standard methods for constrained optimization is covered in the first five chapters to form Part I. This part contains a chapter of detailed worked-out example problems, while other chapters in Part I are supplemented by example problems and exercises that can be done by hand using only pen, paper, and a calculator. In Part II, the focus shifts to computer applications of relatively new and

mainly gradient-based numerical strategies and algorithms that are covered over four chapters. A dedicated computing chapter using *Python* is included as the final chapter of Part II, and the reader is encouraged to consult this chapter as required to complete the exercises in the preceding three chapters. The chapters in Part II are also supplemented by numerical exercises that are specifically designed so as to encourage the students to plan, execute, and reflect on numerical investigations. In summary, the twofold purpose of these questions is to allow the reader, in the first place, to gain a deeper understanding of the conceptual material presented and, secondly, to assist in developing systematic and scientific numerical investigative skills that are so crucial for the modern-day researcher, scientist, and engineer.

**Jan Snyma and Nico Wilke**

Pretoria

30 January 2018

# Preface to the first edition

It is intended that this book is used in senior- to graduate-level semester courses in optimization, as offered in mathematics, engineering, computer science, and operations research departments. Hopefully, this book will also be useful to practicing professionals in the workplace.

The contents of this book represent the fundamental optimization material collected and used by the author, over a period of more than twenty years, in teaching Practical Mathematical Optimization to undergraduate as well as graduate engineering and science students at the University of Pretoria. The principal motivation for writing this work has not been the teaching of mathematics per se, but to equip students with the necessary fundamental optimization theory and algorithms, so as to enable them to solve practical problems in their own particular principal fields of interest, be it physics, chemistry, engineering design, or business economics. The particular approach adopted here follows from the author's own personal experiences in doing research in solid-state physics and in mechanical engineering design, where he was constantly confronted by problems that can most easily and directly be solved via the judicious use of mathematical optimization techniques. This book is, however, not a collection of case studies restricted to the above-mentioned specialized research areas, but is intended to convey the basic optimization principles and algorithms to a general audience in such a way that, hopefully, the application to their own practical areas of interest will be relatively simple and straightforward.

Many excellent and more comprehensive texts on practical mathematical optimization have of course been written in the past, and I am much indebted to many of these authors for the direct and indirect influence



their work has had in the writing of this monograph. In the text, I have tried as far as possible to give due recognition to their contributions. Here, however, I wish to single out the excellent and possibly underrated book of D. A. Wismer and R. Chattergy (1978), which served to introduce the topic of nonlinear optimization to me many years ago, and which has more than casually influenced this work.

With so many excellent texts on the topic of mathematical optimization available, the question can justifiably be posed: Why another book and what is different here? Here, I believe, for the first time in a relatively brief and introductory work, due attention is paid to certain inhibiting difficulties that can occur when fundamental and classical gradient-based algorithms are applied to real-world problems. Often students, after having mastered the basic theory and algorithms, are disappointed to find that due to real-world complications (such as the presence of noise and discontinuities in the functions, the expense of function evaluations, and an excessive large number of variables), the basic algorithms they have been taught are of little value. They then discard, for example, gradient-based algorithms and resort to alternative non-fundamental methods. Here, in Chapter 4 (now Chapter 6) on new gradient-based methods, developed by the author and his co-workers, the above-mentioned inhibiting real-world difficulties are discussed, and it is shown how these optimization difficulties may be overcome without totally discarding the fundamental gradient-based approach.

The reader may also find the organization of the material in this book somewhat novel. The first three chapters present the basic theory, and classical unconstrained and constrained algorithms, in a straightforward manner with almost no formal statement of theorems and presentation of proofs. Theorems are of course of importance, not only for the more mathematically inclined students, but also for practical people interested in constructing and developing new algorithms. Therefore, some of the more important fundamental theorems and proofs are presented separately in Chapter 6 (now Chapter 5). Where relevant, these theorems are referred to in the first three chapters. Also, in order to prevent cluttering, the presentation of the basic material in Chapters 1 to 3 is interspersed with very few worked-out examples. Instead, a generous number of worked-out example problems are presented separately in Chapter 5 (now Chapter 4), in more or less the same order as the

presentation of the corresponding theory given in Chapters 1 to 3. The separate presentation of the example problems may also be convenient for students who have to prepare for the inevitable tests and examinations. The instructor may also use these examples as models to easily formulate similar problems as additional exercises for the students, and for test purposes.

Although the emphasis of this work is intentionally almost exclusively on gradient-based methods for nonlinear problems, this book will not be complete if only casual reference is made to the simplex method for solving linear programming (LP) problems (where of course use is also made of gradient information in the manipulation of the gradient vector  $\mathbf{c}$  of the objective function, and the gradient vectors of the constraint functions contained in the matrix  $\mathbf{A}$ ). It was therefore decided to include, as Appendix A, a short introduction to the simplex method for LP problems. This appendix introduces the simplex method along the lines given by Chvatal (1983) in his excellent treatment of the subject.

The author gratefully acknowledges the input and constructive comments of the following colleagues to different parts of this work: Nielen Stander, Albert Groenwold, Ken Craig, and Danie de Kock. A special word of thanks goes to Alex Hay. Not only did he significantly contribute to the contents of Chapter 4 (now Chapter 6), but he also helped with the production of most of the figures and in the final editing of the manuscript. Thanks also to Craig Long who assisted with final corrections and to Alna van der Merwe who typed the first L<sup>A</sup>T<sub>E</sub>X draft.

**Jan Snyman**

Pretoria

31 May 2004

# Contents

<b>PREFACE SECOND EDITION</b>	<b>vii</b>
<b>PREFACE FIRST EDITION</b>	<b>xi</b>
<b>TABLE OF NOTATION</b>	<b>xxv</b>

## **PART I: BASIC OPTIMIZATION THEORY**

<b>1 INTRODUCTION</b>	<b>3</b>
1.1 What is mathematical optimization? . . . . .	3
1.2 Objective and constraint functions . . . . .	6
1.3 Basic optimization concepts . . . . .	8
1.3.1 Simplest class of problems: Unconstrained one-dimensional minimization . . .	8
1.3.2 Contour representation of a function of two variables ( $n = 2$ ) . . . . .	9
1.3.3 Contour representation of constraint functions . .	12
	xv

1.3.4	Contour representations of constrained optimization problems . . . . .	13
1.3.5	Simple example illustrating the formulation and solution of an optimization problem . . . . .	14
1.3.6	Maximization . . . . .	16
1.3.7	The special case of Linear Programming . . . . .	16
1.3.8	Scaling of design variables . . . . .	17
1.4	Further mathematical concepts . . . . .	18
1.4.1	Convexity . . . . .	18
1.4.2	Gradient vector of $f(\mathbf{x})$ . . . . .	20
1.4.3	Hessian matrix of $f(\mathbf{x})$ . . . . .	22
1.4.4	The quadratic function in $\mathbb{R}^n$ . . . . .	22
1.4.5	The directional derivative of $f(\mathbf{x})$ in the direction $\mathbf{u}$ . . . . .	23
1.5	Unconstrained minimization . . . . .	23
1.5.1	Global and local minima; saddle points . . . . .	25
1.5.2	Local characterization of the behaviour of a multi-variable function . . . . .	26
1.5.3	Necessary and sufficient conditions for a strong local minimum at $\mathbf{x}^*$ . . . . .	28
1.5.4	General indirect method for computing $\mathbf{x}^*$ . . . . .	30
1.6	Test functions . . . . .	33
1.6.1	Unconstrained . . . . .	33
1.6.2	Constrained . . . . .	35
1.7	Exercises . . . . .	38

<b>2</b>	<b>LINE SEARCH DESCENT METHODS FOR UNCONSTRAINED MINIMIZATION</b>	<b>41</b>
2.1	General line search descent algorithm for unconstrained minimization . . . . .	41
2.1.1	General structure of a line search descent method . . . . .	42
2.2	One-dimensional line search . . . . .	43
2.2.1	Golden section method . . . . .	44
2.2.2	Powell's quadratic interpolation algorithm . . . . .	46
2.3	First order line search descent methods . . . . .	48
2.3.1	The method of steepest descent . . . . .	48
2.3.2	Conjugate gradient methods . . . . .	54
2.4	Second order line search descent methods . . . . .	60
2.4.1	Modified Newton's method . . . . .	61
2.4.2	Quasi-Newton methods . . . . .	61
2.5	Zero order methods and computer optimization subroutines . . . . .	63
2.6	Exercises . . . . .	67
<b>3</b>	<b>STANDARD METHODS FOR CONSTRAINED OPTIMIZATION</b>	<b>71</b>
3.1	Penalty function methods for constrained minimization . . . . .	71
3.1.1	The penalty function formulation . . . . .	71
3.1.2	Illustrative examples . . . . .	72
3.1.3	Sequential unconstrained minimization technique (SUMT) . . . . .	73

3.1.4	Simple example . . . . .	74
3.2	Classical methods for constrained optimization problems . . . . .	76
3.2.1	Equality constrained problems and the Lagrangian function . . . . .	76
3.2.2	Classical approach to optimization with inequality constraints: the KKT conditions . . . . .	84
3.3	Saddle point theory and duality . . . . .	87
3.3.1	Saddle point theorem . . . . .	87
3.3.2	Duality . . . . .	88
3.3.3	Duality theorem . . . . .	88
3.4	Quadratic programming . . . . .	90
3.4.1	Active set of constraints . . . . .	91
3.4.2	The method of Theil and Van de Panne . . . . .	92
3.5	Modern methods for constrained optimization . . . . .	94
3.5.1	The gradient projection method . . . . .	95
3.5.2	Multiplier methods . . . . .	102
3.5.3	Sequential quadratic programming (SQP) . . . . .	107
3.6	Exercises . . . . .	110
4	<b>BASIC EXAMPLE PROBLEMS</b>	<b>113</b>
4.1	Introductory examples . . . . .	113
4.2	Line search descent methods . . . . .	119
4.3	Standard methods for constrained optimization . . . . .	132

4.3.1	Penalty function problems . . . . .	132
4.3.2	The Lagrangian method applied to equality constrained problems . . . . .	135
4.3.3	Solution of inequality constrained problems via auxiliary variables . . . . .	144
4.3.4	Solution of inequality constrained problems via the Karush-Kuhn-Tucker conditions . . . . .	146
4.3.5	Solution of constrained problems via the dual problem formulation . . . . .	154
4.3.6	Quadratic programming problems . . . . .	157
4.3.7	Application of the gradient projection method . . . . .	162
4.3.8	Application of the augmented Lagrangian method . . . . .	164
4.3.9	Application of the sequential quadratic programming method . . . . .	165
<b>5</b>	<b>SOME BASIC OPTIMIZATION THEOREMS</b>	<b>169</b>
5.1	Characterization of functions and minima . . . . .	169
5.2	Equality constrained problem . . . . .	173
5.3	Karush-Kuhn-Tucker theory . . . . .	177
5.4	Saddle point conditions . . . . .	180
5.5	Conjugate gradient methods . . . . .	185
5.6	DFP method . . . . .	189

PART II: GRADIENT-BASED ALGORITHMS

6 NEW GRADIENT-BASED TRAJECTORY AND APPROXIMATION METHODS 197

6.1 Introduction . . . . . 197

6.1.1 Why new algorithms? . . . . . 197

6.1.2 Research at the University of Pretoria . . . . . 198

6.2 The dynamic trajectory optimization method . . . . . 199

6.2.1 Basic dynamic model . . . . . 200

6.2.2 Basic algorithm for unconstrained problems (LFOP) . . . . . 201

6.2.3 Modification for constrained problems (LFOPC) . . . . . 202

6.3 The spherical quadratic steepest descent method . . . . . 205

6.3.1 Introduction . . . . . 205

6.3.2 Classical steepest descent method revisited . . . . . 205

6.3.3 The SQSD algorithm . . . . . 206

6.3.4 Convergence of the SQSD method . . . . . 209

6.3.5 Numerical results and conclusion . . . . . 212

6.3.6 Test functions used for SQSD . . . . . 217

6.4 The Dynamic-Q optimization algorithm . . . . . 218

6.4.1 Introduction . . . . . 218

6.4.2 The Dynamic-Q method . . . . . 219

6.4.3 Numerical results and conclusion . . . . . 223



6.5	A gradient-only line search method for conjugate gradient methods . . . . .	225
6.5.1	Introduction . . . . .	225
6.5.2	Formulation of optimization problem . . . . .	227
6.5.3	Gradient-only line search . . . . .	228
6.5.4	Conjugate gradient search directions and SUMT . . . . .	233
6.5.5	Numerical results . . . . .	234
6.5.6	Conclusion . . . . .	237
6.6	Global optimization using dynamic search trajectories . . . . .	239
6.6.1	Introduction . . . . .	239
6.6.2	The Snyman-Fatti trajectory method . . . . .	241
6.6.3	The modified bouncing ball trajectory method . . . . .	243
6.6.4	Global stopping criterion . . . . .	245
6.6.5	Numerical results . . . . .	248
<b>7</b>	<b>SURROGATE MODELS</b>	<b>251</b>
7.1	Introduction . . . . .	251
7.2	Radial basis surrogate models . . . . .	252
7.2.1	Zero order only radial basis surrogate models . . . . .	253
7.2.2	Combined zero and first order radial basis surrogate models . . . . .	255
7.2.3	First order only radial basis surrogate models . . . . .	257
7.3	Basis functions . . . . .	258

7.3.1	Shape parameter . . . . .	259
7.4	Numerical examples . . . . .	261
7.5	Exercises . . . . .	266
<b>8</b>	<b>GRADIENT-ONLY SOLUTION STRATEGIES</b>	<b>273</b>
8.1	Introduction . . . . .	273
8.2	Piece-wise smooth step discontinuous functions . . . . .	275
8.3	Computational minimization procedures . . . . .	281
8.4	Overview of discontinuities . . . . .	284
8.5	Derivatives and gradients of step discontinuous functions . . . . .	286
8.5.1	Associated gradients by finite differences . . . . .	290
8.6	Gradient-only line search descent methods . . . . .	292
8.6.1	One-dimensional gradient-only line search . . . . .	293
8.6.2	Conditions for sufficient improvement . . . . .	295
8.7	Gradient-only sequential approximate optimization . . . . .	296
8.7.1	Constant diagonal Hessian matrix approximations . . . . .	298
8.7.2	Diagonal Hessian matrix approximations . . . . .	300
8.7.3	Symmetric Hessian matrix approximations . . . . .	300
8.7.4	Symmetric inverse Hessian matrix approximations . . . . .	301
8.7.5	Non-symmetric inverse Hessian matrix approximations . . . . .	302
8.7.6	Trust Region Methods and Conservatism . . . . .	303

8.8	Gradient-only optimization problem . . . . .	306
8.9	Exercises . . . . .	307
<b>9</b>	<b>PRACTICAL COMPUTATIONAL OPTIMIZATION USING PYTHON</b>	<b>311</b>
9.1	Introduction to Python . . . . .	311
9.1.1	Installing <i>Anaconda Python</i> . . . . .	312
9.2	Overview . . . . .	312
9.2.1	Basic Arithmetic . . . . .	315
9.3	Symbolic Mathematics . . . . .	315
9.3.1	Differentiation . . . . .	317
9.3.2	Numerical Evaluation . . . . .	318
9.3.3	Optimality Criteria . . . . .	319
9.4	Numerical Linear Algebra . . . . .	320
9.4.1	Overdetermined system of equations . . . . .	321
9.4.2	Underdetermined system of equations . . . . .	323
9.5	Numerical Functions . . . . .	324
9.6	Optimization Algorithms . . . . .	326
9.6.1	Unconstrained minimization . . . . .	326
9.6.2	Constrained minimization . . . . .	330
9.7	Practical Mathematical Optimization (PMO) Algorithms . . . . .	332
9.7.1	User defined Algorithms . . . . .	334
9.8	Exercises . . . . .	338

<b>A THE SIMPLEX METHOD FOR LINEAR PROGRAMMING PROBLEMS</b>	<b>341</b>
A.1 Introduction . . . . .	341
A.2 Pivoting to increase the objective function . . . . .	343
A.3 Example . . . . .	344
A.4 The auxiliary problem for problem with infeasible origin . . . . .	346
A.5 Example of auxiliary problem solution . . . . .	347
A.6 Degeneracy . . . . .	349
A.7 The revised simplex method . . . . .	350
A.8 An iteration of the RSM . . . . .	352
 <b>Bibliography</b>	 <b>355</b>
 <b>Index</b>	 <b>365</b>

# Table of notation

$\mathbb{R}^n$	$n$ -dimensional Euclidean (real) space
$T$	(superscript only) transpose of a vector or matrix
$\mathbf{x}$	column vector of variables, a point in $\mathbb{R}^n$ $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$
$\in$	element in the set
$f(\mathbf{x}), f$	objective function
$\mathbf{x}^*$	local optimizer
$\mathbf{x}_g^*$	non-negative associated gradient projection point
$f(\mathbf{x}^*)$	optimum function value
$g_j(\mathbf{x}), g_j$	$j^{\text{th}}$ inequality constraint function
$\mathbf{g}(\mathbf{x})$	vector of inequality constraint functions
$h_j(\mathbf{x}), h_j$	$j^{\text{th}}$ equality constraint function
$\mathbf{h}(\mathbf{x})$	vector of equality constraint functions
$C^1$	set of continuous differentiable functions
$C^2$	set of continuous and twice continuous differentiable functions
$\min, \min_{\mathbf{x}}$	minimize w.r.t. $\mathbf{x}$
$\mathbf{x}^0, \mathbf{x}^1, \dots$	vectors corresponding to points 0,1,...
$\{\mathbf{x} \mid \dots\}$	set of elements $\mathbf{x}$ such that ...
$\frac{\partial f}{\partial x_i}$	first partial derivative w.r.t. $x_i$
$\frac{\partial x_i}{\partial \mathbf{h}}$	$= [\frac{\partial h_1}{\partial x_i}, \frac{\partial h_2}{\partial x_i}, \dots, \frac{\partial h_r}{\partial x_i}]^T$
$\frac{\partial x_i}{\partial \mathbf{g}}$	$= [\frac{\partial g_1}{\partial x_i}, \frac{\partial g_2}{\partial x_i}, \dots, \frac{\partial g_m}{\partial x_i}]^T$
$\frac{\partial x_i}{\partial x_i}$	
$\nabla$	first derivative operator
$\nabla_A$	first associated derivative operator

$\nabla f(\mathbf{x})$	gradient vector = $\left[ \frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right]^T$
$\nabla^2$	second derivative operator (elements $\frac{\partial^2}{\partial x_i \partial x_j}$ )
$\mathbf{H}(\mathbf{x}) = \nabla^2 f(\mathbf{x})$	Hessian matrix (second derivative matrix)
$\left. \frac{df(\mathbf{x})}{d\lambda} \right _{\mathbf{u}}$	directional derivative at $\mathbf{x}$ in the direction $\mathbf{u}$
$\subset, \subseteq$	subset of
$ \cdot $	absolute value
$\ \cdot\ $	Euclidean norm of vector
$\cong$	approximately equal
$F(\cdot)$	line search function
$F[, ]$	first order divided difference
$F[, , ]$	second order divided difference
$(\mathbf{a}, \mathbf{b})$	scalar product of vector $\mathbf{a}$ and vector $\mathbf{b}$
$\mathbf{I}$	identity matrix
$\theta_j$	$j^{\text{th}}$ auxiliary variable
$L$	Lagrangian function
$\lambda_j$	$j^{\text{th}}$ Lagrange multiplier
$\boldsymbol{\lambda}$	vector of Lagrange multipliers
$\exists$	exists
$\Rightarrow$	implies
$\{\dots\}$	set
$V[\mathbf{x}]$	set of constraints violated at $\mathbf{x}$
$\phi$	empty set
$\mathcal{L}$	augmented Lagrange function
$\langle a \rangle$	maximum of $a$ and zero
$\frac{\partial \mathbf{h}}{\partial \mathbf{g}}$	$n \times r$ Jacobian matrix = $[\nabla h_1, \nabla h_2, \dots, \nabla h_r]$
$\frac{\partial \mathbf{x}}{\partial \mathbf{g}}$	$n \times m$ Jacobian matrix = $[\nabla g_1, \nabla g_2, \dots, \nabla g_m]$
$s_i$	slack variable
$\mathbf{s}$	vector of slack variables
$D$	determinant of matrix $\mathbf{A}$ of interest in $\mathbf{Ax} = \mathbf{b}$
$D_j$	determinant of matrix $\mathbf{A}$ with $j^{\text{th}}$ column replaced by $\mathbf{b}$
$\lim_{i \rightarrow \infty}$	limit as $i$ tends to infinity