

**Universitext**

# Universitext

---

## Series Editors

Sheldon Axler

*San Francisco State University*

Carles Casacuberta

*Universitat de Barcelona*

Angus MacIntyre

*Queen Mary University of London*

Kenneth Ribet

*University of California, Berkeley*

Claude Sabbah

*École Polytechnique, CNRS, Université Paris-Saclay, Palaiseau*

Endre Süli

*University of Oxford*

Wojbor A. Woźczyński,

*Case Western Reserve University*

*Universitext* is a series of textbooks that presents material from a wide variety of mathematical disciplines at master's level and beyond. The books, often well class-tested by their author, may have an informal, personal even experimental approach to their subject matter. Some of the most successful and established books in the series have evolved through several editions, always following the evolution of teaching curricula, to very polished texts.

Thus as research topics trickle down into graduate-level teaching, first textbooks written for new, cutting-edge courses may make their way into *Universitext*.

More information about this series at <http://www.springer.com/series/223>

J. Frédéric Bonnans

# Convex and Stochastic Optimization

 Springer

J. Frédéric Bonnans  
Inria-Saclay  
and  
Centre de Mathématiques Appliquées  
École Polytechnique  
Palaiseau, France

ISSN 0172-5939

Universitext

ISBN 978-3-030-14976-5

<https://doi.org/10.1007/978-3-030-14977-2>

ISSN 2191-6675 (electronic)

ISBN 978-3-030-14977-2 (eBook)

Library of Congress Control Number: 2019933717

Mathematics Subject Classification (2010): 90C15, 09C25, 90C39, 90C40, 90C46

© Springer Nature Switzerland AG 2019

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.

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

*This book is dedicated to Viviane, Juliette,  
Antoine, and Na Yeong*

# Preface

These lecture notes are an extension of those given in the master programs at the Universities Paris VI and Paris-Saclay, and in the École Polytechnique. They give an introduction to convex analysis and its applications to stochastic programming, i.e., to optimization problems where the decision must be taken in the presence of uncertainties. This is an active subject of research that covers many applications. Classical textbooks are Birge and Louveaux [21], Kall and Wallace [62]. The book [123] by Wallace and Ziemba is dedicated to applications. Some more advanced material is presented in Ruszczyński and Shapiro [105], Shapiro et al. [113], Föllmer and Schied [49], and Carpentier et al. [32]. Let us also mention the historical review paper by Wets [124].

The basic tool for studying such problems is the combination of convex analysis with measure theory. Classical sources in convex analysis are Rockafellar [96], Ekeland and Temam [46]. An introduction to integration and probability theory is given in Malliavin [76].

The author expresses his thanks to Alexander Shapiro (Georgia Tech) for introducing him to the subject, Darinka Dentchev (Stevens Institute of Technology), Andrzej Ruszczyński (Rutgers), Michel de Lara, and Jean-Philippe Chancelier (Ecole des Ponts-Paris Tech) for stimulating discussions, and Pierre Carpentier with whom he shared the course on stochastic optimization in the optimization masters at the Université Paris-Saclay.

Palaiseau, France

J. Frédéric Bonnans

# Contents

<b>1</b>	<b>A Convex Optimization Toolbox</b> . . . . .	1
1.1	Convex Functions . . . . .	1
1.1.1	Optimization Problems . . . . .	1
1.1.2	Separation of Convex Sets . . . . .	4
1.1.3	Weak Duality and Saddle Points . . . . .	10
1.1.4	Linear Programming and Hoffman Bounds . . . . .	11
1.1.5	Conjugacy . . . . .	17
1.2	Duality Theory . . . . .	32
1.2.1	Perturbation Duality . . . . .	32
1.2.2	Subdifferential Calculus . . . . .	44
1.2.3	Minimax Theorems . . . . .	49
1.2.4	Calmness . . . . .	51
1.3	Specific Structures, Applications . . . . .	53
1.3.1	Maxima of Bounded Functions . . . . .	53
1.3.2	Linear Conical Optimization . . . . .	56
1.3.3	Polyhedra . . . . .	57
1.3.4	Infimal Convolution . . . . .	60
1.3.5	Recession Functions and the Perspective Function . . . . .	62
1.4	Duality for Nonconvex Problems . . . . .	65
1.4.1	Convex Relaxation . . . . .	65
1.4.2	Applications of the Shapley–Folkman Theorem . . . . .	70
1.4.3	First-Order Optimality Conditions . . . . .	72
1.5	Notes . . . . .	74
<b>2</b>	<b>Semidefinite and Semi-infinite Programming</b> . . . . .	75
2.1	Matrix Optimization . . . . .	75
2.1.1	The Frobenius Norm . . . . .	75
2.1.2	Positive Semidefinite Linear Programming . . . . .	77

2.2	Rotationally Invariant Matrix Functions . . . . .	80
2.2.1	Computation of the Subdifferential . . . . .	80
2.2.2	Examples . . . . .	84
2.2.3	Logarithmic Penalty . . . . .	85
2.3	SDP Relaxations of Nonconvex Problems . . . . .	87
2.3.1	Relaxation of Quadratic Problems . . . . .	87
2.3.2	Relaxation of Integer Constraints . . . . .	90
2.4	Second-Order Cone Constraints . . . . .	91
2.4.1	Examples of SOC Reformulations . . . . .	91
2.4.2	Linear SOC Duality . . . . .	93
2.4.3	SDP Representation . . . . .	94
2.5	Semi-infinite Programming . . . . .	95
2.5.1	Framework . . . . .	95
2.5.2	Multipliers with Finite Support . . . . .	97
2.5.3	Chebyshev Approximation . . . . .	101
2.5.4	Chebyshev Polynomials and Lagrange Interpolation . . . . .	103
2.6	Nonnegative Polynomials over $\mathbb{R}$ . . . . .	106
2.6.1	Nonnegative Polynomials . . . . .	106
2.6.2	Characterisation of Moments . . . . .	111
2.6.3	Maximal Loading . . . . .	114
2.7	Notes . . . . .	115
<b>3</b>	<b>An Integration Toolbox . . . . .</b>	<b>117</b>
3.1	Measure Theory . . . . .	117
3.1.1	Measurable Spaces . . . . .	117
3.1.2	Measures . . . . .	122
3.1.3	Kolmogorov's Extension of Measures . . . . .	125
3.1.4	Limits of Measurable Functions . . . . .	127
3.1.5	Integration . . . . .	128
3.1.6	$L^p$ Spaces . . . . .	134
3.1.7	Bochner Integrals . . . . .	139
3.2	Integral Functionals . . . . .	141
3.2.1	Minimization of Carathéodory Integrals . . . . .	142
3.2.2	Measurable Multimappings . . . . .	143
3.2.3	Convex Integrands . . . . .	145
3.2.4	Conjugates of Integral Functionals . . . . .	146
3.2.5	Deterministic Decisions in $\mathbb{R}^m$ . . . . .	150
3.2.6	Constrained Random Decisions . . . . .	152
3.2.7	Linear Programming with Simple Recourse . . . . .	153
3.3	Applications of the Shapley–Folkman Theorem . . . . .	156
3.3.1	Integrals of Multimappings . . . . .	156
3.3.2	Constraints on Integral Terms . . . . .	159



- 3.4 Examples and Exercises . . . . . 161
- 3.5 Notes . . . . . 164
- 4 Risk Measures . . . . . 165**
  - 4.1 Introduction . . . . . 165
  - 4.2 Utility Functions . . . . . 165
    - 4.2.1 Framework . . . . . 165
    - 4.2.2 Optimized Utility . . . . . 167
  - 4.3 Monetary Measures of Risk . . . . . 168
    - 4.3.1 General Properties . . . . . 168
    - 4.3.2 Convex Monetary Measures of Risk . . . . . 169
    - 4.3.3 Acceptation Sets . . . . . 170
    - 4.3.4 Risk Trading . . . . . 172
    - 4.3.5 Deviation and Semideviation . . . . . 172
    - 4.3.6 Value at Risk and CVaR . . . . . 174
  - 4.4 Notes . . . . . 176
- 5 Sampling and Optimizing . . . . . 177**
  - 5.1 Examples and Motivation . . . . . 177
    - 5.1.1 Maximum Likelihood . . . . . 177
  - 5.2 Convergence in Law and Related Asymptotics . . . . . 178
    - 5.2.1 Probabilities over Metric Spaces . . . . . 178
    - 5.2.2 Convergence in Law . . . . . 179
    - 5.2.3 Central Limit Theorems . . . . . 185
    - 5.2.4 Delta Theorems . . . . . 186
    - 5.2.5 Solving Equations . . . . . 188
  - 5.3 Error Estimates . . . . . 190
    - 5.3.1 The Empirical Distribution . . . . . 190
    - 5.3.2 Minimizing over a Sample . . . . . 191
    - 5.3.3 Uniform Convergence of Values . . . . . 193
    - 5.3.4 The Asymptotic Law . . . . . 194
    - 5.3.5 Expectation Constraints . . . . . 195
  - 5.4 Large Deviations . . . . . 198
    - 5.4.1 The Principle of Large Deviations . . . . . 198
    - 5.4.2 Error Estimates in Stochastic Programming . . . . . 200
  - 5.5 Notes . . . . . 200
- 6 Dynamic Stochastic Optimization . . . . . 201**
  - 6.1 Conditional Expectation . . . . . 201
    - 6.1.1 Functional Dependency . . . . . 201
    - 6.1.2 Construction of the Conditional Expectation . . . . . 201
    - 6.1.3 The Conditional Expectation of Non-integrable  
Functions . . . . . 206
    - 6.1.4 Computation in Some Simple Cases . . . . . 206

- 6.1.5 Convergence Theorems . . . . . 207
- 6.1.6 Conditional Variance . . . . . 208
- 6.1.7 Compatibility with a Subspace . . . . . 209
- 6.1.8 Compatibility with Measurability Constraints . . . . . 212
- 6.1.9 No Recourse . . . . . 213
- 6.2 Dynamic Stochastic Programming . . . . . 214
  - 6.2.1 Dynamic Uncertainty . . . . . 214
  - 6.2.2 Abstract Optimality Conditions . . . . . 215
  - 6.2.3 The Growing Information Framework . . . . . 217
  - 6.2.4 The Standard full Information Framework . . . . . 218
  - 6.2.5 Independent Noises . . . . . 219
  - 6.2.6 Elementary Examples . . . . . 219
  - 6.2.7 Application to the Turbining Problem . . . . . 220
- 6.3 Notes . . . . . 222
- 7 Markov Decision Processes . . . . . 223**
  - 7.1 Controlled Markov Chains . . . . . 223
    - 7.1.1 Markov Chains . . . . . 223
    - 7.1.2 The Dynamic Programming Principle . . . . . 229
    - 7.1.3 Infinite Horizon Problems . . . . . 231
    - 7.1.4 Numerical Algorithms . . . . . 235
    - 7.1.5 Exit Time Problems . . . . . 239
    - 7.1.6 Problems with Stopping Decisions . . . . . 240
    - 7.1.7 Undiscounted Problems . . . . . 243
  - 7.2 Advanced Material on Controlled Markov Chains . . . . . 243
    - 7.2.1 Expectation Constraints . . . . . 243
    - 7.2.2 Partial Information . . . . . 248
    - 7.2.3 Linear Programming Formulation . . . . . 254
  - 7.3 Ergodic Markov Chains . . . . . 255
    - 7.3.1 Orientation . . . . . 255
    - 7.3.2 Transient and Recurrent States . . . . . 256
    - 7.3.3 Ergodic Dynamic Programming . . . . . 263
  - 7.4 Notes . . . . . 266
- 8 Algorithms . . . . . 267**
  - 8.1 Stochastic Dual Dynamic Programming (SDDP) . . . . . 267
    - 8.1.1 Static Case: Kelley’s Algorithm . . . . . 267
    - 8.1.2 Deterministic Dual Dynamic Programming . . . . . 269
    - 8.1.3 Stochastic Case . . . . . 272
  - 8.2 Introduction to Linear Decision Rules . . . . . 275
    - 8.2.1 About the Frobenius Norm . . . . . 275
    - 8.2.2 Setting . . . . . 276

- 8.2.3 Linear Programming Reformulation . . . . . 276
- 8.2.4 Linear Conic Reformulation . . . . . 277
- 8.2.5 Dual Bounds in a Conic Setting . . . . . 278
- 8.3 Notes . . . . . 281
- 9 Generalized Convexity and Transportation Theory . . . . . 283**
  - 9.1 Generalized Convexity . . . . . 283
    - 9.1.1 Generalized Fenchel Conjugates . . . . . 283
    - 9.1.2 Cyclical Monotonicity . . . . . 285
    - 9.1.3 Duality . . . . . 287
    - 9.1.4 Augmented Lagrangian . . . . . 287
  - 9.2 Convex Functions of Measures . . . . . 289
    - 9.2.1 A First Result . . . . . 290
    - 9.2.2 A Second Result . . . . . 293
  - 9.3 Transportation Theory . . . . . 293
    - 9.3.1 The Compact Framework . . . . . 294
    - 9.3.2 Optimal Transportation Maps . . . . . 296
    - 9.3.3 Penalty Approximations . . . . . 297
    - 9.3.4 Barycenters . . . . . 299
  - 9.4 Notes . . . . . 301
- References . . . . . 303**
- Index . . . . . 309**