

Springer Optimization and Its Applications

VOLUME 85

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)

C.A. Floudas (Princeton University)

F. Giannessi (University of Pisa)

H.D. Sherali (Virginia Polytechnic and State University)

T. Terlaky (McMaster 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 work 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.

For further volumes:

<http://www.springer.com/series/7393>

Michael Zabarankin • Stan Uryasev

Statistical Decision Problems

Selected Concepts and Portfolio Safeguard
Case Studies

 Springer

Michael Zabarankin
Department of Mathematical Sciences
Stevens Institute of Technology
Hoboken, NJ, USA

Stan Uryasev
Department of Industrial and Systems
Engineering
University of Florida
Gainesville, FL, USA

ISSN 1931-6828

ISBN 978-1-4614-8470-7

ISBN 978-1-4614-8471-4 (eBook)

DOI 10.1007/978-1-4614-8471-4

Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013946847

Mathematics Subject Classification (2010): 62C05, 62J05, 62J12, 62P05, 62P30, 90B50, 90B90, 90C05, 90C11, 90C20, 90C26, 90C90, 91B06, 91B30, 91B70

© Springer Science+Business Media New York 2014

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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

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.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Scope

In an abstract form, statistical decision making is an optimization problem that uses available statistical data as an input and optimizes an objective function of interest with respect to decision variables subject to certain constraints. Typically, uncertainty encoded in statistical data can be translated into five basic notions: *likelihood*, *entropy*, *error*, *deviation*, and *risk*. As a result, the majority of statistical decision problems can be tentatively divided into four major categories: (i) likelihood maximization, (ii) entropy maximization (relative entropy minimization), (iii) error minimization (regression), and (iv) decision models in which deviation or risk is either minimized or constrained. All these problems may include so-called technical constraints on decision variables, e.g., box and cardinality constraints. It is also common to optimize one of the corresponding five functionals while to constrain another, e.g., maximizing entropy subject to a constraint on deviation, or to find a trade-off between one of the functionals and the expected value of the quantity of interest. The book aims to demonstrate how to use these “building blocks”: likelihood, entropy, error, deviation, and risk to formulate statistical decision problems arising in various risk management applications, e.g., optimal hedging, portfolio optimization, portfolio replication, cash flow matching, and classification and how to solve those problems in optimization package *Portfolio Safeguard* (PSG).

Content

The book consists of three parts: selected concepts of statistical decision theory (Part I), statistical decision problems (Part II), and case studies with PSG (Part III). Part I presents a general theory of error, deviation, and risk measures to be used in various statistical decision problems and also discusses probabilistic inequalities with deviation measures such as generalized Chebyshev’s and Kolmogorov’s

inequalities. Part II covers five major topics: parametric and nonparametric estimation based on the maximum likelihood principle, entropy maximization problems, unconstrained and constrained linear regression with general error measures, classification with logistic regression and with support vector machines (SVMs), and statistical decision problems with general deviation and risk measures. Part III discusses 21 case studies of typical statistical decision problems arising in risk management, particularly in financial engineering, and demonstrates implementation of those problems in PSG. All case studies are closely related to theoretical Part II and are examples of statistical decision problems from the four categories (i)–(iv).

Audience

The book is aimed at practitioners working in the areas of risk management, decision making under uncertainty, and statistics. It can serve as a quick introduction into the theory of general error, deviation, and risk measures for the graduate students, engineers, and statisticians interested in modeling and managing risk in various applications such as optimal hedging, portfolio replication, portfolio optimization, cash flow matching, structuring of collateralized debt obligations (CDOs), classification, sparse signal reconstruction, and therapy treatment optimization, to mention just a few. It can also be used as a supplementary reading for a number of graduate courses including but not limited to those of statistical analysis, models of risk, data mining, stochastic programming, financial engineering, modern portfolio theory, and advanced engineering economy.

Optimization Software: Portfolio Safeguard

PSG is an advanced nonlinear mixed-integer optimization package for solving a wide range of optimization, statistics, and risk management problems. PSG is a product of American Optimal Decisions, Inc. (see www.aorda.com). Although PSG is a general-purpose decision support tool, the focus application areas are risk management, financial engineering, military, and medical applications. PSG is based on a simple but powerful idea: for every engineering area, identify most commonly used nonlinear functions and include them in the package as independent built-in objects. Each function is defined by a function type, parameters, and a matrix of data (e.g., scenario matrix or covariance matrix). Specialized algorithms, built for different types of functions, efficiently optimize large-scale nonlinear functions, such as probability, value-at-risk (VaR), and omega functions, which are typically beyond the scope of commercial packages. The built-in function library provides simple and convenient interface for evaluating functions and their derivatives, for constructing optimization problems and solving them, and for analyzing solutions. No programming experience is required to use PSG.

PSG operates in four programming environments: Shell (Windows), MATLAB, C++, and Run-File (Text). The standard PSG setup includes case studies from

various areas with emphasis on financial engineering applications, such as portfolio optimization, asset allocation, selection of insurance, hedging with derivative contracts, bond matching, and structuring of CDOs.

PSG can be downloaded from the American Optimal Decisions web site: www.aorda.com/aod. Four types of licenses are available: Freeware Express, Regular, Academic, and Regular Business. Freeware Express Edition limits the number of decision variables per function to ten. The Regular PSG edition has a free 30-day trial. After installing PSG, case study projects can be viewed in the Case Studies folder in File tab of the PSG menu. In order to modify an existing case study, it should be copied into Work directory. Tutorials about PSG and case study descriptions can be found in Help tab of the menu.

Acknowledgments

We thank Dr. Bogdan Grechuk at the University of Leicester, UK, and Dr. Anton Molyboha at Teza Technologies for carefully proofreading the portions of the manuscript and for providing valuable comments. We are also grateful to Dr. Viktor Kuzmenko at American Optimal Decisions, Inc. for the technical support and assistance with running numerical experiments for the case studies. We express our deepest gratitude to our families for their constant encouragement, support, and patience.

Hoboken, NJ, USA
Gainesville, FL, USA

Michael Zabrankin
Stan Uryasev

Contents

Part I Selected Concepts of Statistical Decision Theory

1	Random Variables	3
1.1	Probability Distribution Function	3
1.2	Moments and Moment-Generating Function	5
1.2.1	Classical Inequalities with Moments	6
1.2.2	Moment-Generating Function	7
1.3	Partial Moments and Stochastic Dominance	8
1.3.1	Performance Functions	9
1.3.2	Stochastic Dominance	9
1.4	Quantile Functions and Fenchel Duality	10
1.5	Entropy and Distribution Divergence Measures	13
1.5.1	Shannon Entropy and Kullback–Leibler Divergence	13
1.5.2	Renyi Entropy and Renyi Divergence	15
2	Deviation, Risk, and Error Measures	19
2.1	Deviation Measures	21
2.2	Risk Envelopes and Risk Identifiers	23
2.3	Averse Measures of Risk	26
2.4	Error Measures	29
3	Probabilistic Inequalities	33
3.1	Basic Probabilistic Inequalities	33
3.2	Chebyshev’s Inequalities with Deviation Measures	35
3.2.1	One-Sided Chebyshev’s Inequalities	35
3.2.2	Two-Sided Chebyshev’s Inequalities	39
3.3	Kolmogorov’s Inequalities with Deviation Measures	40

Part II Statistical Decision Problems

4	Maximum Likelihood Method	45
4.1	Maximum Likelihood Principle	45
4.1.1	Parametric Estimation	45

4.1.2	Nonparametric Estimation	48
4.1.3	Estimation of Probability Distributions of Random Vectors	50
5	Entropy Maximization	53
5.1	Shannon Entropy Maximization	53
5.2	Relative Entropy Minimization	58
5.3	Renyi Entropy Maximization	59
5.4	Entropy Maximization with Deviation Measures	62
5.5	Entropy Maximization with Constraints on Average α -Quantile for Different Confidence Levels.....	67
6	Regression Models	71
6.1	Generalized Linear Regression and Error Decomposition.....	73
6.2	Least Squares Linear Regression	74
6.2.1	Unconstrained Least Squares Linear Regression	74
6.2.2	Least Squares Linear Regression with Constraints	75
6.3	Median Regression	77
6.3.1	Unconstrained Median Regression	77
6.3.2	Median Regression with Constraints	78
6.4	Quantile Regression	81
6.4.1	Unconstrained Quantile Regression	81
6.4.2	Quantile Regression with Constraints.....	83
6.5	Special Types of Linear Regression	83
6.5.1	Mixed Quantile Regression.....	83
6.5.2	Unbiased Linear Regression.....	84
6.5.3	Risk Acceptable Regression	85
6.5.4	Robust Regression	86
7	Classification	89
7.1	Logistic Regression	89
7.2	Support Vector Machine	92
7.2.1	Standard SVM.....	92
7.2.2	SVM with Soft Margin	93
7.2.3	Alternative Formulations of SVM with Soft Margin	95
7.2.4	Choice of Mapping ϕ : Kernel Approach.....	97
8	Statistical Decision Models with Risk and Deviation	101
8.1	Minimization of Risk and Deviation.....	101
8.2	Deviation Minimization in Portfolio Selection	104
8.2.1	Markowitz Portfolio Selection	104
8.2.2	Markowitz-Type Portfolio Selection	106
8.2.3	Generalized Capital Asset Pricing Model.....	109
8.2.4	CAPM Reinterpretation: Inverse Portfolio Problem	113
8.3	Deviation Minimization in Optimal Hedging	115
8.4	Cash Matching with Deviation Constraints	116
8.5	Risk Minimization in Portfolio Theory	116

- 8.6 Applications with Downside Risk Constraints 119
- 8.7 Portfolio Optimization with Drawdown Measure 122
 - 8.7.1 Drawdown Measure 122
 - 8.7.2 Linear Programming Representations 124
 - 8.7.3 Portfolio Problem Formulations 125
 - 8.7.4 CAPM with CDaR 127

Part III Portfolio Safeguard Case Studies

- 9 Portfolio Safeguard Case Studies 133**
 - 9.1 Case Study: Optimal Hedging of CDO Book
(meanabs_dev, meanabs_pen, polynom_abs, cardn) 133
 - 9.1.1 Problem 1: problem_cdohedge_1 135
 - 9.1.2 Problem 2: problem_cdohedge_2 136
 - 9.2 Hedging Portfolio of Options 137
 - 9.3 Mortgage Pipeline Hedging 142
 - 9.3.1 Problem 1: problem_cvar_dev 144
 - 9.3.2 Problem 2: problem_meanabs_dev 145
 - 9.3.3 Problem 3: problem_std_dev 146
 - 9.3.4 Problem 4: problem_two_tailed_var75 147
 - 9.3.5 Problem 5: problem_two_tailed_var90 147
 - 9.4 Cash Matching Bond Portfolio 148
 - 9.5 Cash Flow Matching with CVaR Constraints 152
 - 9.6 Relative Entropy Minimization 157
 - 9.7 Portfolio Replication with Risk Constraints 159
 - 9.8 Style Classification with Quantile Regression 162
 - 9.9 CVaR Estimation Through Explanatory Factors with
Mixed Quantile Regression 165
 - 9.10 Optimal Test Selection 169
 - 9.10.1 Problem 1: problem_deterministic_LP_model 170
 - 9.10.2 Problem 2: problem_robust_model 171
 - 9.10.3 Problem 3: problem_stochastic_model 172
 - 9.11 Sparse Signal Reconstruction: A Cardinality Approach 173
 - 9.11.1 Problem 1: problem_constr_cardinality 175
 - 9.11.2 Problem 2: problem_minimize_cardinality 176
 - 9.11.3 Problem 3: problem_constr_polynomabs 178
 - 9.12 Sparse Reconstruction Problems from SPARCO Toolbox 179
 - 9.12.1 Problem 1: L1 Relaxed 181
 - 9.12.2 Problem 2: L1 Relaxed D 182
 - 9.12.3 Problem 3: L2 D 183
 - 9.13 Optimal Intensity-Modulated Radiation Therapy
Treatment-Planning Problem 184
 - 9.14 Portfolio Optimization—CVaR Versus Standard Deviation 188
 - 9.14.1 Problem 1: problem_st_dev_covariances_2p9 190
 - 9.14.2 Problem 2: problem_st_dev_scenarios_2p9 191
 - 9.14.3 Problem 3: problem_min_cvar_dev_2p9 192

- 9.15 Portfolio Credit-Risk Optimization Modeled by Scenarios and Mixtures of Normal Distributions 193
 - 9.15.1 Problem 1: problem_var_LLN 197
 - 9.15.2 Problem 2: problem_cvar_LLN 198
 - 9.15.3 Problem 3: problem_avg_var_CLT 199
 - 9.15.4 Problem 4: problem_avg_cvar_CLT 200
 - 9.15.5 Problem 5: problem_avg_var_CLT (Alternative Formulation) 201
 - 9.15.6 Problem 6: problem_avg_cvar_CLT (Alternative Formulation)..... 203
- 9.16 ν -SVM with VaR and CVaR 204
 - 9.16.1 Problem 1a: ν -SVM with CVaR..... 206
 - 9.16.2 Problem 1b: VaR-SVM 207
 - 9.16.3 Problem 2a: Extended ν -SVM with CVaR 208
 - 9.16.4 Problem 2b: Extended VaR-SVM 209
 - 9.16.5 Problem 3a: Robust ν -SVM with CVaR 210
 - 9.16.6 Problem 3b: Robust VaR-SVM..... 211
- 9.17 Implied Copula CDO Pricing Model: Entropy Approach 212
 - 9.17.1 Problem 1a: problem_hull_meansquare 214
 - 9.17.2 Problem 1b: problem_hull_variance 215
 - 9.17.3 Problem 2: problem_CCC 215
- 9.18 Logistic Regression and Regularized Logistic Regression for Estimating the Probability of Cesarean Section 217
 - 9.18.1 Problem 1: Logistic Regression 218
 - 9.18.2 Problem 2: Regularized Logistic Regression 219
- 9.19 Portfolio Optimization with Drawdown Constraints on a Single Sample Path 220
 - 9.19.1 Problem 1: Constraint on Maximum Drawdown 222
 - 9.19.2 Problem 2: Constraint on Average Drawdown 223
 - 9.19.3 Problem 3: Constraint on CDaR 224
- 9.20 Portfolio Optimization with Drawdown Constraints on Multiple Sample Paths 225
 - 9.20.1 Problem 1: Constraint on Maximum Drawdown 228
 - 9.20.2 Problem 2: Constraint on Average Drawdown 230
 - 9.20.3 Problem 3: Constraint on CDaR 231
- 9.21 Portfolio Optimization with Drawdown Constraints: Single Path Versus Multiple Paths 233
 - 9.21.1 Problem 1: Multiple Sample Paths 236
 - 9.21.2 Problem 2: Aggregated Sample Path 238
- References**..... 241
- Index**..... 245

Nomenclature

\sim	“has distribution”
$[x]_+$	$\max\{x, 0\}$ for real-valued x
$\partial\mathcal{F}(X)$	subdifferential of convex functional \mathcal{F} at X
$\succeq_{(n)}$	non-strict preference relation with respect to n^{th} -order stochastic dominance
$\ X\ _p$	p -norm of X in \mathcal{L}^p , $p \in [1, \infty]$
\mathcal{A}	acceptance set
$\text{AvgDD}(\xi)$	average drawdown of sample path ξ
β_i	beta of asset $i \in \{1, \dots, n\}$ in the capital asset pricing model (CAPM)
\mathcal{C}	set of parameters c_0, c_1, \dots, c_n
$\text{CDaR}_\alpha(\xi)$	conditional drawdown-at-risk (CDaR) of sample path ξ
$\text{Cov}(X, Y)$	covariance of X and Y
$\text{CVaR}_\alpha(X)$	conditional value-at-risk (CVaR) of X
$\text{CVaR}_\alpha^\Delta(X)$	conditional value-at-risk (CVaR) deviation of X
$\text{CVaR}_\alpha^-(L)$	conditional value-at-risk (CVaR) for loss L
$\mathcal{D}(X)$	deviation measure \mathcal{D} of X
$D_{\text{KL}}(X\ Y)$	Kullback–Leibler divergence measure (relative entropy) of X and Y
$\mathcal{E}(X)$	error measure \mathcal{E} of X
$E[X]$	expected value of X
e	n -dimensional unit vector $(1, \dots, 1)$
$F_X(t)$	cumulative distribution function (CDF) of X
$f_X(t)$	probability density function (PDF) of X
$F_X^{-1}(s)$	inverse of the cumulative distribution function (CDF) of X
$\Phi(t)$	cumulative distribution function (CDF) of the standard normal random variable
$\Gamma(x)$	Gamma function of real-valued x
$H_\alpha(X)$	Renyi entropy of X
$I_{\{\cdot\}}$	indicator function (1 if the condition in the curly brackets is true and 0 otherwise)
\inf	essential infimum of X ($\text{ess inf } X$)

\mathcal{L}^p	Lebesgue space of integrable functions, $p \in [1, \infty]$
Λ	covariance matrix
Λ^{-1}	inverse of covariance matrix Λ
\mathcal{M}	sigma-algebra over space of probability events
$M_X(t)$	moment generating function (MGF) of X
$\text{MAD}(X)$	mean absolute deviation (MAD) of X
$\text{MaxDD}(\xi)$	maximum drawdown of sample path ξ
$\text{med}(X)$	median of X
μ	mean value of X
μ_M	expected rate of return of master fund (market portfolio)
\mathbb{N}	set of positive integer numbers
$N(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$N_{\mathcal{X}}(X)$	normal cone of nonempty, closed and convex set \mathcal{X} at $X \in \mathcal{X}$
Ω	probability space of elementary events
\mathbb{P}	probability measure
\mathcal{Q}	risk envelope
$\mathcal{Q}_{\mathcal{D}}(X)$	set of risk identifiers for X with respect to deviation \mathcal{D}
$q_X^+(s)$	upper s -quantile of X
$q_X^-(s)$	lower s -quantile of X
$q_X(s)$	s -quantile interval $[q_X^-(s), q_X^+(s)]$
$\bar{q}_X(s)$	average s -quantile of X
$\mathcal{R}(X)$	risk measure \mathcal{R} of X
\mathbb{R}	set of real numbers
\mathbb{R}^+	set of positive real numbers
\mathbb{R}_0^+	set of nonnegative real numbers
\mathbb{R}^n	set of n -dimensional real vectors
r_0	risk-free rate of return
r_i	rate of return of risky asset $i \in \{1, \dots, n\}$
r_M	rate of return of master fund (market portfolio)
$\text{rc}\mathcal{X}$	recession cone of nonempty, closed and convex set \mathcal{X}
$\mathcal{S}(X)$	statistic of X associated with error \mathcal{E}
$S(X)$	Shannon entropy of X
$\sigma(X)$	standard deviation of X
$\sigma^2(X)$	variance of X
$\sigma_+(X)$	standard upper semideviation of X
$\sigma_-(X)$	standard lower semideviation of X
σ_M^2	variance of the rate of return of master fund (market portfolio)
sup	essential supremum of X (ess sup X)
$\hat{\theta}$	statistical estimator of parameter θ
$U(0, 1)$	uniform probability distribution on $[0, 1]$
$\text{VaR}_\alpha(X)$	value-at-risk (VaR) of X
$\text{VaR}_\alpha^\Delta(X)$	value-at-risk (VaR) deviation of X
X, Y, Z	random variables
\mathcal{X}	feasible set of X