

Springer Series in Operations Research

Editors:

Peter Glynn Stephen Robinson

Springer

New York

Berlin

Heidelberg

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo

Springer Series in Operations Research

Altiok: Performance Analysis of Manufacturing Systems

Birge and Louveaux: Introduction to Stochastic Programming

Bonnans and Shapiro: Perturbation Analysis of Optimization Problems

Bramel and Simchi-Levi: The Logic of Logistics: Theory, Algorithms,
and Applications for Logistics Management

Dantzig and Thapa: Linear Programming 1: Introduction

Drezner (Editor): Facility Location: A Survey of Applications and
Methods

Fishman: Discrete-Event Simulation: Modeling, Programming, and Analysis

Fishman: Monte Carlo: Concepts, Algorithms, and Applications

Nocedal and Wright: Numerical Optimization

Olson: Decision Aids for Selection Problems

Yao (Editor): Stochastic Modeling and Analysis of Manufacturing
Systems

John R. Birge François Louveaux

Introduction to Stochastic Programming

With 38 Illustrations



Springer

John R. Birge
McCormick School of Engineering
and Applied Science
Northwestern University
Evanston, IL 60208
USA

François Louveaux
Facultés Universitaires Notre Dame
de la Paix
Département des Méthodes Quantitatives
Rempart de la Vierge 8
B-5000 Namur
Belgium

Series Editors:

Peter Glynn
Department of Operations Research
Stanford University
Stanford, CA 94305
USA

Stephen Robinson
Department of Industrial Engineering
University of Wisconsin
Madison, WI 53786-1572
USA

Library of Congress Cataloging-in-Publication Data

Birge, John R.

Introduction to stochastic programming / John R. Birge, François Louveaux.

p. cm.

Includes bibliographical references and index.

ISBN 0-387-98217-5 (hardcover : alk. paper)

1. Stochastic programming. I. Louveaux, François. II. Title.

T57.79.B57 1997

519.7—dc21

97-6931

© 1997 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

To Pierrette and Marie

Preface

According to a French saying “Gérer, c’est prévoir,” which we may translate as “(The art of) Managing is (in) foreseeing.” Now, probability and statistics have long since taught us that the future cannot be perfectly forecast but instead should be considered random or uncertain. The aim of stochastic programming is precisely to find an optimal decision in problems involving uncertain data. In this terminology, *stochastic* is opposed to *deterministic* and means that some data are random, whereas programming refers to the fact that various parts of the problem can be modeled as linear or nonlinear mathematical programs. The field, also known as *optimization under uncertainty*, is developing rapidly with contributions from many disciplines such as operations research, economics, mathematics, probability, and statistics. The objective of this book is to provide a wide overview of stochastic programming, without requiring more than a basic background in these various disciplines.

Introduction to Stochastic Programming is intended as a first course for beginning graduate students or advanced undergraduate students in such fields as operations research, industrial engineering, business administration (in particular, finance or management science), and mathematics. Students should have some basic knowledge of linear programming, elementary analysis, and probability as given, for example, in an introductory book on operations research or management science or in a combination of an introduction to linear programming (optimization) and an introduction to probability theory.

Instructors may need to add some material on convex analysis depending on the choice of sections covered. We chose not to include such introductory

material because students' backgrounds may vary widely and other texts include these concepts in detail. We did, however, include an introduction to random variables while modeling stochastic programs in Section 2.1 and short reviews of linear programming, duality, and nonlinear programming at the end of Chapter 2. This material is given as an indication of the prerequisites in the book to help instructors provide any missing background. In the Subject Index, the first reference to a concept is where it is defined or, for concepts specific to a single section, where a source is provided.

In our view, the objective of a first course based on this book is to help students build an intuition on how to model uncertainty into mathematical programs, which changes uncertainty brings into the decision process, what difficulties uncertainty may bring, and what problems are solvable. To begin this development, the first section in Chapter 1 provides a worked example of modeling a stochastic program. It introduces the basic concepts, without using any new or specific techniques. This first example can be complemented by any one of the other proposed cases of Chapter 1, in finance, in multistage capacity expansion, and in manufacturing. Based again on examples, Chapter 2 describes how a stochastic model is formally built. It also stresses the fact that several different models can be built, depending on the type of uncertainty and the time when decisions must be taken. This chapter links the various concepts to alternative fields of planning under uncertainty.

Any course should begin with the study of those two chapters. The sequel would then depend on the students' interests and backgrounds. A typical course would consist of elements of Chapter 3, Sections 4.1 to 4.5, Sections 5.1 to 5.3 and 5.7, and one or two more advanced sections of the instructor's choice. The final case study may serve as a conclusion. A class emphasizing modeling might focus on basic approximations in Chapter 9 and sampling in Chapter 10. A computational class would stress methods from Chapters 6 to 8. A more theoretical class might concentrate more deeply on Chapter 3 and the results from Chapters 9 to 11.

The book can also be used as an introduction for graduate students interested in stochastic programming as a research area. They will find a broad coverage of mathematical properties, models, and solution algorithms. Broad coverage cannot mean an in-depth study of all existing research. The reader will thus be referred to the original papers for details. Advanced sections may require multivariate calculus, probability measure theory, or an introduction to nonlinear or integer programming. Here again, the stress is clearly in building knowledge and intuition in the field. Mathematical results are given so long as they are either basic properties or helpful in developing efficient solution procedures. The importance of the various sections clearly reflects our own interests, which focus on results that may lead to practical applications of stochastic programming.

To conclude, we may use the following little story. An elderly person, celebrating her one hundredth birthday, was asked how she succeeded in reaching that age. She answered, “It’s very simple. You just have to wait.”

In comparison, stochastic programming may well look like a field of young impatient people who not only do not want to wait and see but who consider waiting to be suboptimal. We realize how much patience was needed from our friends and colleagues who encouraged us to write this book, which took us much longer than expected. To all of them, we are extremely thankful for their support. The authors also wish to thank the Fonds National de la Recherche Scientifique and the National Science Foundation for their financial support. Both authors are deeply grateful to the people who introduced us to the field, George Dantzig, Roger Wets, Jacques Drèze, and Guy de Ghellinck. Our special thanks go to our wives, Pierrette and Marie, to whom we dedicate this book.

Ann Arbor, Michigan
Namur, Belgium

John R. Birge
François Louveaux

Contents

Preface	vii
Notation	xv
I Models	1
1 Introduction and Examples	3
1.1 A Farming Example and the News Vendor Problem	4
1.2 Financial Planning and Control	20
1.3 Capacity Expansion	28
1.4 Design for Manufacturing Quality	37
1.5 Other Applications	42
2 Uncertainty and Modeling Issues	49
2.1 Probability Spaces and Random Variables	49
2.2 Deterministic Linear Programs	51
2.3 Decisions and Stages	52
2.4 Two-Stage Program with Fixed Recourse	54
2.5 Random Variables and Risk Aversion	61
2.6 Implicit Representation of the Second Stage	63
2.7 Probabilistic Programming	64
2.8 Relationship to Other Decision-Making Models	67
2.9 Short Reviews	73

II	Basic Properties	81
3	Basic Properties and Theory	83
3.1	Two-Stage Stochastic Linear Programs with Fixed Recourse	84
3.2	Probabilistic or Chance Constraints	103
3.3	Stochastic Integer Programs	109
3.4	Two-Stage Stochastic Nonlinear Programs with Recourse	122
3.5	Multistage Stochastic Programs with Recourse	128
4	The Value of Information and the Stochastic Solution	137
4.1	The Expected Value of Perfect Information	137
4.2	The Value of the Stochastic Solution	139
4.3	Basic Inequalities	140
4.4	The Relationship between EVPI and VSS	141
4.5	Examples	144
4.6	Bounds on EVPI and VSS	145
III	Solution Methods	153
5	Two-Stage Linear Recourse Problems	155
5.1	The L-Shaped Method	156
5.2	Feasibility	163
5.3	The Multicut Version	166
5.4	Bunching and Other Efficiencies	169
5.5	Inner Linearization Methods	174
5.6	Basis Factorization Methods	179
5.7	Special Cases—Simple Recourse and Network Problems	192
6	Nonlinear Programming Approaches to Two-Stage Recourse Problems	199
6.1	Regularized Decomposition	199
6.2	The Piecewise Quadratic Form of the L-Shaped Method	206
6.3	Methods Based on the Stochastic Program Lagrangian	215
6.4	Nonlinear Programming in Simple Recourse Problems	225
6.5	Other Nonlinear Programming-Based Methods	231
7	Multistage Stochastic Programs	233
7.1	Nested Decomposition Procedures	234
7.2	Quadratic Nested Decomposition	244
7.3	Other Approaches to Multiple Stages	251
8	Stochastic Integer Programs	253
8.1	Integer L-Shaped Method	253
8.2	Simple Integer Recourse	262

8.3	Binary First-Stage Variables	268
8.4	Other Approaches	276
IV	Approximation and Sampling Methods	283
9	Evaluating and Approximating Expectations	285
9.1	Direct Solutions with Multiple Integration	286
9.2	Discrete Bounding Approximations	288
9.3	Using Bounds in Algorithms	296
9.4	Bounds in Chance-Constrained Problems	301
9.5	Generalized Bounds	305
9.6	General Convergence Properties	323
10	Monte Carlo Methods	331
10.1	General Results for Sampled Problems	332
10.2	Using Sampling in the L-Shaped Method	335
10.3	Stochastic Quasi-Gradient Methods	343
10.4	Sampling Extensions: Uses with Analytical and Empirical Observations	349
11	Multistage Approximations	353
11.1	Bounds Based on the Jensen and Edmundson-Madansky Inequalities	354
11.2	Bounds Based on Aggregation	359
11.3	Bounds Based on Separable Responses	362
11.4	Bounds for Specific Problem Structures	366
V	A Case Study	373
12	Capacity Expansion	375
12.1	Model Development	376
12.2	Demand Distribution Modeling	382
12.3	Computational Comparisons	383
12.4	Results and Extensions	383
A	Sample Distribution Functions	385
A.1	Discrete Random Variables	385
A.2	Continuous Random Variables	386
	References	387
	Author Index	411
	Subject Index	415

Notation

The following describes the major symbols and notations used in the text. To the greatest extent possible, we have attempted to keep unique meanings for each item. In those cases where an item has additional uses, they should be clear from context. We include here only notation used in more than one section. Additional notation may be needed within specific sections and is explained when used.

In general, vectors are assumed to be columns with transposes to indicate row vectors. This yields $c^T x$ to denote the inner product of two n -vectors, c and x . We reserve prime (\prime) for first derivatives with respect to time (e.g., $f' = df/dt$).

Vectors in primal programs are represented by lowercase Latin letters while matrices are uppercase. Dual variables and certain scalars are generally Greek letters. Superscripts indicate a stage while subscripts indicate components followed by realization index. Boldface indicates a random quantity. Expectations of random variables are indicated by a bar ($\bar{\xi}$), μ , or $(E(\boldsymbol{\xi}))$. We also use the bar notation to denote sample means in Chapter 10.

Equations are numbered consecutively in the text by section and number within the section (e.g., (1.2) for Section 1, Equation 2). For references to chapters other than the current one, we use three indices: chapter, section, and equation, (e.g., (3.1.2) for Chapter 3, Section 1, Equation 2). Exercises are given at the end of each section and are referenced in the same manner as equations. All other items (figures, tables, declarations, examples) are labeled consecutively through the entire chapter with a single reference (e.g., Figure 1) if within the current chapter and chapter and number if in a different chapter (e.g., Figure 3.1 for Chapter 3, Figure 1).

Symbol	Definition
$+$	Superscript indicates the positive part of a real (i.e., $a^+ = \max(a, 0)$) or unrestricted variable (e.g., $y = y^+ - y^-$, $y^+ \geq 0, y^- \geq 0$) and its objective coefficients (e.g., q^+), subscript as non-negative values in a set (e.g., \Re_+) or the right-limit ($F^+(t) = \lim_{s \downarrow t} F(s)$)
$-$	Superscript indicates the negative part of a real (i.e., $a^- = \max(-a, 0)$) or unrestricted variable (e.g., $y = y^+ - y^-$, $y^+ \geq 0, y^- \geq 0$) and its objective coefficients (e.g., q^-) or the left-limit ($F^-(t) = \lim_{s \uparrow t} F(s)$)
$*$	Indicates an optimal value or solution (e.g., x^*)
$0 \wedge \sim$	Indicate given nonoptimal values or solutions (e.g., $x^0, \hat{x}, x', \tilde{x}$)
a	Ancestor scenario, real value or vector
A	First-stage matrix (e.g., $Ax = b$), also used to indicate a subset, $A \in \mathcal{A} \subset \Omega$
\mathcal{A}	Collection of subsets
b	First-stage right-hand side (e.g., $Ax = b$)
B	Matrix, basis submatrix, Borel sets, or index set of a basis
\mathcal{B}	Collection of subsets (notably Borel sets)
c	First-stage objective ($c^T x$), t -th stage objective ($(c^t(\omega))^T x^t$) or real vectors
C	Matrix or index set of continuous variables
d	Right-hand side of a feasibility cut in the L-shaped method, a demand, or real vector
D	Left-hand side vector of a feasibility cut in the L-shaped method, a matrix, a set, or an index set of discrete variables
\mathcal{D}	Set of descendant scenarios
e	Exponential, right-hand side of an optimality cut in the L-shaped method, an extreme point, or the unit vector ($e^T = (1, \dots, 1)$)
E	Mathematical expectation operator or left-hand side vector of an optimality cut in the L-shaped method
f	Function (usually in an objective ($f(x)$ or $f_i(x)$) or a density
F	Cumulative probability distribution
g	Function (usually in constraints ($g(x)$ or $g_j(x)$))
h	Right-hand side in second-stage ($Wy = h - Tx$), also $h^t(\omega)$ in multistage problems

Symbol	Definition
H	Number of stages (horizon) in multistage problems
i	Subscript index of functions (f_i) or vector elements (x_i, x_{ij})
I	Identity matrix or index set ($i \in I$)
j	Subscript index of functions (g_j) or vector elements (y_j, y_{ij})
J	Matrix or index set
k	Index of a realization of a random vector ($k = 1, \dots, K$)
K	Feasibility sets (K_1, K_2) or total number of realizations of a discrete random vector
l	Index or a lower bound on a variable
L	The L-shaped method, objective value lower bound, or real value
m	Number of constraints (m_1, m_2) or number of elements ($i = 1, \dots, m$)
n	Number of variables (n_1, n_2) or number of elements ($i = 1, \dots, n$)
N	Set, normal cone, normal distribution, or number of random elements
O	Zero matrix
p	Probability of a random element (e.g., p_k $= P(\boldsymbol{\xi} = \xi_k)$) or matrix of probabilities
P	Probability of events (e.g., $P(\boldsymbol{\xi} \leq 0)$)
q	Second-stage objective vector ($q^T y$)
Q	Second-stage (multistage) value function with random argument ($Q(x, \boldsymbol{\xi})$ or $Q^t(x^t, \xi^t)$)
\mathcal{Q}	Second-stage (multistage) expected value value (recourse) function ($\mathcal{Q}(x)$ or $\mathcal{Q}^t(x^t)$)
r	Revenue or return in examples, real vector, or index
R	Matrix or set
\Re	Real numbers
s	Scenario or index
S	Set or matrix
t	Superscript stage or period index for multistage programs ($t = 1, \dots, H$), a real-valued parameter, or an index
T	Technology matrix ($Wy = h - Tx$ or $T^{t-1}(\omega)(x)$); as a superscript, the transpose of a matrix or vector
u	General vector, upper-bound vector, or expected shortage
U	Objective value upper bound

Symbol	Definition
v	Variable vector or expected surplus
V	Set, matrix or an operator
w	Second-stage decision vector in some examples
W	Recourse matrix ($Wy = h - Tx$)
x	First-stage decision vector or multistage decision vector (x^t)
X	First-stage feasible set ($x \in X$) or t th stage feasible set (X^t)
y	Second-stage decision vector
Y	Second-stage feasible set ($y \in Y$)
z	Objective value ($\min z = c^T x + \dots$)
Z	Integers
α	Real value, vector, or probability level with probabilistic constraints
β	Real value or vector
γ	Real value or function
δ	Real value or function
ϵ	Real value
ζ	Random variable
η	Real value or random variable
θ	Lower bound on $Q(x)$ in the L-shaped method
κ	Index
λ	Dual multiplier, parameter in a convex combination, or measure
μ	Expectation (used mostly in examples of densities) or a parameter for non-negative multiples
ν	Algorithm iteration index
ξ	Random vector (often indexed by time, ξ^t) with realizations as ξ (without boldface)
Ξ	Support of the random vector ξ
π	Dual multiplier
Π	Product or projection operator
ρ	Dual multiplier or discount factor
σ	Dual multiplier, standard deviation, or σ -field
Σ	Summation
τ	Possible right-hand side in bundles or index of time
ϕ	Function in computing the value of the stochastic solution or a measure
Φ	Function, cumulative distribution of standard normal

Symbol	Definition
\emptyset	Empty set
χ	Tender or offer from first to second period ($\chi = Tx$)
ψ	Second stage value function defined on tenders and with random argument, $\psi(\chi, \xi(\omega))$
Ψ	Expected second stage value function defined on tenders, $\Psi(\chi)$
ω	Random event ($\omega \in \Omega$)
Ω	Set of all random events