WILEY

Algorithms for Nonlinear Programming and Multiple-Objective Decisions

Berç Rustem

Algorithms for Nonlinear Programming and Multiple-Objective Decisions

Berç Rustem

Imperial College of Science, Technology and Medicine, London, UK

JOHN WILEY & SONS

Chichester · New York · Weinheim · Brisbane · Singapore · Toronto

Copyright © 1998 by John Wiley & Sons Ltd.
Baffins Lane, Chichester,
West Sussex PO19 1UD, England

National 01243 779777
International (+44) 1243 779777

e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk.
Visit our Home Page on http://www.wiley.co.uk
or http://www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London, W1P 9HE, UK, without the permission in writing of the publisher

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158–0012, USA

WILEY-VCH Verlag GmbH, Papelallee 3,
D-69469 Weinheim, Germany

Jacaranda Wiley Ltd, 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01,
Jin Xing Distripark, Singapore 129809

John Wiley & Sons (Canada) Ltd, 22 Worcester Road,
Rexdale, Ontario, M9W 1L1 Canada

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0 471 97850 7

Typeset in 10/12pt Times by Pure Tech India Ltd, Pondicherry
Printed and bound in Great Britain by Biddles, Guildford and King's Lynn
This book is printed on acid-free paper responsibly manufactured from sustainable forestation, for which at least two trees are planted for each one used for paper production.

Algorithms for Nonlinear Programming and Multiple-Objective Decisions

Wiley-Interscience Series in Systems and Optimization

Advisory Editors

Sheldon Ross

Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA

Richard Weber

Cambridge University, Engineering Department, Management Studies Group, Mill Lane, Cambridge CB2 1RX, UK

GITTINS — Multi-armed Bandit Allocation Indices

KALL/WALLACE — Stochastic Programming

KAMP/HASLER — Recursive Neural Networks for Associative Memory

KIBZUN/KAN — Stochastic Programming Problems with Probability and Quantile Functions

RUSTEM — Algorithms for Nonlinear Programming and Multiple-Objective Decisions

VAN DIJK — Queueing Networks and Product Forms: A Systems Approach

WHITTLE — Optimal Control: Basics and Beyond

WHITTLE — Risk-sensitive Optimal Control

此为试读，需要完整PDF请访问：www.ertongbook.com

Preface

This book is a study of algorithms for decision making with multiple objectives. It is addressed to researchers in computational methods for decision making and optimal design: computer scientists interested in quantitative decision support, in particular numerical optimization; engineers; mathematicians; and those working in management science, operations research, economics and finance. Although most of the mathematics required is reviewed in a series of comments and notes, located at the ends of relevant chapters, the reader is expected to have developed some insight in decision making and associated concepts.

The starting point is the nonlinear optimal decision problem for dynamic systems with multiple objectives under uncertainty. An optimal decision needs to take account of possible future uncertainties. As the process unfolds, and the uncertainty becomes known, the decision is revised and new future uncertainties are considered. This approach to optimal decisions is formulated as a static nonlinear problem, and the question of multiple objective decision making within this framework is considered using quadratic programming, nonlinear programming, nonlinear constrained min–max, mean-variance optimization and noncooperative Nash games. Regarding uncertainty in multiperiod decision problems, the treatment of scenario optimization is omitted from this book. This approach utilises the probability of the scenarios at each period to evaluate the expected value of the objective and ensure the satisfaction of the constraints arising from each scenario. It is essentially the application of optimization algorithms to very large-scale problems. The size of the problem is in particular due to the scenarios and adds a further layer of complexity, often resolved by efficient problem formulation, decomposition and parallel computers. Other areas not covered in this book are network optimization, combinatorial optimization and integer programming. The algorithms and tools for these seem to be beyond the main focus of the book.

We consider the static optimization problem with a single criterion in Chapter 1 and study the optimality conditions under equality and inequality constraints. We also describe in Chapter 1 a simple and approximate algorithm for solving the nonlinear dynamic policy optimization problem. If an approximate search strategy is required, when the multiple objective decision is being formulated, the last two sections of Chapter 1 provide an algorithm that has been tried and tested in numerous macroeconomic decision making and engineering process control problems. Nevertheless, it must be pointed out that the rest of the book is devoted to the discussion of methods for which accuracy is of primary importance.

Chapter 2 is devoted to the solution of the quadratic programming problem, encountered in Chapters 3–5, for the specification of multiple objective problems and, as a subproblem, in Chapters 6, 7, 8 and 12. Three different algorithms are considered in this chapter.

The basic view of multiple criteria is that the decision making process is a cognitive one. It is in the course of this process that the decision maker gains an increasingly concrete knowledge of what can be done and determines the trade-offs and targets. Chapters 3–5 describe iterative methods, involving interactions with the decision maker, for the specification of the relative weights and targets in multiple objective problems. In Chapter 5, the convergence properties of these algorithms are considered.

Chapters 6–8 cover nonlinear programming algorithms required for the solution of the optimization of a single objective. Convex optimization is discussed in Chapter 6, and an efficient version of the Goldstein–Levitin–Polyak algorithm is studied in detail with convergence rate results. The general nonlinear programming problem is considered in Chapter 7 with a detailed study of sequential quadratic programming algorithms. For example, considerations such as convexifying the problem in order to enlarge the region of convergence of the algorithm and augmented Lagrangians are introduced. Techniques for augmenting the Lagrangian are discussed along with stepsize strategies that measure the progress of the algorithm at every iteration. The rates of convergence of these algorithms are dependent on the nature of the approximate Hessian used. These are discussed in Chapter 8, with results concerning the rates for the variable-Lagrange multiplier pair.

A competition model in the presence of multiple decision makers is considered in Chapter 9. This is the case when each objective corresponds to an agent, or player, whose actions affect the system and thereby the objectives of other players. The aim is the computation of Nash equilibria in games. The algorithms considered are an asynchronous relaxation of the best replay algorithm, as well as variants of the Newton algorithm for solving the equilibrium condition.

The mean-versus-variance multiple objective problem is discussed in Chapters 10 and 11. This arises in decision making under uncertainty where we consider the simultaneous optimization of the expected value of the objective and its variance, which represents the associated risk. These are usually conflicting objectives. An example is the classical investment portfolio problem of maximizing expected return and minimizing expected risk. This is studied in Chapters 10 and 11. In the former, an extension of the portfolio problem is discussed, combining the risk of investments and exchange rates. In the latter, a study of the nonlinear case is given for dynamic decision problems with feedbacks.

The final approach to the multiple objective problem, discussed in Chapter 12, is the min–max formulation. The optimization of the worst-case objective requires an algorithm to solve the nonlinearly constrained min–max problem. The algorithm and its convergence rate properties are considered in detail.

The ideas presented in this book are based on experience in designing solutions to optimal decision problems in economics, finance and engineering. These were developed over a period during which I was privileged to have the opportunity to discuss, debate and argue related questions with Robin Becker, Jeremy Bray and Kumaraswamy Velupillai. Without their input, parts of the book would have been considerably weaker. On nonlinear programming, I am indebted to Laurence Dixon and David Mayne for numerous informative discussions and to Ioannis Akrotirianakis for proof reading the manuscripts of most of the related chapters. On various aspects of uncertainty, I am grateful for the comments and advice of Gregory Chow, David Kendrick and Martin Zarrop. Of course, none of the above bear responsibility for any remaining errors or misrepresentations.

The book was finished during a sabbatical year, and I am grateful to Imperial College for the sabbatical programme and to my colleagues in the Department of Computing for giving me this opportunity to finalize the project.

Notation

Sections within a chapter are referred to by their consecutive numbers; sections in other chapters are preceded by the chapter number. For example, in Chapter 1, the third section is referred to as Section 3, whereas within Chapter 2, the same section would be referred to as Section 1.3. The numbering of equations, theorems, etc. also follow the same rule. Within a chapter, an equation is referred to by the section number in which it occurs, followed by the equation or theorem number in the section. Outside the chapter, this is preceded by the chapter number. Although sometimes subsections are used to direct the discussion, these are not used in the referencing system.

SYMBOLS

$t = 1, 2, \ldots, \mathscr{T}$	index of discrete- time periods, starting at period 1 with final time period $\mathscr{T}$;
$u_t \in \mathbb{R}^u$	vector of controls, or decision variables at tth time period;
$y_t \in \mathbb{R}^y$	vector of output, or endogenous values, determined by the system, at tth time period;
$\epsilon_t \in \mathbb{R}^\epsilon$	vector of uncertainties, or random variables, at tth time period;
$\mathscr{E}(\cdot)$	expected value of $(\cdot)$;
$\mathrm{var}\,(\cdot)$	variance of $(\cdot)$;
$U \equiv [u_1^\mathrm{T}, \ldots, u_t^\mathrm{T}, \ldots, u_{\mathscr{T}}^\mathrm{T}]^\mathrm{T} \in \mathbb{R}^{u \times \mathscr{T}}$	vector of controls of all time periods;
$Y \equiv [y_1^\mathrm{T}, \ldots, y_t^\mathrm{T}, \ldots, y_{\mathscr{T}}^\mathrm{T}]^\mathrm{T} \in \mathbb{R}^{y \times \mathscr{T}}$	vector of endogenous variables of all time periods;
U^d, Y^d	desired, or bliss, values of U, Y;
$\epsilon \equiv [\epsilon_1^\mathrm{T}, \ldots, \epsilon_t^\mathrm{T}, \ldots, \epsilon_{\mathscr{T}}^\mathrm{T}]^\mathrm{T} \in \mathbb{R}^{\epsilon \times \mathscr{T}}$	vector of random variables of all time periods;
$\mathbb{R}^n$	n-dimensional real vector space;
$x \in \mathbb{R}^n$	vector of optimization variables; sometimes we denote $x = [Y^\mathrm{T} \vdots U^\mathrm{T}]^\mathrm{T}$;
x^d	desired, or bliss, value of x;
x_p	preferred value of x (Chapters 3–5);
x_c	current optimal value of x (Chapters 3–5);
x_n	new optimal value of x (Chapters 3–5);

$\bar{x}$	unconstrained Newton step at x_k (Chapter 6);
x_k^{p}	projection of $\bar{x}$ (Chapter 6);
d_k	direction of search at x_k;
τ_k	stepsize along direction of search d_k;
c_k	penalty parameter value at x_k;
η_k	barrier parameter at iteration k;
$\mathscr{L}$	binary input length of quadratic programming problem;
$\mathbb{R}^{n \times m}$	set of real matrices of dimensions $n \times m$;
$\mathbb{R}_+^m \equiv \{\eta \in \mathbb{R}^m \mid \eta \geqslant 0\}$:	nonnegative orthant;
$1 \in \mathbb{R}^n$	$[1, 1, \ldots, 1]^{\mathrm{T}}$;
$\langle x, y \rangle$	inner product $x^{\mathrm{T}}y$ for $x, y \in \mathbb{R}^n$;
$\mathbb{E}_+^m \equiv \{\alpha \in \mathbb{R}^m \mid \alpha \geqslant 0; \langle 1, \alpha \rangle = 1\}$;	
$f(x)$	scalar objective function of $x \in \mathbb{R}^n$;
$f \in \mathbb{C}^1(\mathbb{R}^n)$ or $\in \mathbb{C}^1$	function f has continuous first partial derivatives with respect to $x \in \mathbb{R}^n$;
$f \in \mathbb{C}^2$	function f has continuous second partial derivatives with respect to x;
$\nabla_x f(x)$ or $\nabla f(x)$	gradient of f with respect to x; this is a column vector with jth element $\partial f(x)/\partial x^j$:

$$\nabla f(x) \equiv \begin{bmatrix} \dfrac{\partial f(x)}{\partial x^1} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x^j} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x^n} \end{bmatrix};$$

f_k and ∇f_k	$f(x_k)$ and $\nabla f(x_k)$;
$q(x)$ or $q_k(x)$	quadratic objective function or quadratic approximation to the objective function at $x_k j$;
$\mathscr{Y}(U)$	computational mapping between Y and U arising from model of the system $g(Y, U) = 0$ and the model solution algorithm (Chapters 1 and 11);
$\mathscr{F}(x)$ or $\mathscr{F}(U)$	$f(\mathscr{Y}(U), U)$, i.e. the objective function $f(Y, U)$ reduced using the mapping $Y = \mathscr{Y}(U)$;
$\mathfrak{Q} \equiv \nabla^2 f(x)$	Hessian, or second-derivative matrix, of f with respect to x; the j^{ℓ}th element of this matrix is given by $\partial^2 f(x)/\partial x^j \, \partial x^{\ell}$:

$$\nabla^2 f(x) \equiv \begin{bmatrix} \frac{\partial^2 f(x)}{\partial (x^1)^2} & \cdots & \frac{\partial^2 f(x)}{\partial x^j \, \partial x^1} & \cdots & \frac{\partial^2 f(x)}{\partial x^n \, \partial x^1} \\ \vdots & \ddots & \vdots & \vdots & \\ \frac{\partial^2 f(x)}{\partial x^1 \, \partial x^j} & \cdots & \frac{\partial^2 f(x)}{\partial (x^j)^2} & \cdots & \frac{\partial^2 f(x)}{\partial x^n \, \partial x^j} \\ \frac{\partial^2 f(x)}{\partial x^1 \, x^n} & \cdots & \frac{\partial^2 f(x)}{\partial x^j \, \partial x^n} & \cdots & \frac{\partial^2 f(x)}{\partial (x^n)^2} \end{bmatrix};$$

$\mathfrak{Q}_u, \mathfrak{Q}_y$	weighting matrics associated with $U - U^{\mathrm{d}}$, $Y - Y^{\mathrm{d}}$
$g(x) \in \mathbb{R}^e$	vector-valued function of equality constraints (i.e. $\{x \in \mathbb{R}^n \mid g(x) = 0\}$);
g_k	$g(x_k)$;
∇g_k	$\nabla g(x_k) \equiv [\nabla g^1(x_k) \vdots \nabla g^2(x_k) \vdots \ldots \vdots \nabla g^e(x_k)] \in \mathbb{R}^{n \times e}$;
$h(x) \in \mathbb{R}^i$	vector- valued function of inequality constraints (i.e. $\{x \in \mathbb{R}^n \mid h(x) \leqslant 0\}$);
h_k	$h(x_k)$
∇h_k	$\nabla h(x_k) \equiv [\nabla h^1(x_k) \vdots \nabla h^2(x_k) \vdots \ldots \vdots \nabla h^i(x_k)] \in \mathbb{R}^{n \times i}$;
$\mathscr{I}(x_k) \equiv \{i \mid h^i(x_k) = 0\}$	set of active inequality constraints at x_k;
$\mathscr{R}$	set of feasible points, usually given by $\mathscr{R} = \{x \in \mathbb{R}^n \mid g(x) = 0; h(x) \leqslant 0\}$; in Chapter 6, where we constrain $\mathscr{R}$ to be convex, we have $\mathscr{R} = \{x \in \mathbb{R}^n \mid h(x) \leqslant 0\}$; in Chapter 3, we take this set to be defined by linear equalities only: $\mathscr{R} = \{x \in \mathbb{R}^n \mid \mathscr{G}^{\mathrm{T}} x = g\}$;
Ω	set of policies acceptable to the decision maker (Chapters 3–5);
$L(x, \lambda, \mu)$	Lagrangian function for the constrained optimization problem;
$L^{\mathrm{a}}(x, \lambda, \mu, c, \alpha)$	augmented Lagrangian function for the constrained optimization problem;
$\alpha \in \mathbb{R}^i$	vector of offsets for inequality constraints, used in Chapters 7 and 8;
$\lambda \in \mathbb{R}^e$	multiplier vector for the equality constraints;
$\mu \in \mathbb{R}^i$	multiplier vector for the inequality constraints;
trace(A)	$\sum_{i=1}^n a_{ii}$, $A \in \mathbb{R}^{n \times n}$;
diag (x)	$\mathrm{diag}[x^1, x^2, \ldots, x^n] \equiv \begin{bmatrix} x^1 & 0 & \ldots & 0 \\ 0 & x^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x^n \end{bmatrix}; x \in \mathbb{R}^n$;
$\lceil x \rceil$	smallest integer ι such that $\iota \geqslant x$;

此为试读，需要完整PDF请访问：www.ertongbook.com

□	end of a proof, an example or a particular train of thought;
policy maker	decision maker.

Contents

此为试读，需要完整PDF请访问：www.ertongbook.com