INTRODUCTION TO OPTIMIZATION THEORY

Byron S. Gottfried
Joel Weisman

INTRODUCTION TO OPTIMIZATION THEORY

BYRON S. GOTTFRIED

Department of Industrial Engineering,
Systems Management Engineering and Operations Research
University of Pittsburgh

JOEL WEISMAN

Department of Chemical and Nuclear Engineering
University of Cincinnati

PRENTICE-HALL, INC., *Englewood Cliffs, New Jersey*

Library of Congress Cataloging in Publication Data

GOTTFRIED, BYRON S.
Introduction to optimization theory.

(Prentice-Hall international series in industrial and systems engineering)
Includes bibliographical references.
1. Mathematical optimization. 2. Programming (Mathematics) I. Weisman, Joel, joint author. II. Title.
QA402.5.G68 519.7 72-8519
ISBN 0-13-491472-4

Printed in the United States of Ame

10 9 8 7 6

PRENTICE-HALL INTERNATIONAL, INC., *Lon*
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., *Syd*
PRENTICE-HALL OF CANADA, LTD., *Torc*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New D*
PRENTICE-HALL OF JAPAN, INC., *To*

PREFACE

The recent proliferation of optimization techniques and the widespread availability of digital computers have encouraged most universities to offer a collection of courses in optimization theory, with applications to various disciplinary areas. These courses have been received with considerable enthusiasm by students, faculty, and, in many cases, by postgraduate professional people. The pioneer courses of this nature have been difficult to develop, however, owing to the lack of suitable text material which would bring the entire field into a proper perspective for introductory classroom instruction. We have attempted to write such a text, intending it for either the beginning graduate level or, with some modification, the advanced undergraduate level. Our intent is to present the basic ideas of each of the major classes of optimization methods, offering at the same time a basis for unification and some essence of comparison among them.

It was necessary for us to decide whether to orient the text toward certain disciplinary areas, or to write a more general textbook which a competent instructor could supplement with specialized comprehensive problems reflecting his particular area of interest. We have chosen the latter. In using early versions of this manuscript to teach senior- and graduate-level engineering optimization courses at Carnegie-Mellon University, the University of

Pittsburgh, and the University of Cincinnati, we found it very effective supplement the text material with a few carefully chosen disciplinary pr lems requiring a computer solution. Thus our students received an exposı to some analytical model building in their own area of expertise. We stron recommend that others consider this approach, particularly if a few gene purpose computer codes (e.g., linear programming, hill-climbing, etc.) a available.

This text contains more than enough material for a two-semester (30-wee course at the introductory graduate level. The text has been organized so th linear programming is presented as an integrated part of optimization theo Before linear programming is discussed (Chapter 4), optimization terminolo (Chapter 1); the classical Calculus (Chapter 2); Kuhn-Tucker conditio (Chapter 2); and unconstrained optimization (Chapter 3) are presented. Tl sequence allows linear programming to be related to other optimizati techniques as well as allowing duality theory to be derived on the basis of t Kuhn-Tucker conditions.

Chapters 1 through 4, plus Appendix B, can be used for the first semest course. The remaining chapters offer considerable latitude for the secoı semester. These chapters present nonlinear programming (Chapter 5); int ger programming and the method of decomposition (Chapter 6); optimizati of functionals (Chapter 7); dynamic programming and the discrete maximu principle (Chapter 8); and, finally, optimization under risk and uncertain (Chapter 9). If the curriculum calls for a first semester course more heavi oriented towards linear programming, this can be provided by Chapters 1, 4, 6, and Appendix B. A one-semester course in nonlinear programmir can be structured from Chapters 3, 5, 6, and a part of Chapter 9. Also, Cha ters 7 and 8, and a portion of Chapter 9, can be used for a one-semest course in dynamic systems optimization.

Each chapter is supplemented with an extensive set of problems fc student solution. The problems are designed to illustrate and, in some case extend the text material. In a number of problems, some seemingly quit simple, the student will find it difficult to arrive at a numerical solution b hand. These problems are meant to illustrate the real complexity of mos optimization tasks. The student should be encouraged to set up such prob lems for solution but to carry the numerical solution procedure only as fa as seems reasonable. When the class composition is such that developmen of special problems in a given discipline is not feasible, it is suggested that few complex problems be selected from those provided by the text and as signed for solution on a digital computer.

The required mathematical background for this text does not exten beyond elementary calculus and differential equations, though an introduc tory exposure to linear algebra is desirable. Appendix B contains the basi concepts of linear algebra for those who lack background in this important area.

Finally, we wish to thank several people for their assistance and encouragement during the course of this work. To R. H. Curran, who proofread much of the manuscript; D. L. Keefer, who proofread portions of the manuscript and developed some of the problems; and to M. J. Lempel, who made a number of useful suggestions, we extend our sincere appreciation. Our special thanks to Dr. A. G. Holzman who, as a teacher and a colleague, contributed to this project in numerous ways. Last, but not least, we wish to acknowledge the patience and understanding shown by our wives and children for the many hours of preoccupation, which we hope will serve a worthwhile purpose.

BYRON S. GOTTFRIED
Pittsburgh, Pennsylvania

JOEL WEISMAN
Cincinnati, Ohio

NOMENCLATURE

Vector and Scalar Notation

E, *e*	Symbols in standard type refer to scalar quantities
E	Symbols in bold face type indicate vectors or matrices
E′	Primed, bold face symbol indicates transpose of matrix
$\mathbf{E}^{-1}$	Bold face symbol to power of (−1) indicates inverse of matrix
|**E**|	Bold face symbol in | | indicates determinant of matrix within | |

Symbol Definition

Symbol	Definition
a	constant, lower bound on *x*
a_{ij}	coefficients of decision variables in constraint relationships
A	constant
A	a matrix, matrix where components are a_{ij}, vector where components represent lower bounds of **X**

b	constant, upper bound of x
b_i	constraint requirement
B	constant
$\mathbf{B}$	a column vector, vector where components are b_i, vector where components represent upper bound of $\mathbf{X}$, matrix where columns are the basis vectors $\mathbf{P}_1, \ldots \mathbf{P}_m$
c_j	coefficient of x_j in objective function
$\mathbf{C}$	vector whose components are c_j – the objective function coefficients
d_j	coefficient of x_j in Gomory constraint or cut, or in any equation defining a hyperplane
D_k	demand during period k
$\mathbf{D}$	vector whose components are d_j
$\lvert\mathbf{D}_i\rvert$	an $i \times i$ determinant whose elements are second partials of y with respect to the x_j
$E(\ \ )$	expected value of quantity within brackets
$f(x)$	frequency function of x
F_n	n'th member of Fibonacci sequence
$F(x)$	distribution function of x; $F(x) = \int_{-\infty}^{x} f(s)\, ds$, where s is a dummy variable
$g_i(\mathbf{X})$	constraint relationship
G	constraint equation
$G_i(\mathbf{X}, \mathbf{X}^\circ)$	hyperplane approximating $g_i(\mathbf{X})$ at $\mathbf{X}^\circ$
$\mathbf{G}$	constraint set, $[\mathbf{G}(\mathbf{u}, \mathbf{x}, \mathbf{u}', \mathbf{x}', t) = \mathbf{0}]$
H	Hamiltonian function $H(\mathbf{u}, \mathbf{x}, \boldsymbol{\lambda}) = \sum_{j=0}^{n} \lambda_j(t)\, g_j(\mathbf{u}, \mathbf{x})$
$H^{(\alpha)}$	stagewise Hamiltonian function $H^{(\alpha)} = \sum_{j=0}^{n} \lambda_j^{(\alpha)} T_j^{(\alpha)}$
$\mathbf{H}_i$	matrix used to determine search direction in Fletcher-Powell algorithm
$\mathbf{H}$	Hessian matrix
I	integral of function with respect to t
$\mathrm{I}(0)$	$\int_{t_0}^{t_f} \phi(u, u', t)\, dt$
$\mathrm{I}(\epsilon)$	$\int_{t_0}^{t_f} \phi(u + \epsilon\eta, u' + \epsilon\eta', t)\, dt$
$\mathbf{I}$	identity matrix
k	constant

K	positive constant, penalty factor
l	constant,
	augmented integral $= \int_{t_0}^{t_f} (\phi + \lambda G)\, dt$
L_n	interval of uncertainty after n search points
L	Lagrangian function
m	a constant,
	number of constraint equations,
	number of search points
$\mathfrak{M}(\mathbf{u}, \lambda)$	function defined by $\mathfrak{M}(\mathbf{u}, \boldsymbol{\lambda}) = \operatorname{Max}_{\mathbf{x}} H(\mathbf{u}, \mathbf{x}, \boldsymbol{\lambda})$
n	number of decision variables,
	number of terms in series
N	total number of subregions
p	fraction of points in given region
P	probability
$\mathbf{P}$	projection of vector on hyperplane
$\mathbf{P}_j$	search vector,
	vector consisting of jth column of constraint matrix coefficients
r	net worth
r_k	return function for stage k
r_{ij}	fractional part of number
R	set of points
R_j	remainder term
s	distance
$\mathbf{S}$	feasible direction vector
t	time or any quantity for which a functional relationship, $u(t)$, exists,
	standard normal variate $= (x - \mu)/\sigma$
t_i	a polynomial expression with only positive coefficients
$T_j^{(\alpha)}$	transformation function—function which transforms state variable $u_j^{(\alpha-1)}$ into state variable $u_j^{(\alpha)}$
$\mathbf{T}$	matrix whose columns consist of search vectors P_i
u	state variable
u'	first derivative of $u(t)$
u_i	simplex multiplier for row i
u_j	technological coefficient, state variable
u_{ij}	integer part of number
$u_j^{(\alpha)}$	value of j'th state variable at stage α
$u_0^{(\alpha)}$	contributions to objective function from stages 1 through α
U	utility, value
v	value of game
v_j	simplex multiplier for column j,
	fractional advance in stack prices for year j

$v(\boldsymbol{\delta})$	dual function in geometric programming
$\mathbf{V}$	projection matrix
w	dual variable, weighting factor
$\mathbf{W}$	vector of dual variables
x_j	decision variable
$\mathbf{X}$	vector of decision variables
x_{ij}	weight to be given to basis vector $\mathbf{P}_i$ in expression determining $\mathbf{P}_j$
$\mathbf{X}_o$	vector whose components are values of the decision variables at extremum
$\mathbf{X}_0$	vector whose components are values of decision variables at some arbitrarily specified starting point
$\mathbf{X}^*$	minimum feasible solution
y, $y(\mathbf{X})$	objective function
y_j	sum of the products of $c_i\ x_{ij}$ for all the x_{ij} in the jth column of the Simplex tableau
$y_k(u_k)$	optimum value of objective function for k stages $= \underset{x_k, x_{k-1} \cdots x_1}{\max(\min)} \{r_k(u_k, x_k) + r_{k-1}(u_{k-1}, x_{k-1}) + \cdots + r_1(u_1 x_1)\}$
$y'(x)$	first derivative of $y(x)$ with respect to x
$y''(x)$	second derivative of $y(x)$ with respect to x
z	Lagrangian function
α	constant, index indicating stage or subregion number
α_{ij}	weights which replace variables and/or nonlinear functions in separable programming procedure
β	scalar constant between 0 and 1
β_i	scalars determining direction of search vector $\mathbf{P}_i$
δ, δ_{ij}	quantity whose value is 0 or 1
δ_i	change in cost coefficient c_i
δ_i, δ'_{il}	dual variables of geometric programming problem
$\delta \mathrm{I}$	first variation of integral $\mathrm{I}(\epsilon)$ $= \lim_{\epsilon \to 0} \left(\frac{\mathrm{I}(\epsilon) - \mathrm{I}(0)}{\epsilon}\right)$
Δ_i	initial pattern search vector
$\Delta^+ x_j$	magnitude of positive change in x_j
$\Delta^- x_j$	magnitude of negative change in x_j
ϵ	small increment
$\eta(t)$	arbitrary function with continuous second derivatives and vanishing boundary conditions
η'	first derivative of $\eta(t)$

θ	scalar constant; in some cases a scalar restricted to be between 0 and 1
λ	Lagrange multiplier
$\lambda_j^{(\alpha)}$	adjoint variable for state variable u_j at stage α
$\boldsymbol{\lambda}$	vector of Lagrange multipliers
$\boldsymbol{\lambda}_o$	vector whose coordinates are values of Lagrange multipliers at extremum
μ	scalar weighting factor, mean value
μ_1	weight, in geometric programming problem for constraint 1
μ_{ij}	weights applied to extremal point x_{ij}
ξ_j	arbitrary point
π	geometrical constant, failure cost sum
π_k, $\hat{\pi}_k$	dual variables in decomposition algorithm
π_i	simplex multipliers or shadow prices
ρ_j	change in j'th constraint requirement
σ	standard deviation (σ^2 = variance)
τ	time
ϕ	function of u, u' and t
$\phi(\mathbf{X})$	function of $\mathbf{X}$, augmented objective function
ψ	differential or integral constraint function
∇y	gradient of y [vector whose components are $\partial y/\partial x_j$]

CONTENTS

INTRODUCTION TO OPTIMIZATION THEORY

INTRODUCTION

All of us must make many decisions in the course of our day-to-day events in order to accomplish certain tasks. Usually there are several, perhaps many, possible ways to accomplish these tasks, although some choices will generally be better than others. Consciously or unconsciously, we must therefore decide upon the *best*—or *optimal*—way to realize our objectives.

For example, all of us at one time or another find it necessary to drive through city traffic. We could attempt to find the shortest possible route from point A to point B without concern for the time required to traverse this route. Alternatively, we could seek out the quickest, though not necessarily the shortest, route between A and B. As a compromise, we might attempt to find the shortest path from A to B subject to the *auxiliary condition* that the transit time not exceed some prescribed value. Here we have examples of three similar, but different, optimization problems.

The stock market is another example of a fascinating, if not always successful, endeavor to form an optimal strategy. There are several objectives from which to choose when "playing the market," such as maximum rate of growth of capital, maximum rate of return from a fixed amount of capital, minimum chance of loss of capital, and so on. Thus one must first formulate carefully an appropriate objective, and then develop some strategy which