

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

© 1973 by Prentice-Hall, Inc.
Englewood Cliffs, New Jersey

Printed in the United States of America

All rights reserved. No part of this book
may be reproduced in any form or by any means
without permission in writing from the publisher.

10 9 8 7 6 5 4 3 2 1

PRENTICE-HALL INTERNATIONAL, INC., London
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., Sydney
PRENTICE-HALL OF CANADA, LTD., Toronto
PRENTICE-HALL OF INDIA PRIVATE LIMITED, New Delhi
PRENTICE-HALL OF JAPAN, INC., Tokyo

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 problems requiring a computer solution. Thus our students received an exposure to some analytical model building in their own area of expertise. We strongly recommend that others consider this approach, particularly if a few general purpose computer codes (e.g., linear programming, hill-climbing, etc.) are available.

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

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

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

The required mathematical background for this text does not extend beyond elementary calculus and differential equations, though an introductory exposure to linear algebra is desirable. Appendix B contains the basic

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
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
--------	------------

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 $\mathbf{P}_1, \dots, \mathbf{P}_m$
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
$ \mathbf{D}_i $	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 ; $F(x) = \int_{-\infty}^x f(s) ds, \text{ where } s \text{ is a dummy variable}$
$g_i(\mathbf{X})$	constraint relationship
G	constraint equation
$G_i(\mathbf{X}, \mathbf{X}^0)$	hyperplane approximating $g_i(\mathbf{X})$ at \mathbf{X}^0
\mathbf{G}	constraint set, $[\mathbf{G}(\mathbf{u}, \mathbf{x}, \mathbf{u}', \mathbf{x}', t) = 0]$
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
\mathbf{I}	integral of function with respect to t
$\mathbf{I}(0)$	$\int_{t_0}^{t_f} \phi(\mathbf{u}, \mathbf{u}', t) dt$
$\mathbf{I}(\epsilon)$	$\int_{t_0}^{t_f} \phi(\mathbf{u} + \epsilon \boldsymbol{\eta}, \mathbf{u}' + \epsilon \boldsymbol{\eta}', t) dt$
\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
$\mathcal{M}(\mathbf{u}, \lambda)$	function defined by $\mathcal{M}(\mathbf{u}, \lambda) = \text{Max}_{\mathbf{x}} H(\mathbf{u}, \mathbf{x}, \lambda)$
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 j th 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 \mathbf{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 stock prices for year j

$v(\delta)$	dual function in geometric programming
V	projection matrix
w	dual variable, weighting factor
W	vector of dual variables
x_j	decision variable
X	vector of decision variables
x_{ij}	weight to be given to basis vector P_i in expression determining P_j
X_0	vector whose components are values of the decision variables at extremum
X_0	vector whose components are values of decision variables at some arbitrarily specified starting point
X^*	minimum feasible solution
$y, y(X)$	objective function
y_j	sum of the products of $c_i x_{ij}$ for all the x_{ij} in the j th column of the Simplex tableau
$y_k(u_k)$	optimum value of objective function for k stages $= \max(\min) \{r_k(u_k, x_k) + r_{k-1}(u_{k-1}, x_{k-1}) + \cdots + r_1(u_1, x_1)\}$ $x_k, x_{k-1} \cdots x_1$
$y'(x)$	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 P_i
δ, δ_{ij}	quantity whose value is 0 or 1
δ_i	change in cost coefficient c_i
δ_i, δ'_{ii}	dual variables of geometric programming problem
δI	first variation of integral $I(\epsilon)$ $= \lim_{\epsilon \rightarrow 0} \left(\frac{I(\epsilon) - I(0)}{\epsilon} \right)$
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
λ	vector of Lagrange multipliers
λ_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
$\pi_k, \hat{\pi}_k$	dual variables in decomposition algorithm
π_i	simplex multipliers or shadow prices
ρ_j	change in j 'th constraint requirement
σ	standard deviation ($\sigma^2 = \text{variance}$)
τ	time
ϕ	function of u, u' and t
$\phi(X)$	function of X , augmented objective function
ψ	differential or integral constraint function
∇_y	gradient of y [vector whose components are $\partial y / \partial x_j$]

CONTENTS

Preface xi

Nomenclature xiv

1 Introduction 1

1.1 Optimization Present and Past	4
1.2 Systems and Mathematical Models	5
1.3 Variables and Objective Functions	6
1.4 Multifactor Objectives	7
1.5 Stationary Values and Extrema	9
1.6 Relative and Absolute Extrema	9
1.7 Equivalence of Minimum and Maximum	10
1.8 Convex, Concave and Unimodal Functions	10
1.9 Constraints	14
1.10 Mathematical Programming Problems	19
1.11 Dimensionality	20
1.12 Simultaneous and Sequential Optimization	20
References	21
Problems	22

2	Optimization Fundamentals	26
2.1	One-dimensional Functions	28
2.2	Multidimensional Functions	32
2.3	Optimization with Constraints—Lagrange Multipliers	43
2.4	Behavior of the Lagrangian Function	47
2.5	Inequality Constraints and Lagrange Multipliers	49
2.6	Necessary and Sufficient Conditions for an Equality Constrained Optimum with Bounded Independent Variables	56
2.7	Necessary and Sufficient Conditions for an Inequality Constrained Optimum	57
2.8	Applicability of Classical Methods to Practical Optimization Problems	62
	References	62
	Problems	63
3	Unconstrained Optimization	66
3.1	One-dimensional Optimization by Search Methods	68
3.2	Multi-dimensional Optimization by Gradient Methods	84
3.3	Multi-dimensional Optimization by Direct-search Techniques	112
3.4	Multi-dimensional Optimization by Random Search Techniques	120
3.5	Concluding Remarks	129
	References	130
	Problems	131
4	Linear Programming	136
4.1	Statement of the Linear Programming Problem	138
4.2	Characteristics of Linear Objective Functions and Linear Inequality Constraints	142
4.3	Properties of Convex Sets	147
4.4	The Simplex Algorithm	154
4.5	Selecting the Initial Basis	165
4.6	Degeneracy	170
4.7	The Revised Simplex Algorithm	173
4.8	Post-optimal Analysis	175
4.9	Duality	186
4.10	Unrestricted Dual Variables	192
	References	195
	Problems	195
5	Nonlinear Programming	206
5.1	Convex and Concave Programming	207
5.2	Linearization	211
5.3	Methods of Feasible Directions	237
5.4	Penalty-function Techniques	250
5.5	Geometric Programming	265
5.6	Comparison of Nonlinear Programming Techniques	282
	References	285
	Problems	286

6	Mathematical Programming Extensions	294
6.1	The Origin of Integer Programming Problems	295
6.2	Integer Programming Problems Solvable by Normal Linear Programming Techniques	301
6.3	Cutting-plane Methods for Solution of Integer Linear Programming Problems	309
6.4	Branch-and-bound Methods	320
6.5	Decomposition of Large Programming Problems	327
6.6	Other Applications of the Decomposition Principle	341
6.7	Conclusion	347
	References	348
	Problems	349
7	Optimization of Functionals	356
7.1	Necessary Conditions: One Unknown Function and One Independent Variable	357
7.2	Necessary Conditions: Problems of Higher Dimensionality	364
7.3	Integral Constraints	366
7.4	Differential Constraints	371
7.5	The Continuous Maximum Principle	380
7.6	Variants of the Problem	386
7.7	Behavior of the Hamiltonian Function	392
7.8	Generalizations of the Basic Problem	396
7.9	Direct Methods for Optimization of Functionals	406
	References	407
	Problems	408
8	Optimization of Staged Systems	412
8.1	A Minimum-path Problem	413
8.2	Dynamic Programming Fundamentals	418
8.3	Generalized Mathematical Formulation	422
8.4	The Principle of Optimality	428
8.5	Discrete Decision Models—An Optimal Automobile Replacement Policy	429
8.6	Nonlinear Continuous Models	440
8.7	Dynamic Programming and the Optimization of Functionals	445
8.8	Dynamic Programming and Variational Calculus	456
8.9	The Discrete Maximum Principle	458
8.10	Comparison of the Methods	473
	References	474
	Problems	475
9	Optimization Under Uncertainty and Risk	482
9.1	Optimization Under Uncertainty	484
9.2	Probability and Risk	493
9.3	Risk Elements Only in the Objective Function	502
9.4	Chance-constrained Programming	504

9.5	Risk Minimization by Failure Penalties	510
9.6	Risk Programming	512
9.7	Sequential Decisions Under Risk	517
9.8	Status of Optimization Under Risk and Uncertainty	526
9.9	Quantitative Decision Making	527
	References	528
	Problems	529
Appendices		535
A	Guide To Selection of Optimization Techniques	535
B	Elements of Matrix Algebra	538
	<i>Matrices</i>	538
	<i>Vectors</i>	539
	<i>Special Matrices</i>	540
	<i>Matrix Addition and Multiplication</i>	542
	<i>Orthogonal Vector Sets</i>	544
	<i>Determinants</i>	545
	<i>Inverse Matrices</i>	547
	<i>Simultaneous Equations: Gauss-Jordan Elimination</i>	548
	<i>Linear Independence and Bases</i>	551
	<i>Solution of m Linear Equations in n Unknowns</i>	555
	<i>Quadratic Forms</i>	557
	<i>Problems</i>	559
Index		562

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